Society for Molecular Biology & Evolution

SMBE 2024

PUERTO VALLARTA

Scientific Program

Contents

Welcome Message

Welcome Message

Organizing Committee

Delegate Information

Registration

Internet Access

Social Program

Plenary Lecture and Welcome Cocktail

Gala Dinner

Code of Conducte

Harassment, Discrimination and Liability

SMBE code of conduct and SMBE Allies

Broadcasting Policy

Guideline for Presenters

Program Overview

Keynote Speakers

Dr Alicia Mastretta-Yanes

Dr Aida Andrés

Scientific Program

Floor Map

Ground Floor Distribution

Upper Level Distribution

Online attendees guidelines

Welcome Message

Dear Delegates

¡Bienvenidos! We warmly welcome you to the 2024 Annual Meeting of the Society for Molecular Biology and Evolution (SMBE 2024) in Puerto Vallarta, Mexico. We are thrilled to have you here in our beautiful country.

It has been a long journey since 2018 when we dreamed of hosting this world-class meeting in Mexico and now it is a reality. This is the first time the Annual SMBE meeting is hosted in mainland Latin America and in a developing/middle-income country. It is a significant gesture from the Society to hold the meeting outside the Global North, highlighting the long-term goal of becoming a more inclusive and global Society. We are honored to have received such trust from the Society and you, our guest, and we hope the meeting meets or even exceeds your scientific and personal expectations.

Mexico is one of the most megadiverse countries in the world and, therefore, has a long history of research in ecology, evolution, molecular biology, and genomics. In the last two decades, Mexico has seen the birth of at least three new research centers with complementary strengths: the National Laboratory of Genomics for Biodiversity (LANGEBIO, Cinvestav), the Center for Genomic Sciences Research (CCG, UNAM), and the International Laboratory fo Human Genome Research (LIIGH, UNAM). Our local organizing committee combines members from these different institutions and our joint interests cover a wide range of bleeding-edge topics in Evolutionary Biology which we have reflected in a rich scientific program with an unprecedented number of 39 symposia. This includes Phylogenetics, Evolutionary and Population Genetics, Paleogenomics, Conservation, Crop Evolution, Machine Learning, Education, Medical Genomics, and more. We hope you'll get to discuss lots of science, interact with as many delegates as possible and strengthen the Society's mission.

We trust this year's venue, Puerto Vallarta's International Convention Center, will be an excellent place for our cherished Annual Meeting and that you'll get the opportunity to enjoy some of the outstanding natural, entertainment, culinary and cultural experiences Puerto Vallarta has to offer, either the natural charm of its beaches embraced by the Sierra Madre mountain range; the essence of its culture, its gastronomy and traditions; and the warmth of their hospitality of its people. It is a magical destination with an authentic Mexican spirit.

We have strived to provide excellence in your travel and conference experience, including childcare, travel, accommodation and transport to and from the venue from key points across Puerto Vallarta. You can find more about these on the following pages, or you can approach any of us and our staff during the meeting

We want to thank Artefacto Eventos, our Professional Congress Organizer for embarking in this effort and making this event a success. We also want to thank all the sponsors supporting this meeting. We strongly encourage you to visit their booths, as they are a key part of what has made this meeting possible.

We hope our efforts translate into memorable scientific and personal experiences and connections that will last well beyond the end of this meeting. Thank you for attending SMBE 2024, and we hope you enjoy your stay in Mexico

On behalf of the Local Organizing Committee - ¡Gracias!

Selene Fernandez-Valverde and Andrés Moreno-Estrada

Local Organizing Committee

Andrés Moreno Estrada

Advanced Genomics Unit, Center for Research and Advanced Studies (CINVESTAV), Mexico.

María Avila Arcos

International Laboratory for Human Genome Research (LIIGH), National Autonomous University of Mexico (UNAM), Mexico.

Valeria Souza

Institute of Ecology, National Autonomous University of Mexico (UNAM), Mexico.

Diego Ortega Del Vecchyo

International Laboratory for Human Genome Research (LIIGH), National Autonomous University of Mexico (UNAM), Mexico.

Federico Sánchez Quinto

International Laboratory for Human Genome Research (LIIGH), National Autonomous University of Mexico (UNAM), Mexico.

Mashaal Sohail

Center for Genomic Sciences (CCG), National Autonomous University of Mexico (UNAM), Mexico.

Diego Cortez Quezada

Center for Genomic Sciences (CCG), National Autonomous University of Mexico (UNAM), Mexico.

Selene Fernández Valverde

UNSW RNA Institute / School of Biotechnology & Biomolecular Sciences, UNSW Sydney, Australia.

Karla Sandoval Mendoza

Equity and Gender Office of the Center for Research and Advanced Studies (CODIGO-C), CINVESTAV, Mexico.

LOC Event Planner Coordinator

Karla Benítez Castrejón

Center for Research and Advanced Studies (CINVESTAV), Mexico.

Professional Conference Organizer (PCO) Coordinator

Annette Aguilar Bustamante

ARTEFACTO CONSULTORES, Mexico.

Professional Conference Organizer (PCO) Assistant

Santiago Iker Aguilar

ARTEFACTO EVENTOS, Mexico.

Delegate Information

1. Registration

Date: Sunday, 7 July 2024.

Time: From 14:00 to 18:00

Venue: Convention Center, Puerto Vallarta, México.

SMBE 2024 registration includes:

Registration fees include the welcome cocktail, lunches, coffee breaks and poster session refreshments.

2. Internet Access

SSID: SMBE2024

Password: holaammigo

Social Program

Plenary Lecture and Welcome Cocktail

Date: Sunday, 7 July 2024.

Time: 18:00

Venue: Convention Center, Puerto Vallarta, México.

Gala Dinner

Date: Wednessday, 10 July 2024.

Time: 19:00

Venue: Paradise Village Hotel, Nuevo Nayarit, México.

Code of Conducte

Harassment, Discrimination and Liability

SMBE 2024 attendees are expected to adhere to SMBE's Code of Conduct. SMBE is dedicated to providing a safe, hospitable, and productive environment for all members. Accordingly, SMBE prohibits all forms of discrimination and harassment (including among members, volunteers, and staff) in all Society activities, including events (online or in person), journal correspondence, social media, etc). Behaviour that undermines the integrity of intellectual discourse is an impediment to scientific progress and will not be tolerated. Unacceptable treatment of others includes (but is not limited to) harassment, stalking, belittling, discrimination, bullying/intimidation, assault, trolling, and retaliation (see above link for descriptions). Any misconduct on the basis of gender, gender identity, sexual orientation, age, disability, physical appearance, race, religion, national origin, ethnicity, or other descriptive categories will not be tolerated. If you need to report an incident, please contact our Safety Officer, Paula Brantner at pbworksolutions@gmail.com

SMBE code of conduct and SMBE Allies

At the next SMBE 2024 meeting in Puerto Vallarta, Mexico, you will notice some attendees wearing a ribbon indicating their status as SMBE Allies. These allies are trained to provide support to those who are victims or witnesses of inappropriate behavior that may violate the SMBE Code of Conduct. Their role is to listen, help affected individuals deal with the situation and provide them with information about their options, including the possibility of filing complaints. However, the allies are not involved in investigations or the imposition of sanctions. In addition, they are committed to creating safe spaces by acting as active observers during the event.

Broadcasting Policy

The SMBE Annual Meeting supports the communication and discussion of science. Information presented at the Meeting (in oral or poster format) may be reported and discussed by attendees and science writers via blogs, social media, or other formats, unless any of the authors requests otherwise. We do request that communications are respectful and do not directly reproduce visual materials (e.g., no posting of photos of slides or posters) unless permission is obtained from the presenter or if they have already made this information freely available in an open-source forum. If a presenter does not want information from his/her presentation to be photographed at all, or broadcast, they should make this clear in their talk/poster and we ask that attendees respect this. If you have questions or concerns about this policy, or would like to report an abuse of it, please contact any member of the Local Organizing Committee or email Paula Brantner at pbworksolutions@gmail.com

Guideline for Presenters

Dear speaker,

Hope this email finds you well, our meeting is just around the corner so we would like to share with you the following instructions and recommendations.

-Presentation size 16:9 (1920 x 1080 px)
NOT 4:3 (720 x 480 px)

It is recommended the font size to be at least 22 points in Arial font. When used for titles, the font size should be larger than that of the body text.

Classic Sans Serif fonts such as Helvetica, Garamond, Futura, Gill Sans, and Rockwell are some of the best options for body text in presentations. Helvetica and Futura are legible in most font sizes.

When the presentation includes one or more videos, they must be embedded within the presentation. NOT in a folder outside the presentation and NOT as an internet link.

Recommended poster size: 0.70 x 1 m (width x height)
Dimension limits: 0.80 x 1.20 m (width x height)

If any doubts or concerns may rise, please do not hesitate to contact us, it will be our pleasure to help out. We are really excited for your valuable participation in this distinguished scientific event! See you soon!

Kind regards,

Program Overview

SMBE 2024 PUERTO VALLARTA

SCIENTIFIC PROGRAM OVERVIEW

Day 1 — Sunday 7 July

8:00–19:00	SMBE Council Meeting (Council members only)
11:00–16:00	Workshop On site "Introduction to programmatic access to EnsEMBL" (Hall: Costa Alegre)
12:00–14:00	Workshop On site "Introdution to STDPOPSIM workshop" (Hall: Boca de Tomates)
14:00–16:00	Workshop On site "GENHIST kickoff at SMBE" (Hall: Boca de Tomates)
14:00–18:00	Registration
18:00–19:00	Plenary Hall (Los Tules) – Opening Keynote Speaker: Stephen Wright
19:00–20:00	Welcome Cocktail

Day 2 — Monday 8 July

9:00–11:00	Plenary Hall (Los Tules) – Graduate Student Excellence Awards
11:00– 11:30	COFFEE BREAK

HALL NAME	Los Tules (Plenary)	Boca de Tomates	Olas Altas	Las Pilitas	Costa Alegre (upper floor)
11:30 – 13:20 110 min.	Associate Editors' Symposium	Open Symposum 1	Not just Ne Ne–more: New applications for SMC from ecology to phylogenies	Clustering of human cohorts beyond race and ancestry: Towards relational thinking in genomics	Greener and Sustainable Computing in Molecular Evolution: Methods, Algorithms, Tools, and Protocols

13:20– 14:20	LUNCH BREAK
14:20 – 15:50	Poster session 1 Boca de Tomates: MBE editors meeting

15:50 –18:10 140 min.	human evolution in the genomic era	Exploring the Evolutionary Effects of Admixture	Unlocking the hidden dimensions of genomic diversity within species	Structural phylogenetics: investigating deep evolutionary history using protein structure	Progress and challenges for understanding the molecular evolution of sex chromosomes across Eukarya

20:00–22:00	Paradise Village Hotel:MBE Editors dinner

Day 3 — Tuesday 9 July

9:00–10:00	Plenary Hall (Los Tules) – Keynote Speaker: Alicia Mastretta–Yanes
10:00– 10:30	COFFEE BREAK

HALL NAME	Los Tules (Plenary)	Boca de Tomates	Olas Altas	Las Pilitas	Costa Alegre (upper floor)
10:30 – 12:20 110 min.	Exploring the Frontiers of Single–Cell Biology in Diverse Organisms	IDEA symposium	Aging from a multidisciplinary overview: evolution, longevity and biomedicine	Haplotype–based methods and frameworks for inference of evolutionary history	Everything that is old becomes new: comparative genomics and museum specimens

12:20– 13:20	LUNCH BREAK Boca de Tomates: SMBE Business meeting
13:20 – 14:50	Poster session 2 Boca de Tomates: GBE editors meeting

14:50 –17:10 140 min.	Evolutionary medical genomics	Paleogenomics and human evolutionary history: new insights and novel methods	Pushing the frontiers of conservation genomics	Molecular evolution through metagenomics	Transposable elements in the population genomics era: towards a better understanding of their contribution to evolution
17:10 – 17:40	COFFEE BREAK				
17:40–19:30 110 min.	One Health and microbial evolution: New ideas and perspectives	Going local: Using engaged research practices to understand regional–scale interactions	Impact of environmental changes on agrobiodiversity and strategies for resilience	Animal paleogenomics beyond higher latitudes	Genetics, Molecular Biology, and the Future of Forensic Science
20:00–22:00	Paradise Village Hotel: GBE Editors dinner				

Day 4 — Wednesday 10 July

9:00–10:00	Plenary Hall (Los Tules) – Keynote Speaker: Aida Andrés
10:00–10:30	COFFEE BREAK

HALL NAME	Los Tules (Plenary)	Boca de Tomates	Olas Altas	Las Pilitas	Costa Alegre (upper floor)
10:30 – 12:20 110 min.	Novel approaches to study plant domestication: disentangling the complex evolutionary history of crops	Genomes and Generations: Mapping Health and Heritage Through African Genomics	Advances in Machine Learning for Evolutionary Genomics	Unveiling the evolutionary history of pathogens through paleogenomics	Science in the Spotlight: Empowering Education and Public Engagement with Cutting–Edge Science in Molecular Evolution

12:20– 13:20	LUNCH BREAK Boca de Tomates: MBE & GBE 'Q&A session'
13:20 – 14:50	Poster session 3

14:50 –17:10 140 min.	Evolution of microbial communities: is the sum of parts greater than the whole?	Human population demography and adaptation signals in the Americas	Spatial population genetics: where are we now?	Mitochondria: from powerhouse to processor and from marker to meaning	Deciphering the functional and adaptive effects of genomic structural variation
19:00–23:00	GALA DINNER at Paradise Village Hotel				

Day 5 — Thursday 11 July

HALL NAME	Los Tules (Plenary)	Boca de Tomates	Olas Altas	Las Pilitas	Costa Alegre (upper floor)
09:00– 10:50 110 min.	New computational approaches to estimate past demographic events and natural selection	Open Symposium 2	Epigenetic inheritance: from molecular mechanisms to evolutionary consequences	Decoding the past to safeguard the future	Human genetic variability in the Pangenomic era.

10:50 –11:20	COFFEE BREAK
11:20–12:20	Plenary Hall (Los Tules) – Closing Keynote Speaker: Scott Edwards
12:20–14:20	Awards Ceremony and Introduction to SMBE 2025

Full Program available at www.smbe2024.org

Keynote Speakers

Dr Alicia Mastretta-Yanes

CONAHCYT - UNAM

Alicia is a biologist from the National Autonomous University of Mexico (UNAM), and holds a PhD from the University of East Anglia, England. She currently Works as a CONAHCYT Research Fellow at Instituto de Ecología - UNAM. Her research focuses on incorporating genetic diversity into conservation using evolutionary approaches, spanning diverse ecosystems such as the Lacandona jungle, sky-islands, polluted forests, and Mexican agroecosystems. She contributed to the development of genetic diversity indicators, now adopted by the Global Biodviersity Framework, and she is actively involved in co-formulating approaches for nations to monitor and safeguard genetic diversity on a global scale.

Dr Aida Andrés

University College London, UK

Aida is Professor of Population and Evolutionary Genomics at University College London (UK). She is interested in how organisms adapt to their environment. In her work, this means analysing genomes, both modern and ancient, to infer how natural selection mediates genetic adaptations. Her group works mostly on humans, as they have an interesting history of fast colonisation of diverse habitats, and in endangered primates, where the ability to adapt to quickly changing environments is crucial for survival. Her group tackles these questions using genomic approaches to study the processes of adaptation, population genetic techniques to make inferences on the history of selected alleles, and functional information to infer the consequences in present-day phenotypes of previously adaptive alleles. The group is particularly interested in the types of natural selection that maintain diversity within populations (e.g. balancing selection) or that create differences among populations (e.g. local adaptation), as well as adaptive introgression.

Keynote Speakers

Scott V. Edwards

Harvard University, USA

Scott holds a B.A. from Harvard University (1986) and a PhD from the University of California, Berkeley (1992). Following his doctoral studies, he served as an Alfred P. Sloan Postdoctoral Fellow in Molecular Evolution at the University of Florida, Gainesville (1992-94). As a scientist, he has broad interests in the evolution of life on earth and the processes that have generated biodiversity. His research primarily utilizes birds as models to study patterns of speciation, biogeography, evolution of the genome, and the process of adaptation. This work has exposed his lab to a wide range of questions, from the evolution of immune genes and disease resistance to how best to reconstruct the tree of life.

Stephen Wright

SMBE President

University of Toronto, Canada

Stephen is the current President of the Society for Molecular Biology and Evolution (SMBE). His research focuses on understanding nucleotide polymorphism and genome evolution in natural plant populations. He explores questions such as the role of mutation vs. natural selection in genome structure evolution, how population history and mating systems affect genetic variability, and the rate and genetic basis of adaptive evolution and deleterious mutation. He uses DNA sequence polymorphism and molecular evolution analysis, large-scale genome sequence data analysis, and population genetic theory to investigate these questions. He applies these approaches in a comparative context, studying both model and non-model organisms, to understand genome diversity and structure patterns. His current projects include studying transposable element evolution, the impact of recombination and mating systems on genome evolution, the effects of demographic history and natural selection on genetic diversity, and the role of gene and genome duplication in species diversification and genome evolution.

Scientific Program

Day 1 (7 July)

14:00-18:00	**Registration**	
18:00-19:00	**Opening Keynote Speaker** Stephen Wright Adaptation and maladaptation in plant genomes.	Los Tules
19:00-20:00	**Welcome Cocktail**	

Day 2 (8 July)

9:00-11:00	**Graduate Student Excellence Awards**	Los Tules
9:00-9:13	Bárbara Sousa da Mota Investigating the ecological suicide (“ecocide”) theory in Rapa Nui with ancient DNA data	
9:13-9:26	Benjamin Moran A lethal mitonuclear incompatibility in complex I of natural hybrids	
9:26-9:39	Christopher Blake Evolutionary shift of a tipping point forestalls collapse in a microbial community.	
9:39-9:52	Jaison Sequeira The provenance of Proto-North Dravidian ancestry in the Indus valley	
9:52-10:05	María Palma Martínez Beyond Continental Groups: Unveiling Dynamic Human Genetic Communities with a Novel Network Analysis Pipeline	
10:05-10:18	Mariela Tenorio MAYEX is an ancient long non-coding RNA recruited for X chromosome dosage compensation in lizards	
10:18-10:31	Meaghan Marohn Revisiting the evolution of lactase persistence: insights from South Asian genomes	
10:31-10:44	Siliang Song Human sex ratio at birth has virtually no detectable genetic variation: Was Fisher wrong about the evolution of sex ratio?	
10:44-10:57	Sung-Ya Lin Rapid coevolution preserves the epigenetic establishment of telomere protection	
11:00- 11:30	**COFFEE BREAK**	
11:30 - 13:20	**Associate Editors’ Symposium**	Los Tules
11:30-11:45	Aida Ouangraoua Transcript Homology Relationships and Phylogeny Reconstruction	
11:45-12:00	Carolin Kosiol PoMo via RevBayes: Inferring Phylogenies, Disentangling GC-bias and Balancing Selection	
12:00-12:15	Jianzhi Zhang Evidence for the role of selection for reproductively advantageous alleles in human aging	
12:15-12:30	Joanna Kelley The evolution of hibernation-related shifts in gene expression across hibernating mammals	
12:30-12:45	Kirk Lohmueller The distribution of fitness effects varies with phylogeny across animals	
12:45-13:00	Maud Tenaillon Detecting barriers to gene flow from genomic patterns: application to maize and teosintes	
13:00-13:15	Adam Eyre-Walker Is there nepotism in the allocation of research grants?	
11:30 - 13:20	**Open Symposium 1**	Boca de Tomates
11:30-11:45	Ulises Rosas Size matters: revealing the genes controlling shot apical meristem gigantism in cacti	
11:45-12:00	Navya Shukla Investigating Functional Convergence Across Millions Of Years Of Evolution	
12:00-12:15	Christian Landry Compensatory mutations potentiate constructive neutral evolution by gene duplication	

Scientific Program

Day 2 (8 July)

11:30 - 13:20 — Open Symposium 1 — Boca de Tomates

12:15-12:30 **Dulce Valdivia Martínez**

Exploring the role of tRNA regulation in the transition to multicellularity of the amoeba Dictyostelium discoideum.

12:300-12:45 **Landen Gozashti**

How repeats rearrange chromosomes in deer mice

12:45-13:00 **Ying Zhen**

Firefly toxin lucibufagins evolved after the origin of bioluminescence

13:00-13:15 **Leon Hilgers**

Seven new reference-quality genomes illuminate the genomic evolution of genetic sex determination in turtles

11:30 - 13:20 — Not just Ne Ne-more: New applications for SMC from ecology to phylogenies. — Olas Altas

11:30-11:55 **Noah Rose**

Mosquito evolution and the emergence of mosquito-borne disease in the human era

11:55-12:10 **Anastasia Ignatieva**

Detecting structural variation in reconstructed genealogies

12:10-12:25 **Toby Kovacs**

Drivers of historical population declines in Australian Marsupials: are humans behind the wheel?

12:25-12:40 **Alba Nieto Heredia**

AI and demographic inference in structured populations

12:40-12:55 **Arun Durvasula**

Accurate inference of population history in the presence of background selection

12:55-13:20 **Stefan Strütt**

Joint inference of evolutionary transitions to self-fertilization and demographic history using whole-genome sequences

11:30 - 13:20 — Clustering of human cohorts beyond race and ancestry: Towards relational thinking in genomics — Las Pilitas

11:30-11:55 **Hussein Mohsen**

What is in a discrete category? Clustering genomics cohorts beyond race, ethnicity—and ancestry

11:55-12:10 **Simon Gravel**

Genetically informed representations of ancestry and ethnicity

12:10-12:25 **Luke Anderson-Trocmé**

Breaking Boundaries: Leveraging Continuous Models for Improved Population Genetics Analyses

12:25-12:40 **Benjamin Peter**

Quantitative Interpretations of Principal Component Analysis

12:40-12:55 **John Novembre**

Non-parametric representations of structure in genetic ancestry through time

12:55-13:20 **Marta Ciccarella**

Multiple periods of admixture and isolation during and after the Transatlantic Slave Trade on the island of São Tomé.

11:30 - 13:20 — Greener and Sustainable Computing in Molecular Evolution: Methods, Algorithms, Tools, and Protocols — Costa Alegre

11:30-11:55 **Beatriz Mello**

Green Computing in Molecular Evolution and Phylogenetics for the Global South

11:55-12:20 **Bui Minh**

CMAPLE: efficient phylogenetic inference in the pandemic era

12:20-12:35 **Sudhir Kumar**

MEGA 12: Advancing Green Computing and Phylogenomics

12:35-12:50 **Gavin Huttley**

A spectral step on the divide and conquer path to energy efficient phylogenetic reconstruction

Scientific Program

Day 2 (8 July)

11:30 - 13:20 **Greener and Sustainable Computing in Molecular Evolution: Methods, Algorithms, Tools, and Protocols** — Costa Alegre

12:50-13:05 Jiansi Gao

Causes and impacts of the widespread lack of topological convergence in Bayesian phylodynamic inference on large viral datasets

13:05-13:20 Li Liu

Shifting Focus from Gene-centric to Evolution-centric Analyses in Cancer Genomics

13:20- 14:20 **LUNCH BREAK**

14:20 - 15:50 **Poster session 1**

Clustering of human cohorts beyond race and ancestry: Towards relational thinking in genomics.

Exploring the Evolutionary Effects of Admixture.

Human evolution in the genomic era.

Human genetic variability in the Pangenomic era.

Not just Ne Ne-more: New applications for SMC from ecology to phylogenies.

Open Symposium 1

Progress and challenges for understanding the molecular evolution of sex chromosomes across Eukarya

Structural phylogenetics: investigating deep evolutionary history using protein structure.

Unlocking the hidden dimensions of genomic diversity within species.

14:20 - 15:50 **Event: MBE Editors Meeting** — Boca de Tomates

15:50 -18:10 **Human evolution in the genomic era.** — Los Tules

15:50-16:15 Katherine Pollard

Decoding the evolution and functions of Human Accelerated Regions with deep learning

16:15-16:30 Erin Gilbertson

Machine learning reveals the diversity of human 3D chromatin contact patterns

16:30-16:45 Ryder Easterlin

Discovering the cis-regulatory basis of archaic human-derived phenotypes

16:45-17:00 Maddy Comerford

Functional characterisation of the regulatory activity of archaic DNA in Island Southeast Asia

17:00-17:15 Samvardhini Sridharan

Worldwide patterns of diversity at the 17q21.31 locus in modern human genomes\r\n

17:15-17:30 Eva Brigos Barril

Why do we get sick? Unraveling Genetic Trade-offs between Fertility, Longevity, and Complex Diseases

17:30-17:45 Moisès Coll Macià

Archaic introgression in modern humans modelled by ancestral components

17:45-18:10 Bo Xia

On the genetic basis of tail loss evolution in humans and apes

15:50 -18:10 **Poster session 1** — Boca de Tomates

Alexande Ioannidis

Disentangling the effects and history of human admixture from the Americas to Oceania

Alexander Kirschel

Alternative genomic and biochemical pathways to red carotenoid pigments in birds

Daniel Powell

Pervasive gene flow despite strong and varied reproductive barriers in swordtails

Lucia Morales Reyes

Genomic Insights into the Recurrent Interspecific Hybridizations of Agave-Associated Saccharomyces Yeasts

Maria Fariello Rico

Detecting selection in admixed populations using haplotypic information\r\n

Scientific Program

Day 2 (8 July)

15:50 -18:10 **Poster session 1** Boca de Tomates

Andrew Vaughn

Bayesian Inference of Admixture Graphs on Native American and Arctic Populations

Aaron Pfennig

Complex neutral processes drive the evolutionary fate of Neanderthal alleles in 30,780 admixed genomes with African-like and European-like ancestry

Emilia Huerta Sanchez

Evolutionary insights from admixed genomes

15:50 -18:10 **Unlocking the hidden dimensions of genomic diversity within species.** Olas Altas

15:50-16:15 Mayra Furlan-Magaril

Chromatin 3D conformation of a silent genome: the case of nucleated erythrocytes

16:15-16:30 Pac Majic Bergara

On mutational neighbourhoods and their influence on evolutionary processes

16:30-16:45 James Phipps-Tan Sheng Yi

Beyond averages: transcriptional variability in outbred Drosophila melanogaster and its environmental and genetic dependence

16:45-17:00 Thea Rogers

Novel topological units characterise a distinct regulatory architecture underlying the evolution of complex traits in coleiod cephalopods

17:00-17:15 Christabel-Floi Bucao

Gene expression variability implies stable cis-regulation of brain genes across non-nervous organs

17:15-17:30 Justin Wilcox

Population genomics of wild avian embryos uncovers hidden genetic diversity

17:30-17:45 Gabriela Santos Rodriguez

The evolution of transcriptome complexity using Oxford Nanopore long-read sequencing technology.

17:45-18:10 Aurora Ruiz-Herrara

Plasticity of 3D chromatin folding across timescales

15:50 -18:10 **Structural phylogenetics: investigating deep evolutionary history using protein structure** Olas Altas

15:50-16:15 Martin Steinegger

Supercharged Protein Analysis in the era of AI

16:15-16:30 Hector Romero

Exploring protein evolution through natural and simulated sequences using experimental and modeled 3D structures\r\n

16:30-16:45 Stefano Pascarelli

AlphaHOG, a protein structure-based reference classification to improve orthology inference

16:45-17:00 David Moi

Foldtree. Empirical benchmarking of structural phylogenetics methods

17:00-17:25 Betul Kaçar

Early protein evolution using deep time sequence and structure reconstructions

17:25-17:40 Desiree Langer

Structural phylogenetics of the jelly roll fold sheds light on viral evolution

17:40-17:55 Claudia Alvarez Carreno

Polyphyletic insertions in the bacterial DNA-directed RNA polymerase

17:55-18:10 Charles Carter, Jr

Structural Enzymology, Phylogenetics, Differentiation and Symbolic Reflexivity at the Dawn of Biology

Scientific Program

Day 2 (8 July)

15:50 -18:10 **Structural phylogenetics: investigating deep evolutionary history using protein structure** — Olas Altas

15:50-16:15 Martin Steinegger

Supercharged Protein Analysis in the era of AI

16:15-16:30 Hector Romero

Exploring protein evolution through natural and simulated sequences using experimental and modeled 3D structures\r\n

16:30-16:45 Stefano Pascarelli

AlphaHOG, a protein structure-based reference classification to improve orthology inference

16:45-17:00 David Moi

Foldtree. Empirical benchmarking of structural phylogenetics methods

17:00-17:25 Betul Kaçar

Early protein evolution using deep time sequence and structure reconstructions

17:25-17:40 Desiree Langer

Structural phylogenetics of the jelly roll fold sheds light on viral evolution

17:40-17:55 Claudia Alvarez Carreno

Polyphyletic insertions in the bacterial DNA-directed RNA polymerase

17:55-18:10 Charles Carter, Jr

Structural Enzymology, Phylogenetics, Differentiation and Symbolic Reflexivity at the Dawn of Biology

15:50 -18:10 **Progress and challenges for understanding the molecular evolution of sex chromosomes across Eukarya** — Costa Alegre

15:50-16:15 Gabriel Marais

The Silene latifolia genome and its giant Y chromosome

16:15-16:30 Erik Sorensen

The struggle of being a sex chromosome: Why X has a lower diversity than expected in primates

16:30-16:45 Claudia Weber

Do Z chromosomes in Butterflies and Moths Show Evidence of Increased Rates of Adaptation?

16:45-17:00 Josué Barrera-Redondo

The UV sex chromosomes of brown algae act as genomic cradles for new genes that evolve de novo

17:00-17:25 Sophia MacRae Orzechowski

Shining a light on early stages of neo-sex chromosome evolution in Australian honeyeaters (Aves: Meliphagidae)

17:25-17:40 Gabrielle Coffing

Cephalopod sex determination and its ancient evolutionary origin revealed by chromosome-level assembly of the California two-spot octopus

17:40-17:55 Sarah Carey

Using phased genome assemblies to examine the evolution of heteromorphic sex chromosomes in Cannabaceae

17:55-18:10 Melissa Wilson

Alignment and Variant Calling in Reference Genomes to Improve Sex Chromosome Comparative Genomics

Day 3 (9 July)

9:00-10:00 **Keynote Speaker** — Los Tules

Alicia Mastretta-Yanes

Genetic diversity in the Anthropocene and the role of molecular studies for the future of biodiversity

Scientific Program

Day 3 (9 July)

10:30 - 12:20 — Exploring the Frontiers of Single-Cell Biology in Diverse Organisms — Los Tules

10:30-10:55 **Carlos Ortiz Ramírez**

Single-cell genomics for understanding trait evolution and rapid adaptation

10:55-11:10 **Jesús Murga-Moreno**

Quantifying adaptive evolution of the human immune cell landscape

11:10-11:25 **Roberto Arbore**

The molecular bases of red-to-yellow color variation in parrots

11:25-11:40 **Meritxell Riera**

Evolutionary impact of transcriptomic differences in X and Y carrying gametes

11:40-11:55 **Taylor Cooper**

Characterizing neuronal evolution between recently diverged species at single-cell resolution

11:55-12:10 **Ines Hellmann**

Cross-species comparison of gene regulatory networks using single cell sequencing

10:30 - 12:20 — IDEA Symposium — Boca de Tomates

10:30- 11:00 **Miriam Miyagi**

Contextual Analysis of Genetic Studies of Gender, Sex, and Sexuality

11:00- 11:30 **Rori Rohlfs**

Science Wise: From 'hidden figures' to scientific foremothers getting their flowers

11:30- 11:45 **Joanna Kelley**

Title to be confirmed

11:45-12:00 **Samantha López Clinton**

Title to be confirmed

12:00-12:20 **Discussion**

10:30 - 12:20 — IDEA Symposium — Olas Altas

10:30-10:55 **Luis Gutierrez Robledo**

Evolutionary medicine, aging and longevity

10:55-11:20 **Tamas Szekely**

Sex difference in longevity: causes and implications

11:20-11:35 **Mariangela Iannello**

Signatures of extreme longevity: a perspective from bivalve molecular evolution

11:35-11:50 **Miriam Merenciano**

Contribution of transposable elements in the sex gap longevity of different Drosophila species

11:50-12:05 **Gillian Meeks**

Accounting for common genetic variation improves chronological age prediction in African populations

12:05-12:20 **Eugenio Mancera**

Exploring the chronological lifespan spectrum of Candida species

10:30 - 12:20 — Haplotype-based methods and frameworks for inference of evolutionary history — Las Pilitas

10:30-10:55 **Pier Palamara**

Inference of recent effective population size from high and low coverage DNA data

10:55-11:20 **Katia Bougiouri**

How to paint a dog: limits of local ancestry inference in ancient genomes

11:20-11:35 **Eduardo González Orozco**

Archaic ancestry fragment calling using wavelet decomposition

Scientific Program

Day 3 (9 July)

10:30 - 12:20 **Haplotype-based methods and frameworks for inference of evolutionary history** — Las Pilitas

11:35-11:50 Amy Williams

Reconstructing parent genotypes at genotyping array accuracy using siblings and other relatives

11:50-12:05 Jazeps Medina Tretmanis

Improving Local Ancestry Inference through Neural Networks

12:05-12:20 Amy Goldberg

Ancestry-based approaches for inference of evolutionary history on short timescales

10:30 - 12:20 **Everything that is old becomes new: comparative genomics and museum specimens.** — Costa Alegre

10:30-10:55 Elinor Karlsson

Title to be confirmed

10:55-11:20 Alexandra Keene

Viral and host metagenomics of 100-year-old insect museum specimens

11:20-11:35 Eucharist Kun

The trait specific timing of accelerated genomic change in the human lineage.

11:35-11:50 Lucas Rocha Moreira

Evolutionary paths to thermal resilience in placental mammals: functional genomics insights from the Frozen Zoo Biobank

11:50-12:05 Michelle Stitze

From swampy ancestors to modern maize: tracing the allopolyploid origins of Zea

12:05-12:20 Nico Franz

Title to be confirmed

12:20- 13:20 **LUNCH BREAK**

13:20 - 14:50 **Poster session 2**

Aging from a multidisciplinary overview: evolution, longevity and biomedicine.

Animal paleogenomics beyond higher latitudes.

Everything that is old becomes new: comparative genomics and museum specimens.

Evolutionary medical genomics.

Exploring the Frontiers of Single-Cell Biology in Diverse Organisms.

Genetics, Molecular Biology, and the Future of Forensic Science.

Going local: Using engaged research practices to understand regional-scale interactions.

Haplotype-based methods and frameworks for inference of evolutionary history.

Impact of environmental changes on agrobiodiversity and strategies for resilience.

Molecular evolution through metagenomics.

One Health and microbial evolution: New ideas and perspectives.

Paleogenomics and human evolutionary history: new insights and novel methods.

Pushing the frontiers of conservation genomics.

Transposable elements in the population genomics era: towards a better understanding of their contribution to evolution.

13:20 - 14:50 **Event: GBE Editors Meeting** — Boca de Tomates

14:50 -17:10 **Evolutionary Medical Genomics** — Los Tules

14:50-15:15 Pleuni Pennings

Why aren't all E. coli resistant to antibiotics?

15:15-15:30 Imroze Khan

Rapid evolution of mutation rate and germline maintenance under strong pathogen selection: new insights from experimental evolution experiments

15:30-15:45 Cara Love

Radiation-induced stress and signatures of selection on anti-tumor immunity in Chornobyl wolves

Scientific Program

Day 3 (9 July)

14:50 -17:10 Evolutionary Medical Genomics — Los Tules

15:45-16:00 **Daniel Jordan**

Deep learning phenotype imputation to measure dominance of rare variant effects on rare disease in UK Biobank

16:00-16:15 **Alief Moulana**

Epistasis and pleiotropy constrain and expand the evolution of SARS-CoV-2 Omicron lineage

16:15-16:30 **Helena Hopson**

Interactions between sickle cell and Plasmodium falciparum genotypes in asymptomatic malaria

16:30-16:45 **Evan Irving-Pease**

The evolutionary origins of autoimmune and infectious disease risks

16:45-17:10 **Shamil Sunyaev**

Evolutionary perspective on human polygenic traits

14:50 -17:10 Paleogenomics and human evolutionary history: new insights and novel methods. — Boca de Tomates

14:50-15:15 **Kay Prüfer**

Ancient Genomes of Neandertals and Modern Humans from Europe

15:15-15:30 **Marco Rosario Capodiferro**

Archaic ancestry inference in imputed ancient human genomes

15:30-15:45 **Yulin Zhang**

Reconstructing mutation patterns over the course of human evolution

15:45-16:00 **Stéphane Peyrégne**

A high-quality genome from a 200,000-year-old Denisovan

16:00-16:15 **Maanasa Raghavan**

Human genetic histories at the Himalayan frontiers

16:15-16:30 **Dilek Koptekin**

Overcoming bias and postmortem damage to improve the accuracy of ancient genome analysis

16:30-16:45 **Sandra Oliveira**

Tracing admixture in European early farmers using local ancestry inference

16:45-17:10 **Hugh McColl**

Steppe Ancestry in western Eurasia and the spread of the Germanic Languages

14:50 -17:10 Pushing the frontiers of conservation genomics. — Olas Altas

14:50-15:15 **Carolyn Hogg**

Conservation Genomics: innovations, applications & where to next?

15:15-15:30 **Genís Garcia-Erill**

The genetic architecture of an adaptive phenotype conditions vulnerability and populations evolutionary response to climate change

15:30-15:45 **Jonas Aguirre Liguori**

The evolutionary genomics of population's responses to climate change

15:45-16:00 **Paulina Nuñez-Valencia**

Predicting Pathogenic Variation for conservation: a deep learning approach

16:00-16:15 **Rebekah Rogers**

Low Rates of Gene Duplication in an Endangered Freshwater Mussel

16:15-16:30 **Eugenie Charley Yen**

Epigenome-wide scans identify DNA methylation markers for monitoring sublethal thermal stress in endangered sea turtles

16:30-16:45 **Theo Phanu Serivichyaswat**

Towards genome-wide data of Arctic species obtained from snow footprints

16:45-17:10 **Eduardo Eizirik**

Conservation Genomics of Neotropical Wild Cats

Scientific Program

Day 3 (9 July)

14:50 -17:10	**Molecular evolution through metagenomics**	Las Pilitas
14:50-15:15	Gabriela Olmedo Álvarez Clues to the Evolution of Antibiotic resistance from bacteria and ancient communities from pristine sites in Cuatrocienegas, Coahuila, Mexico.	
15:15-15:30	Abigail Lind Host adaptation and genomic change in intestinal eukaryotes revealed from sequencing microbial mixtures	
15:30-15:45	Maya Lemmon-Kishi What's the Rate: Molecular Clock Calibration Using Sedimentary Ancient DNA	
15:45-16:00	Edder Daniel Bustos-Diaz Nitrogen cycling in symbiotic bacterial communities associated with ancient American plants	
16:00-16:15	Irina Velsko De novo assembly of ancient human metagenomes improves community diversity estimates and reveals varied microbial taxon-specific evolutionary trajectories	
16:15-16:30	Mariana Guerrero-Osornio Time-lapse of the microbial composition during agave fermentation using Meta-HiC	
16:30-16:45	Davide Bozzi Genomic characterization of novel Treponema species from the oral microbiome of Aboriginal Australians	
16:45-17:10	Nandita Garud Pervasive selective sweeps within and across human gut microbiomes	
9:00-10:00	**COFFEE BREAK**	
14:50 -17:10	**Transposable elements in the population genomics era: towards a better understanding of their contribution to evolution.**	Costa Alegre
14:50-15:15	Grace Yuh Chwen Lee How transposable element shapes genome evolution through epigenetic mechanisms?	
15:15-15:30	Leandro Quadrana Transposon-Mediated Environmental Regulation: Catalyst for Adaptive Responses?	
15:30-15:45	Marie Lebherz DNA Transposons favour de novo transcript emergence through enrichment of transcription factor binding motifs	
15:45-16:00	Riccardo Pianezza Double trouble: two retrotransposons triggered a cascade of horizontal transfers in Drosophila species within the last 50 years	
16:00-16:15	Jilong Ma Independent transitions to sociality in Stegodyphus spiders are associated with transposable element expansions	
16:15-16:30	Austin Daigle TEforest: Enhancing short-read transposable element detection with machine learning	
16:30-16:45	Cei Abreu-Goodger Evolution of transposable elements as the source of extracellular RNA	
16:45-17:10	Jordana Oliveira Evolution of transposable elements in arbuscular mycorrhizal fungi: insights into genome variability and gene regulation	
17:10 - 17:40	**LUNCH BREAK**	
17:40-19:30	**One Health and microbial evolution: New ideas and perspectives**	Los Tules
17:40-17:55	Mirna Vázquez-Rosas-Landa Microbial dynamics in mangrove sediments during the anthropogenic era	
17:55-18:10	Matthew Shepherd Within-patient evolution of ciprofloxacin resistance in Pseudomonas aeruginosa across a large-scale clinical trial	

Scientific Program

Day 3 (9 July)

17:40-19:30	**One Health and microbial evolution: New ideas and perspectives**	Los Tules
18:10-18:25	Rahgavi Poopalarajah Evolution of the human gut resistome across diverse lifestyles and environments	
18:25-18:40	Manuel Ochoa-Sánchez Feather microbiota compositional and functional landscapes across phenology and biogeography in Magellanic penguins in the Magellan Strait, Chile	
18:40-19:05	Santiago Elena An evolving world for evolving pathogens: ecological drivers of virus diversification and adaptation	
19:05-19:30	Fernando González Candelas Evolutionary genomics and the one health approach to AMR	
17:40 - 19:30	**Going local: Using engaged research practices to understand regional-scale interactions.**	Boca de Tomates
17:40-18:05	María Ávila-Arcos Community engagement experiences of the Afromexico Genomics Project	
18:05-18:20	Eleni Seferidou An ancient genetic insight into pre-colonial Trinidad.	
18:20-18:35	Daniela Orozco-Perez Paleogenomic analyses of archaeological remains reveal sex roles and mobility of ancient families in central Mexico	
18:35-18:50	Esha Bandyopadhyay Distinct positions of genetic and oral histories: Perspectives from India	
18:50-19:05	Constanza de la Fuente Castro Conversations in community-based genomic research: Showcasing a platform from Chile to integrate Indigenous engagement and representation	
19:05-19:30	Laura Weyrich Uncovering the global origins of an Industrialized microbiome using regionally focused studies	
17:40 - 19:30	**Impact of environmental changes on agrobiodiversity and strategies for resilience.**	Boca de Tomates
17:40-18:05	Brandon Gaut Crop wild relatives, climate change and genomic diversity	
18:05-18:20	Luis Eguiarte Population genomics of tequila and mezcal (Agave spp.): Unexpected findings and lessons for these Mexican industries in the face of climate change.	
18:20-18:35	Mirte Bosse A boaring story of hybridization and adaptation in wild boars, domestic pigs and feral swine	
18:35-18:50	Alejandra Hernandez-Teran Disentangling microbiome effects on plant adaptation	
18:50-19:05	Ayelet Salman-Minkov Strategies for mining useful alleles for climate change adaptation	
19:05-19:30	Azalea Guerra-García Building agroecosystems resilience by fostering native crops and their wild relatives	
17:40 - 19:30	**Animal paleogenomics beyond higher latitudes.**	Las Pilitas
17:40-18:05	Viviane Slon Sedimentary ancient DNA studies in challenging preservation contexts: Insights from the southern Levant	
18:05-18:20	Alida de Flamingh Compacted hair in broken carnivore teeth reveal dietary prey of historic lions	
18:20-18:35	Gloria Gonzalez Fortes The last especimen of Eurasian lynx (Lynx lynx) from the Iberian Peninsula: new genetic data leads to review the biogeographic history of lynx matrilineal linages in western Europe.	

Scientific Program

Day 3 (9 July)

17:40 - 19:30	**Animal paleogenomics beyond higher latitudes.**	Las Pilitas
18:35-18:50	David Ledesma Small but mighty: ancient DNA analyses of late Pleistocene and Holocene microvertebrates from Central Texas	
18:50-19:05	Federico Sánchez Quinto The evolutionary origin of the Columbian mammoths from the Basin of Mexico and its implications for the Mammuthus genus	
19:05-19:30	Selina Brace Museomics and the Tropics: occasionally vexing, but a rewarding partnership	
17:40 - 19:30	**Genetics, Molecular Biology, and the Future of Forensic Science**	Costa Alegre
17:40-18:05	Rori Rohlfs Quantifying the accuracy of forensic genetic technologies	
18:05-18:20	Anne Stone The recovery of DNA from burned forensic contexts.	
18:20-18:35	Feriel Ouerghi How should we report genetic matches following an Investigative Genetic Genealogy search?	
18:35-18:50	Meghna Swayambhu Highly accurate predictions indicating sexual activity using microbiome-based analyses in forensic settings	
18:50-19:05	Maria Flores Construction of Epigenetic Clock using Cell-Free DNA	
19:05-19:30	Jaehee Kim Population genetics approaches for forensic genetic record-matching	

Day 4 (10 July)

9:00-10:00	**Keynote Speaker**	Los Tules
	Aida Andres Local genetic adaptation in humans and other primates	
10:00- 10:30	**COFFEE BREAK**	
10:30 - 12:20	**Novel approaches to study plant domestication: disentangling the complex evolutionary history of crops**	Los Tules
10:30-10:55	June Simpson Exploring the fructan syndrome in Agave tequilana-a non-typical crop plant	
10:55-11:10	Xianran Li Deciphering Sorghum Domestication: The Pivotal Role of Condensed Tannins in Co-evolution Among Plants, Humans, and Birds	
11:10-11:25	Arthur Wojcik Meta-analysis of wild and domesticated crop phenotypic spaces	
11:25-11:40	Laura Botigué Wheat domestication and adaptation to new environments: the role of wild emmer.	
11:40-11:55	Miriam Ferrer Multipurpose species domestication in Mayan homegardens	
11:55-12:20	Alejandro Quintero Exploring the patterns of microbiota recruitment and inheritance on the chocolate tree, Theobroma cacao L	

Scientific Program

Day 4 (10 July)

10:30 - 12:20	**Genomes and Generations: Mapping Health and Heritage Through African Genomics.**	Boca de Tomates
10:30-10:55	Brenna Henn A Rich Resource of African-Descent Genomes (CAAPA2)	
10:55-11:20	Dhriti Sengupta Impact of Khoe-San gene flow in phenotypic variation of anthropometric and cardiometabolic traits	
11:20-11:35	Jeremy Choin Genomic Insights into the population history of southern-central Africa	
11:35-11:50	Rohini Janivara Heterogeneous genetic architectures and evolutionary genomics of prostate cancer in Sub-Saharan Africa	
11:50-12:05	Esther Brielle Sudanic, Central, and Western African ancestry contributions to ancient Ugandan Populations	
12:05-12:20	Jasmin Rees Evolutionary History of Transient Receptor Proteins in Sub-Saharan Africa	
10:30 - 12:20	**Advances in Machine Learning for Evolutionary Genomics.**	Olas Altas
10:30-10:55	Sara Mathieson Interpreting deep learning methods for population genetic inference	
10:55-11:20	Siavash Mirarab Metric learning for phylogenetic placement	
11:20-11:35	Carina Cornejo Paramo A Bag-Of-Motif Model Captures Context-Specific Distal Regulatory Elements	
11:35-11:50	Steven Reilly Interrogating the relationship between conservation and regulatory function across all human promoters using machine-learning variant effect predictions	
11:50-12:05	Murillo Rodrigues A deep learning framework for evolutionary inference using ancestral recombination graphs	
12:05-12:20	Zhengting Zou Sequence embeddings by pretrained language model indicate adaptive convergence of high-order protein features.	
10:30 - 12:20	**Unveiling the evolutionary history of pathogens through paleogenomics**	Las Pilitas
10:30-10:55	Tanvi Honap Exploring oral microbial evolution in the context of European colonization of the Americas	
10:55-11:20	Laura Carrillo-Olivas Exploring Pathogens and Demographic Dynamics in Colonial Mexico City through Paleogenomics	
11:20-11:35	Paola Campos Deciphering the emergence and evolutionary history of a bacterial crop pathogen: insights from historical herbarium specimens	
11:35-11:50	Sojung Han RNA virus genomes from historical specimens of great apes	
11:50-12:05	Nasreen Broomandkhoshbacht Characteristics of a 5,500-year-old Treponema pathogen from Colombia	
12:05-12:20	Elizabeth Nelson Tuberculosis and Sociocultural Dynamics: Critical Insights from Paleogenomics in South America and Beyond	

Scientific Program

Day 4 (10 July)

10:30 - 12:20 **Science in the Spotlight: Empowering Education and Public Engagement with Cutting-Edge Science in Molecular Evolution.** Costa Alegre

10:30-10:55 **Roberto Torres**

Melanogaster Catch The Fly! Rural high schools committed to frontier research on genomics and evolution through citizen science

10:55-11:10 **Sarah Schaack**

TE Hub: A Community for Developing Tools, Trainings, and Resources to Analyze DNA Repeats

11:10-11:25 **Kate Duggan**

Disease Detectives: Using Minecraft to explore virus evolution and epidemiology

11:25-11:40 **Sònia Casillas**

Adopting ESG principles in PopLife, the population genomics browser across the tree of life

11:40-11:55 **Rebecca Siford**

Evaluating community perceptions and ethical considerations in genetics research in small scale northern Kenyan populations

11:55-12:20 **Nicolette Caperello**

Engaging the Public on Biology's Moonshot: The Earth BioGenome Project

12:20- 13:20 **LUNCH BREAK**

13:20 - 14:50 **Poster session 3**

Advances in Machine Learning for Evolutionary Genomics.

Deciphering the functional and adaptive effects of genomic structural variation.

Decoding the past to safeguard the future.

Epigenetic inheritance: from molecular mechanisms to evolutionary consequences.

Evolution of microbial communities: is the sum of parts greater than the whole?

Genomes and Generations: Mapping Health and Heritage Through African Genomics.

Human population demography and adaptation signals in the Americas.

Mitochondria: from powerhouse to processor and from marker to meaning.

New computational approaches to estimate past demographic events and natural selection.

Novel approaches to study plant domestication: disentangling the complex evolutionary history of crops

Open Symposium 2

Spatial population genetics: where are we now?

Unveiling the evolutionary history of pathogens through paleogenomics.

14:50 -17:10 **Evolution of microbial communities: is the sum of parts greater than the whole?** Los Tules

11:55-12:20 **Paul Rainey**

Communities, lineages, and the evolution of evolvability

José Flores-Uribe

Experimental evolution of synthetic phototrophic communities

Mike McDonald

A new genome from the pangenome - experimental evolution with horizontal gene transfer.

Pu Wang

Eco-evolutionary feedbacks promoted by hypermutation rate in a predator-prey microcosm

Edmund Moody

The nature of the Last Universal Common Ancestor and its impact on the early Earth system

Autumn Peterson

Disentangling the evolutionary impacts of relatedness and facultative/obligate life cycles during the transition to multicellularity

Sarit Avrani

Tradeoffs between phage resistance and nitrogen fixation drive the evolution of multicellular cyanobacteria

Scientific Program

Day 4 (10 July)

14:50 -17:10 — Human population demography and adaptation signals in the Americas — Boca de Tomates

14:50-15:15 **Cosimo Posth**

The genetic history of early Americans

15:15-15:30 **Santiago Medina-Muñoz**

Inferring fine-scale demography of indigenous lineages from 500 whole genomes across Mexico

15:30-15:45 **Valeria Anorve-Garibay**

The contribution of Neanderthal introgression to the heritability of human complex traits in Mexico

15:45-16:00 **Esther Brielle**

Adaptation in Ancient Maya

16:00-16:25 **Tábita Hünemeier**

Evolutionary Changes Throughout Human History in America

16:25-16:40 **Epifanía Arango Isaza**

Uncovering South American genomic diversity in the Southern Cone

16:40-16:55 **Marisol Espitia Fajardo**

Colombian Arawak genomes shed light on population demography of South America

16:55-17:10 **Thomaz Pinotti**

Picuris Pueblo oral history and genomics reveal genetic continuity in North American Southwest

14:50 -17:10 — Spatial population genetics: where are we now? — Olas Altas

14:50-15:15 **Alison Feder**

Spatial drivers of disease progression in bacteria and viruses

15:15-15:30 **Alessandro Lopez-Hernandez**

A space-time model for jointly inferring gene flow and natural selection

15:30-15:45 **Clara Rehmann**

In space no one can hear you sweep: a novel signal of selective sweeps in continuous geographic space

15:45-16:00 **Jack Harper**

Using biological invasions to understand rapid adaptation to new environments: a genomic reconstruction of the house sparrow global spread

16:00-16:25 **Andy Kern**

Deep Space– supervised machine learning methods for spatial population genetics

16:25-16:40 **Gilia Patterson**

A spatially explicit close kin mark recapture method for estimating census size from genetic data

16:40-16:55 **Vivaswat Shastry**

Inference of long-range gene-flow events in a background of spatially heterogeneous isolation-by-distance

16:55-17:10 **Michael Grundler**

Spatial ancestry reconstruction using parsimony on tree sequences

14:20 - 15:50 — Mitochondria: from powerhouse to processor and from marker to meaning — Las Pilitas

14:50-15:15 **Dennis Lavrov**

Evolution of mitochondrial mRNA editing and linear multipartite mt-genome architecture in calcaronean sponges

15:15-15:30 **Vaishali Katju**

No passive sidekick: a substantial role for mitochondrial compensatory evolution during adaptation of ETC-deficient strains of Caenorhabditis elegans

15:30-15:45 **Toni Gossmann**

Phylonumtomics uncovers diverse evolutionary trajectories of mitogenomic fossils buried in mammalian and avian genomes

15:45-16:00 **Dennis Lavrov**

Evolution of mitochondrial mRNA editing and linear multipartite mt-genome architecture in calcaronean sponges (invited talk)

Scientific Program

Day 4 (10 July)

14:20 - 15:50	**Mitochondria: from powerhouse to processor and from marker to meaning**	Las Pilitas
16:00-16:15	Konstantin Popadin Decoding the Puzzle of Human Somatic mtDNA Mutagenesis: Bridging Species Life-History Traits and Organ-Specific Cellular Properties	
16:15-16:30	Leah Darwin The contribution of within- and between-species mitochondrial variation to adaptation in experimental Drosophila populations	
16:30-16:45	Wei-Chin Ho Population-Genetic Environment of Genome Retention and Reduction in Mitochondria	
16:45-17:10	Michael Garvin The Importance of Mito-nuclear Epistasis in Evolutionary Genetics	
14:20 - 15:50	**Deciphering the functional and adaptive effects of genomic structural variation**	Costa Alegre
14:50-15:15	Mia Levine DNA satellite evolution triggers a cross-species incompatibility	
15:15-15:30	Julie Chuong DNA replication errors are a major source of gene amplification in adaptive evolution	
15:30-15:45	Elli Cryan Molecular evolution of a maize hybrid barrier over 12 million years suggests epistatic silencing	
15:45-16:00	Samuel Bogan Convergent Structural Variation of Antifreeze Proteins in Polar and Deep Sea Fishes	
16:00-16:15	Tristram Dodge Intergenic structural variation and ancient gene duplication underpin pigmentation diversification in swordtail fish	
16:15-16:30	Diego Hartasánchez The role of structural variation in clownfish adaptive radiation\r\n	
16:30-16:45	Megan Dennis Gene expansions contributing to human brain evolution	
16:45-17:10	Christine Beck Repetitive sequences drive genomic variation and vary regulatory landscapes	
19:00	**GALA DINNER**	

Day 5 (11 July)

09:00- 10:50	**New computational approaches to estimate past demographic events and natural selection.**	Los Tules
09:00-09:25	Jazlyn Mooney Considerations for Inferring Demography and Detecting Selection in Understudies Species	
09:25-09:40	Charleston Chiang A genealogy-based framework to estimate population structure and demographic history	
09:40-09:55	Ekaterina Noskova Hidden Diffusion: Accurate Joint Inference of Selection and Demographic History from Time-series Data for Multiple Populations	
09:55-10:10	Yun Deng Understanding the interactions between demography and selection landscapes in human populations: insights from Ancestral Recombination Graphs (ARGs)	
10:10-10:25	Anastasia Lyulina Linkage equilibrium between rare alleles	
10:25-10:40	Joshua Schraiber Inference of ultra recent demography, recurrent mutation, and natural selection from biobank-scale site frequency spectra	
10:40-10:55	M. Elise Lauterbur Bayesian mediation models reveal patterns of host genomic disease adaptation driven by environment	

Scientific Program

Day 5 (11 July)

09:00- 10:50 **Open Symposium 2** Boca de Tomates

09:00-09:25 Nadia Haghani

The genetic and structural basis of a novel iridescence phenotype in the platyfish, Xiphophorus variatus

09:25-09:40 Tamsen Dunn

Inferring the characteristics of ancient polyploidization events from modern plant genomes

09:40-09:55 Sara Miller

Genomic architecture in social insects is more strongly associated with phylogeny than social behavior

09:55-10:10 Héléna Vassilieff

A new scenario for the macroevolution of the Caulimoviridae based on the analysis of endogenous caulimovirids using CAULIFINDER\r\n

10:10-10:25 Maëlle Daunesse

Phylogenetic modeling of gene expression shifts in the mole-rat clade

10:25-10:40 Jay Storz

Life at extreme elevations: genomic and physiological insights into mechanisms of hypoxia adaptation in Andean mice

10:40-10:55 Angel Rivera-Colón

Population genetics of the Pacific acorn barnacle and insights into evolution in large populations

09:00- 10:50 **Epigenetic inheritance: from molecular mechanisms to evolutionary consequences.** Olas Altas

9:00-9:25 Peter Sarkies

Contribution of epimutation to evolution in C. elegans and beyond

9:25-9:40 Elisa Israel

MBD2/3 lost its methyl-CpG binding ability in multiple families of Holometabola

9:40-9:55 Pierre Baduel

Molecular determinants of transgenerational epigenetic variation at transposable element sequences in plants and its impact in nature

9:55-10:10 Mariana Gómez-Schiavon

Exploring the evolutionary dynamics of epigenetic switches and beyond: Insights from gene regulatory circuit models

10:10-10:25 Gustavo Tapia

The role of CTCF in the transcriptional activation of the hematopoietic-specific gene HBA2 is evolutionary conserved

10:25-10:50 Sito Torres-García

Exploring the potential for epimutations to drive cancer cell adaptation

09:00- 10:50 **Decoding the past to safeguard the future.** Las Pilitas

9:00-9:25 Mafalda Sousa Ferreira

Harnessing the past to predict the future: the evolution and future adaptive potential of seasonal camouflage in white-tailed jackrabbits

9:25-9:50 Laurent Frantz

Tracking human-mediated animal translocation in Island Southeast Asia using ancient DNA

9:50-10:05 Germán Hernández-Alonso

Exploring the Rock Dove's (Columba livia) Genetic Diversity: Insights into the Species' Evolutionary History and Conservation Challenges

10:05-10:20 Chenyu Jin

Metagenomic reconstruction of the plant community in the stomach content of a steppe bison living over 48,000 C14 years ago

10:20-10:35 Maria Zicos

Reconstructing the population history of the endangered Hispaniolan solenodon, one of the most evolutionary distinct mammals on earth

10:35-10:50 Xuejing Wang

Reviving from four birds: post-bottleneck overview of temporal genomics of Mauritius kestrel

Scientific Program

Day 5 (11 July)

09:00- 10:50	**Human genetic variability in the Pangenomic era..**	Costa Alegre
9:00-9:25	Erik Garrison Implicit pangenomics	
9:25-9:50	Gabriel Renaud Maximum-likelihood inferences from ancient environmental DNA using panmitogenomes	
9:50-10:05	Runyang Nicolas Lou Global diversity, recurrent evolution, and recent selection on amylase structural haplotypes in humans	
10:05-10:20	Joana Rocha A Pan-pangenome captures the full spectrum of genetic variation in humans, chimpanzees and bonobos	
10:20-10:35	Rick McLaughlin Pangenome assemblies reveal the evolution and recent in vivo activity of human LINE-1 retrotransposons	
10:35-10:50	Arang Rhie Living with your dynamic genome: lessons learned from T2T genomes	
10:50 -11:20	**COFFEE BREAK**	
11:20-12:20	**Closing Keynote Speaker**	Los Tules
	Scott Edwards PhyloG2P: the new science of connecting genomes to phenotypes via phylogenies	
12:20-14:20	**Awards Ceremony and Introduction to SMBE 2025**	

Floor Map

SMBE 2024
PUERTO VALLARTA

CONVENTION CENTER
GROUND FLOOR DISTRIBUTION

1. LOBBY
Registration
Front desk

2. FOYER
Sponsors Stands
Coffee break & Lunch Break

3. POSTER SESSIONS

4. POSTER SESSIONS

5. LOUNGE AREA

6. LOS TULES
Plenary

6. LOS TULES
7. BOCA DE TOMATES
8. OLAS ALTAS
9. LAS PILITAS

Concurrent Symposia

7. BOCA DE TOMATES

Workshops:
Introduction to STDPOPSIM
July 7th 12 - 14 hrs

GENHIST kickoff at SMBE
July 7th 14 - 16 hrs

SMBE Business Meeting.
July 9th 12:20 - 13:20 hrs

MBE & GBE Editors meetings.
July 8th 14:20 - 15:50 hrs
July 9th 13:20 - 14:50 hrs
July 10th 12:20 - 13:20 hrs

F FOOD STATION

V VEGAN & VEGETARIAN FOOD STATION

10. KITCHEN

11. COWORKING AREA

Breastfeeding room

Undergraduate's dinner
July 7th 19 - 21 hrs

12. KIDS CLUB
(0 - 10 years old)

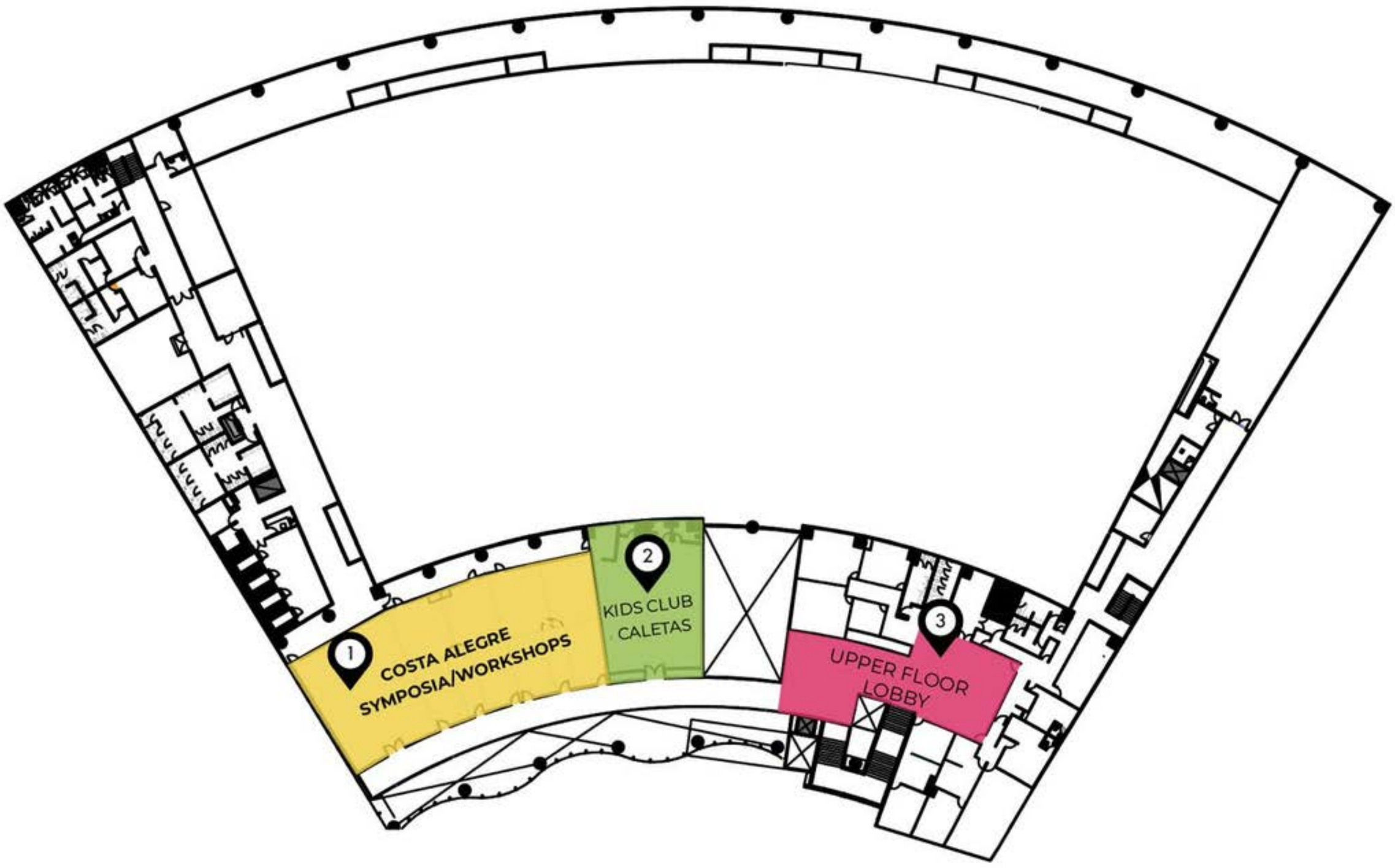
13. DELI & COFFEE

14. MEDICAL SERVICE

Floor Map

SMBE 2024
PUERTO VALLARTA

CONVENTION CENTER
UPPER LEVEL DISTRIBUTION



1. COSTA ALEGRE
Concurrent Symposia

2. CALETAS
Kids Club (11 - 17 years old)

1. COSTA ALEGRE

WORKSHOPS

July 7th 11 - 16 hrs
Introduction to programmatic access to EnsEMBL.

Online attendees guidelines

Dear online attendee,

Broadcasting of SMBE 2024 will take place through an user/password access system available through the SMBE 2024 website (https://smbe2024.org/). Live broadcasting will take place the duration of the event (program under UTC -6 time zone schedule), and only to attendees who specifically registered to attend the conference online.

Online attendees will receive an email from SMBE 2024 Organizing Committee previous to the event with their respective credentials to access the system.

SMBE 2024
PUERTO VALLARTA

Smbe
Society for Molecular Biology & Evolution

SMBE 2024
PUERTO VALLARTA

Abstract book

Oral presentations

334 - A Bag-Of-Motif Model Captures Context-Specific Distal Regulatory Elements

Carina Paola Cornejo Paramo

Carina Paola Cornejo Paramo, Xuan Zhang, Lithin K Louis, Yi-Hua Yang, Zelun Li, David Humphreys, Emily S Wong
University of New South Wales, Victor Chang Cardiac Research Institute, Victor Chang Cardiac Research Institute, Victor Chang Cardiac Research Institute, Victor Chang Cardiac Research Institute, Victor Chang Cardiac Research Institute, Victor Chang Cardiac Research Institute, University of New South Wales, Victor Chang Cardiac Research Institute

Enhancers are cis-regulatory regions that orchestrate the precise control of spatiotemporal gene expression by mediating transcription factor (TF) interactions. Deciphering their sequence properties is challenging due to rapid evolution, low conservation, and binding complexity (1,2). We developed "Bag-of-Motifs" (BOM), a tree-ensembleâ€“based machine-learning framework designed to identify cell type-specific enhancers using TF binding motif composition. BOM is an adaptable framework for dissecting the key TF binding motifs within regulatory sequences between specific cell types, states, and conditions. We use Shapley values, a concept rooted in game theory, to elucidate the individual contribution of binding motifs to the classification tasks (3). BOM excels in classifying candidate enhancers from multiple cell types and states across various species and developmental stages and outperforms deep learning models in the same classification task (auROC = 0.98). Shapley values also allow us to correlate motifs with transcription factor expression to gain a deeper insight into the motifs crucial for enhancer function and gene regulation. In conclusion, our 'Bag-of-Motifs' framework not only excels in classifying candidate enhancers across diverse biological contexts but also sheds light on the critical motifs governing enhancer function and gene regulation.

1. Smith, G. D., Ching, W. H., Cornejo-PÃ¡ramo, P., & Wong, E. S.(2023). Decoding enhancer complexity with machine learning and high-throughput discovery. Genome Biology, 24(1),1-25.
2. Spitz, F., & Furlong, E. E.(2012). Transcription factors: from enhancer binding to developmental control. Nature reviews genetics, 13(9),613-626.
3. Shapley, L. S.(1953). A value for n-person games.

409 - Sequence embeddings by pretrained language model indicate adaptive convergence of high-order protein features.

Zhengting Zou

Zhenqiu Cap, Zhengting Zou
University of Science and Technology of China, Institute of Zoology, Chinese Academy of Sciences, Institute of Zoology, Chinese Academy of Sciences

Various computational methods have been proposed to investigate the genetic basis for organismal functional convergence, specifically focusing on convergence of amino acid state at individual sites in proteins. However, beside site-level convergence, convergence of high-order features in protein sequence can also lead to functional similarity. In this study, we derive fixed-length embeddings of protein sequences by a pretrained protein language model (PLM), and propose that these embeddings reflect high-order features of the proteins. We investigated multiple cases of distantly related but functionally similar protein orthologs putatively having experienced adaptive convergence during evolution, e.g. proteins in hyperthermophilic bacteria and archaea species. We found that in most cases, these

proteins have similar embeddings despite their site-level phylogenetic divergence, supporting that PLM embeddings can reflect high-order feature convergence of proteins. We then proposed a convergence test pipeline PEAC (protein embedding adaptive convergence) for multiple lineages of species with functional convergence, based on comparing the PLM embedding similarity of proteins with simulated neutral background. Using PEAC in a genome-wide case study of echolocating mammals, we discovered proteins with putative adaptive convergence that are not reported previously, alongside with known convergent cases such as Prestin.

729 - A deep learning framework for evolutionary inference using ancestral recombination graphs

Murillo F Rodrigues

Murillo F Rodrigues, Nathaniel S Pope, Peter L Ralph, Andrew D Kern
University of Oregon, University of Oregon, University of Oregon, University of Oregon

Applications of deep learning in population genetics typically take as input the genotype matrix or compressed summaries thereof.
While these methods have proven powerful and versatile, a major hurdle is that they do not scale well to biobank-size datasets which are currently being generated.
An alternative lies in using the ancestral recombination graph (ARG), which encodes the transmission of genetic material from ancestors to extant individuals across the genome.
The ARG can solve many issues in population genetic inference because it is:
(i) sufficient, in that it contains all the information obtainable from the genotype matrix,
(ii) compact, in that shared genetic variation is stored as relationships between ancestral haplotypes, and
(iii) multiscale, in the sense that the span of ancestral haplotypes and edges naturally reflects the spatial (along the genome) and temporal scales at which shared ancestry impacts variation, without arbitrary discretization into windows.
We developed an architecture, tsNN, that takes ARGs (as tree sequences) as input for evolutionary inference by borrowing ideas from the field of graph neural networks (GNNs).
The sequence-like nature of the ARG is captured by updating the node embedding with edge insertions and deletions.
We applied tsNN to two tasks: dating mutations and inferring demographic parameters.
tsNN outperforms a naive GNN despite having fewer learnable parameters, demonstrating that the biology-informed inductive bias can improve learning.
Further, tsNN performs better at dating mutations than the current model-based methods.
Taken together, our results demonstrate the utility of graph neural networks in ARG-based inference.

853 - Interrogating the relationship between conservation and regulatory function across all human promoters using machine-learning variant effect predictions

Steven Reilly

John C Butts, Stephen Rong, Rodrigo I Castro, Sager J Gosai, Mackenzie Noon, Rohit Ghosh, Ryan Tewhey, Steven K Reilly
The Jackson Laboratory , University of Maine, Yale University, The Jackson Laboratory, The Broad Institute, Yale University, Yale University, The Jackson Laboratory , Graduate School of Biomedical Sciences and Engineering, University of Maine, Graduate School of Biomedical Sciences, Tufts University School of Medicine, Yale University School of Medicine, Dept. of Genetics, Yale Center for Genomics Health

The ability to predict an alleleâ€™s regulatory function would improve our identification of causal, non-coding variants underlying human traits and disease. Evolutionary conservation is commonly used as a functional proxy, but the relationship between conservation and

regulatory impact are poorly understood. While massively parallel reporter assays (MPRAs) are powerful tools to identify functional regulatory variants, they can maximally assay ~100K genetic variants at a time, limiting their applicability to a small subset of currently observed human variation. Our recently developed deep learning model, Malinois, accurately predicts empirical MPRA results (R>0.89) in multiple cell types and is scalable to the whole genome.

Here we present the first comprehensive analysis of regulatory function and evolutionary conservation at the single nucleotide level in all promoters. We generate saturation mutagenesis MPRA predictions for ~14M SNVs in promoters of all protein-coding genes and provide these functional scores for all promoters as prospective maps for variant interpretation. Genome-wide, we find weak purifying selection against both loss-of-function (LoF) and gain-of-function (GoF) promoter variants, even for genes that show no evidence of coding LoF-intolerance. Some promoters show strong purifying selection against LoF variants (e.g. LDLR and U2AF2), GoF (e.g. LCP2), or both (e.g. MAG), indicative of different modes of selection acting on promoter activity. Finally, we find recent purifying selection against LoF and GoF promoter variants using gnomAD. We deploy these maps to interpret rare variant allelic series in biobanks and identify causal variation underlying human selection scans.

138 - Contribution of transposable elements in the sex gap longevity of different Drosophila species

Miriam Merenciano

Miriam Merenciano, Sonia Janillon, Camille Mermet-Bouvier, Nelly Burlet, Matthieu Boulesteix, Marie Fablet, Cristina Vieira
Laboratory of Biometry and Evolutionary Biology, „ Laboratory of Biometry and Evolutionary Biology, Laboratory of Biometry and Evolutionary Biology, Laboratory of Biometry and Evolutionary Biology, Laboratory of Biometry and Evolutionary Biology, Laboratory of Biometry and Evolutionary Biology, Laboratory of Biometry and Evolutionary Biology

In Drosophila, like in many other animal species, females tend to live longer than males, a phenomenon known as sex gap in longevity (SGL). One of the possible causes underlying this phenomenon could be related to the high number of transposable elements (TE) in the Y chromosome (toxic Y effect). TE activity is normally repressed by epigenetic mechanisms. However, it is known that this regulation is disrupted with age. Since the Y chromosome is rich in TEs, more TEs may become active in old males than in old females, generating more somatic mutations, and reducing longevity in males. In this work, we studied the natural variation in SGL in several natural populations of three different Drosophila species that vary in their TE content: Drosophila melanogaster, Drosophila simulans, and Drosophila suzukii. Furthermore, we found that the replacement of the Y chromosome between strains with different SGL reduces male lifespan over generations and thus increases SGL, suggesting an important role of the Y chromosome in male longevity. Finally, RNA-seq analysis from old and young flies suggested that there is an increased number of upregulated TE families in old samples, and more specifically in old males compared to old females, and that the total fraction of transcripts derived from repeats increase during aging depending on the species and the population tested. Overall, this work tries to better understand the genomic differences that lead to variation in longevity patterns between sexes, and emphasizes the importance of TEs in male longevity.

375 - Signatures of extreme longevity: a perspective from bivalve molecular evolution

Mariangela Iannello

Mariangela Iannello, Giobbe Forni, Giovanni Piccinini, Ran Xu, Jacopo Martelossi, Fabrizio Ghiselli, Liliana Milani
University of Bologna, University of Bologna, University of Bologna, University of Bologna, University of Bologna, University of Bologna, University of Bologna

Longevity is a complex and poorly characterized trait: we know that some species can live longer than others, but the molecular players

responsible for the extended longevity phenotype are largely unknown. In this context, bivalve molluscs can provide novel perspectives: the class Bivalvia shows the highest lifespan disparity within Metazoa, ranging from 1 to 500+ years, and includes the longest-lived non-colonial animal species known so far, the clam Arctica islandica. Bivalves therefore represent important resources to provide insights into the molecular mechanisms of longevity.

In this work, we leveraged transcriptomic resources spanning thirty bivalve species and we investigated genes with signature of convergent molecular evolution in bivalves showing extreme longevity. We found that a majority of genes showing convergent evolution in long-lived bivalves constitute a continuous network of protein-protein interactions. Such network is enriched for factors with experimental support for a role in longevity in other animals. The network highlights that an integration of different genes and pathways is required for the extended longevity phenotype, and genes involved in cell proliferation control, translational machinery, and response to hypoxia seem to have a central role in lifespan extension.

Our results suggest that the mechanisms underlying extended longevity are, at least partially, similar across metazoans: while some genes in the network have experimental support of a role in longevity in model species, other genes in the network may represent new possible candidates with a role in extending lifespan, both in bivalves and other animals.

458 - Accounting for common genetic variation improves chronological age prediction in African populations

Gillian L Meeks

Gillian L Meeks, Brenna M Henn, Hana Asnake, Brooke Scelza, Shyamalika Gopalan

University of California Davis, University of California Davis, University of California Davis, University of California Los Angeles, Clemson University

Aging is associated with genome-wide DNA methylation changes in humans, facilitating the development of epigenetic age prediction methods for forensic and health-care applications. However, most prior research studying these changes have focused on European-ancestry individuals, which represent only a fraction of extant human genetic diversity. We test one of the most popular epigenetic age predictors, the Horvath model, on data from central African Baka foragers, southern African â‰ Khomani San foragers, and southern African Himba pastoralists, and observe substantially higher mean errors in these cohorts compared to European-ancestry samples. Common genetic variants are known to influence baseline methylation levels, introducing noise to the relationship between methylation and age, and potentially affecting prediction performance. However, these interactions are not well characterized for genetically diverse African-ancestry populations. We hypothesize that this unaccounted-for variation partially explains why published age predictors perform worse in our cohorts. Supporting this, we find that 25% of the predictor loci used by the Horvath model are influenced by genetic variants that segregate in at least one of our cohorts. We show that accounting for these types of variants improves age association at Horvathâ€™s predictor loci, and furthermore enables the discovery of 211 additional age associations. Next, we developed a novel age prediction model that corrects for the effect of nearby genetic variants on DNA methylation levels at age-associated loci. Our work shows that accounting for a broad range of human genetic diversity enables more accurate and universally-applicable epigenetic age prediction methods.

783 - Exploring the chronological lifespan spectrum of Candida species

Eugenio Mancera

Alberto Luna-Carrera, Alexander DeLuna, Eugenio Mancera

1Departamento de IngenierÃ-a GenÃ©tica, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Unidad

Irapuato, 2Unidad de GenÃ³mica Avanzada, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, 1Departamento de IngenierÃ-a GenÃ©tica, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Unidad Irapuato

Fungal microorganism, including those that inhabit our bodies and cause infection, often undergo stationary phases characterized by cell-cycle arrest. The capacity to survive in this state is known as the chronological lifespanâ€”a widely used model to understand postmitotic cellular agingâ€”allowing organisms to renter a proliferative phase and to survive in their natural environments. Despite its relevance, we know little about the diversity of chronological lifespan among species and how it relates to the specific environments in which they live. In this talk, we will present the implementation of novel experimental approaches to portray the chronological lifespan of Candida species, a group of microorganisms best known for causing a variety of human diseases, but mostly composed of non-pathogenic yeasts. Our results show great lifespan diversity among these fungi and changes in longevity by caloric restriction. Lifespan is not directly associated to pathogenicity, as there are short- and long-lived species among the usual inhabitants of the human body. Surprisingly, lifespan does not vary according to the phylogenetic position of these species, suggesting that stationary-phase survival is a plastic trait in evolutionary terms. We will present analysis of specific gene knockout mutants shedding light into the molecular underpinnings of lifespan diversity, with possible implications for the control of these clinically important fungi.

311 - The evolutionary origin of the Columbian mammoths from the Basin of Mexico and its implications for the Mammuthus genus

Federico Sánchez Quinto

Federico SÃ¡nchez Quinto, Eduardo Arrieta Donato, Ã•ngeles Tavares GuzmÃ¡n, Rigoberto Padilla Bustos, Alejandra Castillo Carbajal, Ernesto Garfias Morales, Miriam Bravo LÃ³pez, Viridiana Villa Islas, Vanssy Li, Alejandro LÃ³pez JÃ-menez, Mashaal Sohail, Juan Camilo ChacÃ³n-Duque, JoaquÃ-n Arroyo-Cabrales, MarÃ-a C Ã•vila Arcos

International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), Centro de Ciencias GenÃ³micas, Universidad Nacional AutÃ³noma de MÃ©xico, International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), Centre for Palaeogenetics, Laboratorio de ArqueozoologÃ-a, SubdirecciÃ³n de Laboratorios y Apoyo AcadÃ©mico, Instituto Nacional de AntropologÃ-a e Historia, Centro de Ciencias GenÃ³micas, Universidad Nacional AutÃ³noma de MÃ©xico, Centre for Palaeogenetics, Department of Zoology, Stockholm University, Laboratorio de ArqueozoologÃ-a, SubdirecciÃ³n de Laboratorios y Apoyo AcadÃ©mico, Instituto Nacional de AntropologÃ-a e Historia, International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM)

Recent paleogenomic studies have suggested a hybrid origin for the Columbian mammoth deriving from the admixture of an ancestral woolly mammoth and a steppe mammoth lineage. While its geographical range once extended from North to Central America, only mitochondrial data is currently available from across the USA. Moreover, only one Columbian mammoth autosomal genome has been generated to date, obtained from a sample close to the putative woolly-Columbian hybrid zone. Such genetic sampling hinders broad conclusions about its evolutionary trajectory. Recently, multiple Columbian mammoth remains have been discovered in the Basin of

Mexico, providing a unique opportunity to better understand their evolutionary origin. We screened 72 libraries from M. columbi, authenticating their endogenous origin and molecularly sexing individuals. We performed mitochondrial DNA (mtDNA) capture-enrichment experiments recovering 28 genomes with >10x coverage, and generated a >1.5x nuclear genome. We compared this data to mitochondrial and nuclear reference sequences, respectively, from Columbian and woolly mammoths across North America and Eurasia, and performed phylogenetic and demographic analyses. Inferences stemming from both datasets provide a more detailed picture of the evolutionary trajectory of the Columbian mammoth compared to what has been previously reported. Thus, these results highlight the importance of recovering ancient genomic data from wider geographical ranges in order to fully understand the evolutionary history of extinct species. Our study provides a better understanding of the evolutionary origin of the Columbian mammoth in Mexico and America, as well as valuable insights into the evolutionary trajectory of the mammuthus genus.

756 - Compacted hair in broken carnivore teeth reveal dietary prey of historic lions

Alida de Flamingh

Alida de Flamingh, Tom P Gnoske, Julian C Kerbis Peterhans, Vellizar Simeonovski, Nduhiu Gitahi, Ogeto Mwebi, Bernard R Agwanda, Julian M Catchen, Alfred L Roca, Ripan S Malhi

University of Illinois at Urbana-Champaign, Field Museum of Natural History, Field Museum of Natural History, Field Museum of Natural History, University of Nairobi, National Museums of Kenya, National Museums of Kenya, University of Illinois at Urbana-Champaign, University of Illinois at Urbana-Champaign, University of Illinois at Urbana-Champaign

The synergistic advancement of molecular and computational technologies has pushed genomics into a new era; complete nuclear genomes can now be sequenced from minuscule quantities of DNA and from specimens that are more than a million years old. DNA analysis from hair samples is a well-established approach widely used in human forensics and wildlife conservation science. Hair samples are less prone to contamination from external DNA sources, and can be used to identify the species of animal from which the hair originated. Here we used ancient DNA and bioinformatic methodologies optimized for low quantity and quality degraded DNA, to systematically identify the dietary prey species from four individual hairs and three hair clumps compacted in the teeth of the infamous Tsavo lions that are known to have also preyed on humans during the late 1890â€™s in Kenya. Molecular and bioinformatic analysis of hair DNA identified giraffe, human, oryx, waterbuck, wildebeest and zebra as prey species. DNA preservation allowed for the reconstruction of complete mitogenome profiles for zebra and giraffe. Giraffe mitogenomes are phylogeographically partitioned, and we found that the Tsavo lions ate at least two individuals that belong to a subspecies of Masai giraffe (Giraffa tippelskirchi tippelskirchi) typically found in southeast Kenya. The protocol and approach reported here enable a better understanding of the hunting behaviors, diets, and ecology of historic individuals, populations, and species and holds promise for elucidating these characteristics in extinct populations and species.

776 - The last especimen of Eurasian lynx (Lynx lynx) from the Iberian Peninsula: new genetic data leads to review the biogeographic history of lynx matrilineal linages in western Europe.

Gloria Gonzalez Fortes

Gloria Gonzalez Fortes, Eva Fernandez Bejarano, Aurora Grandal D'Anglade

University of A CoruÃ±a, University of A CoruÃ±a, University of A CoruÃ±a, Univeresity of A CoruÃ±a

The Eurasian lynx (Lynx lynx), once widespread across Eurasia, was recently extirpated from western Europe, mainly due to anthropogenic pressures and habitat fragmentation. Although the past distribution of the species encompassed the Iberian Peninsula from the Late Pleistocene up until the 17th century, little is known about the relationship of these Iberian populations with other Eusarian lynxes. Here

we present the paleomolecular study of the last recorded specimen of Eurasian lynx from the Iberian Peninsula (dated to ~ 200 yBP, from Sima Topinoria, Cantabria), which has been analysed at genetic level using cutting edge technologies for the sequencing of ancient DNA. The specimen showed an excellent state of biomolecular conservation, preserving up to 60% of endogenous DNA, which allowed the reconstruction of its complete mitochondrial genome and the analysis of genome-wide nuclear markers. The phylogenetic analysis of the mitogenome revealed that the lynx from Sima Topinoria (ST) belonged to an extinct lineage that is basal to all modern populations of Eurasian lynx, distributed from Northwestern Europe to Northeast Asia. This lineage identifies ST as belonging to a relict population of Pleistocene lynxes that survives in the Cantabrian mountains until recent times. Indeed, the molecular identification of this specimen revises the date of the species extirpation from the Iberian Peninsula to only 200 years ago.

934 - Small but mighty: ancient DNA analyses of late Pleistocene and Holocene microvertebrates from Central Texas

David Treviño Ledesma

David T Ledesma, Melissa E Kemp

The University of Texas at Austin, The University of Texas at Austin

Fossil identifications are foundational for making robust interpretations of past evolutionary and ecological processes. Morphology has traditionally been the main method for fossil identification; however, ancient DNA (aDNA) has emerged as an exciting new tool for identifying fossils to lower taxonomic levels than what morphology can often provide. The ability to recover aDNA from fossils is known to be dependent on age and environment, with older and lower latitude sites being less suitable for long-term DNA preservation. Furthermore, aDNA research is relatively limited for microvertebrates in part because recovery of aDNA becomes increasingly difficult with smaller amounts of starting material for the extraction process. Here we investigate our ability to obtain aDNA from late Pleistocene to Holocene microvertebrate herpetofauna remains from a low latitude site in Central Texas. We successfully extracted aDNA from 15 out of 20 specimens ranging from 0.9 to 56 mg. We found a positive linear relationship between fossil mass and DNA yield with few fossils below 10 mg producing detectable amounts of DNA. Preliminary shotgun sequence data of aDNA extracted from salamander fossils confirms the presence of endogenous DNA. We performed phylogenetic analyses with extant species and corroborated morphological identification of fossils to a clade of tiger salamanders. aDNA indicates that the fossils are closely related to species that today occur further west and south of the fossil locality. These results reinforce the merit of aDNA for microvertebrate material from low-latitude sites and provide insights into past diversity and biogeography of Central Texas salamanders.

243 - Detecting barriers to gene flow from genomic patterns: application to maize and teosintes

Maud I Tenaillon

Maud I Tenaillon, Ewen Burban, Arthur Wojcik, Stella Huyn, AgnÃ¨s Rousselet, Marie Couderc, Harry Belcram, Yves Vigouroux, Sylvain GlÃ©min

GQE-Le Moulon, UniversitÃ© Paris-Saclay, INRAE, CNRS, AgroParisTech, ECOBIO-UMR 6553, UR1-CNRS, GQE-Le Moulon, UniversitÃ© Paris-Saclay, INRAE, CNRS, AgroParisTech, UMR DIADE, UniversitÃ© Montpellier, IRD, CIRAD, GQE-Le Moulon, UniversitÃ© Paris-Saclay, INRAE, CNRS, AgroParisTech, UMR DIADE, UniversitÃ© Montpellier, IRD, CIRAD, GQE-Le Moulon, UniversitÃ© Paris-Saclay, INRAE, CNRS, AgroParisTech, UMR DIADE, Univ Montpellier, IRD, CIRAD, UMR DIADE, ECOBIO-UMR 6553, UR1-CNRS, Department of Ecology and Genetics, Evolutionary Biology Centre

Domestication is the process of divergent selection between wild forms undergoing natural selection in their habitats, and domesticates evolving under combined natural and human-mediated selection. Domestication has been considered as a choice example to study

adaptation. It also offers an excellent opportunity to catch the very-first processes at work in ecological speciation, where adaptive divergence between nascent lineages triggers the onset of reproductive isolation (RI). Here, we present RIDGE â€“ Reproductive Isolation Detection using Genomic polymorphisms â€“ a tool tailored for quantifying gene flow barrier proportion and identifying the relevant genomic regions. RIDGE relies on an Approximate Bayesian Computation with a model-averaging approach to accommodate diverse scenarios of lineage divergence. It captures heterogeneity in effective migration rate (m) along the genome while accounting for variation in linked selection and recombination, estimates barrier proportion (Q) and provides a test at the locus scale to detect gene flow barriers. Simulations and analyses of published datasets indicate that RIDGE is highly effective at both detecting ongoing migration and identifying barrier loci, even for recent divergence times (<0.1 2Ne generations). We applied RIDGE to 40 resequenced genomes from maize and teosintes. We estimated Q to 1.7% of the maize genome (49Mb). Interestingly, one locus known to be a barrier between maize and teosintes was detected. RIDGE also demonstrated its ability to distinguish barrier loci from domestication loci.

308 - Evidence for the role of selection for reproductively advantageous alleles in human aging

Jianzhi George Zhang

Jianzhi Zhang
University of Michigan

The antagonistic pleiotropy hypothesis posits that natural selection for pleiotropic mutations that confer earlier or more reproduction but impair the post-reproductive life causes aging. This hypothesis of the evolutionary origin of aging is supported by case studies but lacks unambiguous genomic evidence. Here we genomically test this hypothesis using the genotypes, reproductive phenotypes, and death registry of 276,406 UK Biobank participants. We observe a strong, negative genetic correlation between reproductive traits and lifespan. Individuals with higher polygenetic scores for reproduction (PGS_R) have lower survivorships to age 76 (SV_76), and PGS_R increased over birth cohorts from 1940 to 1969. Similar trends are seen from individual genetic variants examined. The antagonistically pleiotropic variants are often associated with cis-regulatory effects across multiple tissues or on multiple target genes. These and other findings support the antagonistic pleiotropy hypothesis of aging in humans and point to potential molecular mechanisms of the reproduction-lifespan antagonistic pleiotropy.

697 - Single-cell Patterns and Evolutionary Mechanisms of Subgenome Expression Differentiation in Common Wheat

Wenfeng Qian

Wenfeng Qian
Institute of Genetics and Developmental Biology Chinese Adademy of Sciences

Polyploidization enhances plant adaptability and plasticity beyond their diploid ancestors, largely through subgenome expression diversification. In this study, we developed a single-cell expression atlas for common wheat seedlings (hexaploid) at the subgenomic level. We categorized genes into four groups based on their expression variations across cell types and subgenomes, observing varying evolutionary pressures. We explored the evolutionary drivers behind subgenome expression diversity by comparing our data with the single-cell transcriptomes of diploid or tetraploid ancestors. The expression patterns of homoeologous genes in common wheat closely mirrored those in its tetraploid ancestor. This study showcases the use of single-cell technology to elucidate the patterns and evolutionary processes of plant polyploidization.

830 - PoMo via RevBayes: Inferring Phylogenies, Disentangling GC-bias and Balancing Selection

Carolin Kosiol

Svitlana Braichenko, Rui Borges, Carolin Kosiol
University of St Andrews, Vetmeduni Vienna, University of St Andrews

The interplay between mutation, genetic drift, directional and balancing selection shapes the diversity of populations, however, it can be highly complex and challenging to disentangle. To address this, sophisticated models with a high degree of flexibility and the ability to handle multiindividual data are required. We have developed a set of polymorphism-aware phylogenetic models called PoMos. These models are based on the Moran model and have recently demonstrated their effectiveness in inferring species trees, as well as capturing mutational effects, fixation biases, and GC-bias rates. To enhance accessibility, we have implemented these models in the open-source Bayesian inference framework RevBayes. In this study, we further expanded the capabilities of PoMos to investigate neutral, GC-bias, and for the first time, balancing selection. The novel aspect of our approach lies in PoMosâ€™ ability to account for long standing ancestral polymorphisms and incorporate parameters that measure frequency-dependent selection. We implemented validation tests and assessed the model on the data simulated with SLiM and a custom Moran model simulator. We examined real sequences from Drosophila populations to gain insights into the evolutionary dynamics of regions subject to frequency-dependent balancing selection, particularly in the context of sex-limited colour dimorphism.

840 - Transcript Homology Relationships and Phylogeny Reconstruction

Aida Ouangraoua

Wend Yam D D Ouedraogo, Aida Ouangraoua
University of Sherbrooke, University of Sherbrooke

Alternative Splicing (AS) is a ubiquitous process in eukaryotes by which different combinations of introns are spliced to produce distinct transcript isoforms from a gene. It has garnered interest in the last few decades because of its significant role in gene function diversification, as well as in cancer. Recent studies have highlighted that the transcript isoforms of human genes are often conserved in orthologous genes from various species. The conserved transcripts are referred to as transcript orthologs, and the identification of transcript ortholog groups provides valuable insights for studying their functions. Exploring the evolutionary history of homologous transcripts enhances our understanding of their protein functions and their origins. It allows us to better understand the role of alternative splicing in transcript evolution. It also provides additional information on gene evolution, such as the history of transcript isoform gains and losses in the evolution of a gene family, and the evolution of alternative splicing.

In this presentation, we will describe a computational method and a database we have developed to study transcript evolution. The method first infers orthology relations between transcripts of homologous genes [Ouedraogo-Ouangraoua 2023, https://doi.org/10.1007/978-3-031-36911-7_2], and then uses this information to estimate a transcript phylogeny based on precomputed clusters of orthologous transcripts [Ouedraogo-Ouangraoua 2024, to appear in RECOMB-CG 2024 proceedings]. The TranscriptDB database provides access to the transcript phylogenies and transcript orthology relations inferred for all eukaryotic gene families of the Ensembl Compara database [Ouedraogo-Ouangraoua 2023, https://transcriptdb.cobius.usherbrooke.ca/].

987 - The evolution of hibernation-related shifts in gene expression across hibernating mammals

Joanna L. Kelley

Joanna L Kelley, Anne Nakamoto, Blair Perry, Heiko Jansen, Charles Robbins

UCSC, WSU Bear Center, UCSC, WSU, WSU, WSU

Hibernation in bears involves a series of seasonally reversible changes in metabolism and physiology. Understanding the mechanisms behind the reversible nature of hibernation has implications for both animal and human health. One of the factors driving changes in hibernation is the significant alteration in gene expression across various tissues. The extent to which hibernation-related shifts in gene expression are shared across hibernating mammals is not well understood. Moreover, the evolutionary origins of hibernation remain unclear. In this study, we utilized RNA-sequencing datasets within various hibernating mammals and performed comparative transcriptomic analysis of multiple tissues to identify hibernation-related genes in each species. By comparing orthologous genes across species, we found that the sharing of hibernation-related differential gene expression is limited. Functional analysis showed that many of the shared genes are involved in metabolic processes, while genes unique to certain hibernating rodents were associated with functions such as immunity and blood coagulation. Additionally, we observed that many groups of orthologous genes contain genes that are upregulated during hibernation in certain species, while downregulated in others, highlighting the complexity and potential rewiring of hibernation regulatory pathways throughout evolution. We couple these analyses with analyses of population-specific selection and broader scale molecular evolution to determine the selective-regimes underlying hibernation-related genes. We delve into the evolution of regulatory regions in brown bears to uncover population-specific selection. Investigating the similarities and differences between hibernation phenotypes at the molecular level can contribute to understanding the evolutionary origins of hibernation.

1036 - The distribution of fitness effects varies with phylogeny across animals

Kirk E. Lohmueller

Meixi Lin, Eduardo G. Amorim, Sergio F. Nigenda-Morales, Annabel C. Beichman, Paulina G. NuÃ±ez-Valencia, Sneha Chakraborty, Jonathan Mah, Jacqueline A. Robinson, Chris C. Kyriazis, Christian D. Huber, Andrew E. Webb, Sarah D. Kocher, Andres Moreno-Estrada, Kirk E. Lohmueller

UCLA, UC Berkeley, California State University Northridge, National Laboratory of Genomics for Biodiversity , University of Washington, National Laboratory of Genomics for Biodiversity , UCLA, UCLA, UCSF, UCLA, Pennsylvania State University, Princeton University, Princeton University, Princeton University, National Laboratory of Genomics for Biodiversity, UCLA

The distribution of fitness effects (DFE) describes the selection coefficients (s) of newly arising mutations and fundamentally influences population genetics processes. Despite its importance, the extent of DFE variation in natural populations has not been systematically investigated across species with divergent phylogenetic histories and distinct ecological roles. Here, we inferred the DFE in natural populations of eleven animal (sub)species across mammals, birds, and insects using the site frequency spectrum (SFS). We find a significant phylogenetic signal in the DFE parameters, with the expected selection coefficient (E[s]) being more similar in more closely related species (Pagel's lambda = 0.84, P = 0.04). Additionally, mammals have a higher proportion of strongly deleterious mutations (22.4% to 47.4% in mammals; 0.0% to 16.0% in insects and birds) and a lower proportion of weakly deleterious mutations than insects and birds. Population size is negatively correlated with the average s of new deleterious mutations (P = 0.03). This result may imply a drift barrier to s, similar to that for mutation rates. Specifically, larger populations experience more efficient positive selection for improved protein stability, allowing subsequent deleterious mutations to be only mildly deleterious. In smaller populations, however, selection is less efficient at increasing protein stability, and so subsequent mutations are more strongly deleterious. Several of our results are also consistent with Fisherâ€™s geometric model. Overall, our study provides insights into the evolution of a fundamental population genetics parameter and suggests the evolutionary stability of the DFE across animal species.

341 - Multiple periods of admixture and isolation during and after the Transatlantic Slave Trade on

the island of São Tomé.

Marta Ciccarella

Marta Ciccarella, Romain Laurent, Zachary Szpiech , Etienne Patin, FranÃ§oise Dessarps-Freichey , JosÃ© UtgÃ© , Laure LÃ©mÃ©e , Armando Semo, Jorge Rocha, Paul Verdu

UMR7206 Eco-anthropologie, CNRS, MNHN, UniversitÃ© Paris CitÃ©, CIBIO-InBIO, BIOPOLIS Program, Departamento de Biologia, Faculdade de CiÃªncias, Universidade do Porto, UMR7206 Eco-anthropologie, CNRS, MNHN, UniversitÃ© Paris CitÃ©, Department of Biology, Penn State University, Institute for Computational and Data Sciences, Penn State University, Human Evolutionary Genetics Unit, Institut Pasteur, UniversitÃ© Paris CitÃ©, CNRS UMR2000, UMR7206 Eco-anthropologie, CNRS, MNHN, UniversitÃ© Paris CitÃ©, UMR7206 Eco-anthropologie, CNRS, MNHN, UniversitÃ© Paris CitÃ©, Plate-forme Technologique Biomics, C2RT, Institut Pasteur, CIBIO-InBIO, BIOPOLIS Program, Departamento de Biologia, Faculdade de CiÃªncias, Universidade do Porto, CIBIO-InBIO, BIOPOLIS Program, Departamento de Biologia, Faculdade de CiÃªncias, Universidade do Porto, UMR7206 Eco-anthropologie, CNRS, MNHN, UniversitÃ© Paris CitÃ©

Human admixture history is rarely a simple process in which distinct populations, previously isolated for a long time, come into contact once to form an admixed population. In this study, we aim to reconstruct the complex admixture histories of the population of SÃ£o TomÃ©, an island in the Gulf of Guinea that was first settled by the Portuguese in the 15th century, was the site of the first slave-based plantation economy, and experienced successive waves of slavery and post-slavery migration from Africa. We examined 2.5 million SNPs newly genotyped in 96 SÃ£o TomÃ©ans and found that geography alone cannot explain the observed patterns of genetic differentiation within the island. We defined five genetic groups in SÃ£o TomÃ© based on the hypothesis that individuals sharing the most haplotypes are more likely to share similar admixture histories. Using IBD and different local ancestry inference methods, we inferred shared ancestries between SÃ£o TomÃ©ans and 70 African and European populations, and we used ancestry tract distributions to reconstruct the admixture histories of each SÃ£o TomÃ©an genetic group separately. We identify admixture events between admixed groups that were previously isolated on the island. This study demonstrates how complex admixture and isolation histories during and after the Transatlantic Slave-Trade shaped extant individual genetic patterns at a local scale in Africa; it explores cases in which recently admixed populations are themselves at the sources of other admixture events; and it further shows the need for extreme caution when assigning anthropological labels to genetic clusters.

780 - Non-parametric representations of structure in genetic ancestry through time

John Novembre

Ahmed Selim, NÃºria Hermosilla, Luke Anderson-TrocmÃ©, John Novembre

University of Chicago, Universitat Pompeu Fabra, University of Chicago, University of Chicago

Understanding population history often requires the assumption of discrete, bounded populations even when the underlying history is more complex. Currently, there is a need for tools that can reveal the structure of genetic ancestry in a time-varying way, without assumptions about whether the structure has discrete units or continuous gradients. Here, we present progress on this problem by describing structure in genetic ancestry through time, leveraging reconstructed Ancestral Recombination Graphs (ARGs). Specifically, we assess how individuals sampled at present descend from sets of genetic ancestors in the past. This approach naturally lends itself to revealing discrete population groupings when they are present and to revealing continuous gradients when the history is more complex. Attractively, the method produces flexible representations of genetic histories, in effect, adding a time dimension to popular low dimensional summary approaches such as PCA. The method is also helpful for revealing the timescale of genetic structure in a way that is more instructive about how genetic structure impacts allelic sharing than many population history models. The method is

individual-based and does not rely on population descriptors, though labels might be used to facilitate interpretation in light of historical processes if that is a focus. Overall, the method is a promising alternative to single snap-shot representations of population structure (e.g. PCA) and to demographic history tools that assume discrete populations. It also has potential as a teaching tool to help reveal structure in the â€œcloud of gene historiesâ€• (W. Maddison) underlying genetic variation.

837 - Breaking Boundaries: Leveraging Continuous Models for Improved Population Genetics Analyses

Luke Anderson-Trocmé

Luke Anderson-TrocmÃ©, Jeremy Guez, Konrad Karczewski, John Novembre
The University of Chicago Department of Human Genetics, Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Analytic and Translational Genetics Unit, Massachusetts General Hospital, Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Analytic and Translational Genetics Unit, Massachusetts General Hospital, The University of Chicago Department of Human Genetics

Human genetic variation is continuous and yet categorical labels are still commonly used in population genetics analyses to describe groups of individuals and to facilitate matching to reference samples. Such approaches necessarily exclude individuals to maintain specificity, result in possibly suboptimal groupings, and lead to decreased power for analysis. Two particular applications that currently adopt discrete groupings are calculating allele frequencies in a reference sample, e.g. for genetic diagnosis, or linkage disequilibrium (LD) matrices, e.g. for fine-mapping. Here, we propose a method to move beyond discrete genetic ancestry groupings towards a continuous model based on genetic similarity to improve accuracy in both these applications.

Our method computes local allele frequencies and LD patterns using a distance-weighted method in a low dimensional representation for samples in a genetic reference panel. Using cross-validation, we compared our method to more traditional categorical approaches by evaluating the error and bias from estimated allele frequencies to the known frequencies of groups of individuals. We show an improvement in predicting allele frequencies and LD patterns in hold-out sets of individuals, and we show how the accuracy of our method outperforms discrete methods at a range of genetic divergences.

We deploy this method in the largest public aggregated database of human exomes and genomes (gnomAD) which is widely used for genetic diagnosis. Our method provides a critical first step in abandoning the use of discrete labels in population genetics and simultaneously improves downstream applications, including fine-mapping and rare disease investigations.

954 - Quantitative Interpretations of Principal Component Analysis

Benjamin M Peter

Benjamin M Peter, Divyaratan Popli
University of Rochester, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology

Principal Component Analysis (PCA) is a widely used method to assess population structure. However, interpreting PCA-plots remains difficult, and is often done qualitatively. In contrast to previous approaches that interpret the principal components themselves, we develop a framework that uses PCA to simplify the geometry of the genotype space, allowing us to link PCA to underlying population genetic models.

We show that population trees, admixture graphs and general migration models all leave clear geometric patterns in a PCA. These patterns can be used for statistical inference and also to directly assess and visualize deviations from the model.

Our results have numerous practical implications. First, we find that different versions of PCA should be used for different purposes; probabilistic PCA enables us to isolate population structure from sampling noise, and is best used to visualize-or correct for-populaton structure. In contrast, ordinary PCA does not make this distinction and is best suited for investigating statistical artifacts and biases. Second, our results give suggestions on how the results of a PCA should be plotted, how many PCs should be used and how the axes should be scaled.
Finally, building on our framework, we are able to generate fast, accurate and flexible tests for admixture, as we can use the full data set rather than a small subset for inference.

In summary, while most population genetic models assume a discrete population structure, the geometric perspective on PCA illustrates how models can be extended to populations that vary in a more continuous manner.

216 - Gene expansions contributing to human brain evolution

Megan Y Dennis

Daniela C Soto, JosÃ© M Uribe-Salazar, Gulhan Kaya, Aarthi Sekar, Alfie Gleeson, Nicholas K Haghani, Keiko Hino, Gabriana N La, Natasha A Mariano, Cole Ingamells, Aidan E Baraban, Li-En Jao, Sergi SimÃ³, Gerald Quon, Aida M AndrÃ©s, Megan Y Dennis
University of California, Davis, University of California, Davis, University of California, Davis, University of California, Davis, University College London, University of California, Davis, University of California, Davis, University of California, Davis, University of California, Davis, University of California, Davis, University of California, Davis, University of California, Davis, University of California, Davis, University College London, University of California, Davis

Genomic drivers of human-specific traits remain largely undiscovered. Duplicated genes expanded uniquely in the human lineage likely contributed to brain evolution, as demonstrated by SRGAP2C and its association with improved sensory-task performance when introduced in mouse models. Challenges exist in discovering duplicated genes due to the paralog similarity making them prone to sequence-assembly errors. Mitigating this issue, we used a complete telomere-to-telomere human genome sequence to identify 491 duplicated gene families comprising likely human-specific paralogs (>98% identity). Using genome sequence and fetal-brain transcriptomic datasets, we narrowed in on 741 brain-expressed paralogs universal across modern humans. This included 42 paralogs co-expressed in modules enriched for autism-associated genes, connecting them with human language and cognition. Further study of ten duplicate gene families using long-read DNA sequencing allowed us to characterize variation across ~200 modern humans of diverse ancestries, showing that a subset of paralogs retain conservation at the amino-acid level and are likely still functional, evolving under strong purifying selection. To understand their roles in brain development, we generated zebrafish CRISPR â€œknockoutâ€• models of orthologs and transiently introduced mRNA-encoding paralogs, effectively â€œhumanizingâ€• the larvae. Morphometric, behavioral, and single-cell RNA-seq screening highlighted a possible new mechanism for SRGAP2C function in synaptogenesis and, for the first time, a role for GPR89B in dosage-mediated brain expansion. Our holistic approach provides new insights and a useful resource to the community for studying gene expansion drivers of human brain evolution.

259 - DNA replication errors are a major source of gene amplification in adaptive evolution

Julie Chuong

Julie N Chuong, Nadav Ben Nun, Ina Suresh, Julia Matthews, Titir De, Grace Avecilla, Farah Abdul-Rahman, Yoav Ram, David Gresham
New York University, Dept. of Biology, Center for Genomics & Systems Biology, Tel Aviv University, New York University, New York University, New York University, New York University, Tel Aviv University, New York University

Copy number variants (CNVs) are an important source of genetic variation underlying rapid adaptation and genome evolution. However, little is known about factors contributing to the structure, formation rate, and fitness effects of CNVs. Local genome elements are likely important determinants. Previously, we found the GAP1 gene in Saccharomyces cerevisiae undergoes repeated amplification and selection under glutamine-limited selection. This gene has a unique genomic architecture consisting of two flanking long terminal repeats (LTRs) and a proximate origin of DNA replication (autonomously replicating sequence (ARS)), which contribute to GAP1 CNV formation. To test the role of these elements on CNV-mediated adaptation, we experimentally evolved in glutamine-limited chemostats strains lacking either the LTRs, ARS, or all elements. Using a CNV reporter and neural network simulation based inference we quantified the rate and fitness effect of CNVs for each genotype. Removal of local DNA elements has minimal impact on the rate or fitness effect of CNVs. We defined the mechanisms of CNV formation for ~40 CNV lineages from each genotype. Regardless of genotype, 49% of CNVs are mediated by the DNA replication based mechanism Origin-Dependent Inverted-Repeat Amplification (ODIRA). In the absence of the local ARS, distal ones can mediate ODIRA CNV formation. Homologous recombination can mediate gene amplification in the absence of LTRs in part initiated by de novo insertion of retrotransposons at the locus. Our study reveals the remarkable plasticity of the genome and demonstrates that DNA replication errors are a predominant source of adaptive CNVs.

306 - Intergenic structural variation and ancient gene duplication underpin pigmentation diversification in swordtail fish

Tristram Dodge

Tristram O Dodge, Daniel L Powell, John J Baczenas, Theresa R Gunn, Manfred Schartl, Molly Schumer
Stanford University, Centro de Investigaciones CientÃ-ficas de las Huastecas "Aguazarca", Stanford University, Centro de Investigaciones CientÃ-ficas de las Huastecas "Aguazarca", Stanford University, Stanford University, University of Wuerzburg, Texas State University, Centro de Investigaciones CientÃ-ficas de las Huastecas "Aguazarca", Stanford University, Centro de Investigaciones CientÃ-ficas de las Huastecas "Aguazarca", Howard Hughes Medical Institute

Pigment patterns provide striking examples of phenotypic diversification and are a tractable system to study the genetic basis and evolution of adaptive traits. While melanic patterns in mammals are best studied, teleost fish present an exciting system due to their high levels of pattern variation and increased pigment cell diversity. Swordtail fish in the genus Xiphophorus are polymorphic for dozens of pigmentation traits under balancing selection, providing excellent opportunities to characterize both genetic architecture and selection. Here, we study four melanic patterns in two diverged Xiphophorus clades. Using GWAS and QTL mapping in five species, we find three distinct patterns on the body and tail map to intergenic regions upstream of kitlga, a gene involved in melanocyte migration. RNAseq and allele-specific expression data show increased kitlga expression in these specific tissues, and the nonoverlapping GWAS peaks hint that distinct regulatory elements may drive each pattern. Surprisingly, we find tailspot phenotypes in a second Xiphophorus clade map to kitlgb, a gene that arose by duplication 350 million years ago. Using long-read sequencing, we identify and resolve intergenic structural variants perfectly associated with each trait. Pairwise alignment and pangenomic approaches show these structural variants are complex and often include multiple duplications. Haplotype comparisons across species reveal unexpected patterns of allele sharing and divergence, suggesting prominent roles of introgression and allelic turnover. Our results suggest that ancient duplications can set the stage for recent phenotypic diversification, and intergenic structural variants may play an under-appreciated role in gene regulation and adaptation.

611 - The role of structural variation in clownfish adaptive radiation

Diego A. Hartasánchez

Anna Marcionetti, Diego A. HartasÃ¡nchez, Nicolas Salamin
University of Lausanne, University of Lausanne, University of Lausanne

Clownfishes are a complex of 28 species that rapidly diversified following the acquisition of mutualism with sea anemones around 15 Mya. While the genomic signatures associated with this adaptive radiation have started to emerge, little attention has been paid to structural genomic variation in clownfish. We know, however, that structural variation can disrupt gene function and regulation and alter gene dosage, and that it has been demonstrated to play a central role in adaptive evolution and species diversification in multiple taxa. Hence, we have investigated the number, type and role of structural variation across the clownfish phylogeny, paying particular attention to chromosome 18, which has recently been suggested to harbor two large inversions. Using long sequencing reads, we obtained high-quality genomes for several clownfish species spanning the main lineages identified in the genus. We found a large number of insertions, deletions, duplications, and inversions that occurred during clownfish diversification. We were able to reconstruct and date their evolutionary history, confirming the presence of large inversions on chromosome 18. We show that hybridization events during clownfish diversification might have contributed to generating the striking pattern of presence/absence of one of these large inversions across the genus, and suggest a possible link between this inversion and clownfish phenotype, in turn associated with anemone choice. This study represents the first evidence of considerable structural variation in clownfish genomes. It also provides new, high-quality, valuable genomic resources for pursuing research on this emerging model system and on its adaptation to sea anemones.

1014 - Molecular evolution of a maize hybrid barrier over 12 million years suggests epistatic silencing

Elli Cryan

Elli Cryan, Garnet A Phinney, Arun S Seetharam, Matthew M S Evans, Daniel J Kliebenstein, Jeffrey Ross-Ibarra
UC Davis Plant Sciences, UC Davis Evolution & Ecology, UC Davis Evolution & Ecology, Iowa State University Ecology, Evolution and Organismal Biology, Iowa State University Genetics, Development and Cell Biology, Iowa State University Genetics, Development and Cell Biology, Carnegie Institution for Science, UC Davis Plant Sciences, UC Davis Evolution & Ecology

Maize and its wild relatives teosinte can all readily hybridize, yet they remain distinct. Mating incompatibility loci provide one mechanism that can allow populations to maintain reproductive isolation. In Zea mays, three complex mating incompatibility loci encode genes that disrupt directional pollen tube growth down the silk. When a maternal plant receives pollen with incompatible alleles, fertilization is impeded, creating a prezygotic reproductive barrier. However, this barrier is not complete. Infrequent fertilization of incompatible gametes facilitates introgression of incompatibility genes into other populations. Previous modeling shows that each factor should only undergo brief periods of strong selection, on a timescale shorter than time to speciation in this clade. Against this expectation of transient benefit, these loci are hypothesized to have aided in speciation. All of the loci display presence absence variation and copy number variation. We used public and new genome sequences to reconstruct the evolutionary history of these complex loci. We classified haplotype diversity at all three loci in over 30 Zea mays genomes, identified syntenic loci in related species, constructed gene trees of known and candidate functional genes, and analyzed rates of molecular evolution. We find evidence of potentially functional reproductive barrier loci and genes in lineages estimated to be twelve million years diverged. Over millions of years, these genes have undergone selection and evolved to be distinct barriers, which is reflected in predicted protein structures. We also find evidence of loci driving epigenetic silencing of interacting loci. These loci may have played a role in Zea diversification.

1039 - Convergent Structural Variation of Antifreeze Proteins in Polar and Deep Sea Fishes

Samuel N Bogan

Samuel N. Bogan, Nathan Surendran, Scott Hotaling, Isabel Kline, Muh-Ching Yee, Thomas Desvignes, Iliana Bista, Luana S.F. Lins, Mari H. Eilertsen, Nathalie R. Le FranÃ§ois, Tait Algayer, Scott L. Hamilton, Paul B. Frandsen, Federico G. Hoffmann, Joanna L. Kelley
University of California, Santa Cruz, University of California, Santa Cruz, Utah State University, Washington State University, University of California, Santa Cruz, University of California, Santa Cruz, University of Oregon, Senckenberg LOEWE Center for Translational Biodiversity Genomics, Australian National Insect Collection, CSIRO, Washington State University, Centre for Deep Sea Research, University of Bergen, Laboratoire de Physiologie, Aquaculture et Conservation, Division des Collections Vivantes, Conservation et Recherche, BiodÃ´me de MontrÃ©al, Washington State University, San Jose State University, Moss Landing Marine Laboratories, Brigham Young University, Mississippi State University, University of California, Santa Cruz

Zoarcoidei is the only suborder of fishes adapted to both the Arctic and Southern Oceans and is uniquely valuable for studying convergent environmental adaptation. A key innovation permitting zoarcoid polar invasions was the evolution of type III antifreeze protein (AFP III). There is evidence that zoarcoids adapted to colder environments possess more AFP copies. It is unknown (i) whether AFP copy number variation (CNV) has convergently evolved during adaptation to both poles and (ii) what genomic mechanisms influence AFP CNV. Addressing these questions requires long read sequencing to obtain accurate measures of AFP copy number, synteny, and structural variation. We sequenced and assembled high quality, long read genomes for nine zoarcoids including deep-water, shallow, temperate, and polar species. Leveraging these assemblies and five other zoarcoid long read genomes, we integrated oceanographic measures of temperature, salinity, and pressure across depth-informed species ranges to model AFP gene family evolution as a function of temperature and pressure, which jointly affect freezing point. AFP copy number convergently increased in species from colder habitats and decreased in deep-water species under high pressure. High AFP copy number in cold, shallow water species was driven by expansions of AFPs independently translocated out of an AFP tandem array in which ancestral AFP III arose. AFP CNV positively correlated with transposable element (TE) density in AFP tandem arrays, and TE density was greater in arrays of translocated AFPs. Our findings suggest roles for translocation and AFP CNV facilitated by TEs during convergent adaptation to extreme cold.

120 - Exploring the Rock Dove’s (Columba livia) Genetic Diversity: Insights into the Species’ Evolutionary History and Conservation Challenges

Germán Hernández-Alonso

GermÃ¡n HernÃ¡ndez-Alonso, JazmÃ-n Ramos-Madrigal, Hein van Grouw, Marta Maria Ciucani, Emily Louisa Cavill, Mikkel-Holger S. Sinding, Shyam Gopalakrishnan, George Pacheco, M. Thomas P. Gilbert
Department of Organismal Biology, Uppsala University, The Globe Institute, University of Copenhagen, Department of Life Sciences, Natural History Museum, The Globe Institute, University of Copenhagen, The Globe Institute, University of Copenhagen, Department of Biology, University of Copenhagen, The Globe Institute, University of Copenhagen, The Globe Institute, University of Copenhagen, The Globe Institute, University of Copenhagen

The domestic pigeon (Columba livia) is one of the most important model organisms, playing a key role in developing Darwin's Theory of Evolution. In contrast, its parental species, the rock dove, remains largely understudied, and questions regarding its evolution, taxonomy, and conservation status remain unresolved. Additionally, experts warn of the rock doveâ€™s imminent risk of extinction due to genetic replacement by the almost ubiquitous feral pigeon. In order to describe, for the first time, the rock doveâ€™s genetic diversity prior to the feral pigeon worldwide expansion, we generated whole-genome sequencing data from 65 historical samples representing all recognized rock dove subspecies. Remarkably, the characterization of the rock dovesâ€™ diversity and gene-flow patterns revealed a highly divergent population from the West of Africa that potentially represents a different species. We hypothesize that this population diverged as a result

of a long history of allopatric cycles, which occurred during the quaternary glacial and interglacial periods. We also detect high levels of inbreeding and low heterozygosity across the West African rock doves, highlighting the need for its conservation status revaluation. Finally, our historical dataset revealed admixture signals between domestic/feral pigeons and wild rock doves, confirming that the expansion of the feral birds, favoured by ongoing urbanization and environmental deterioration, represents the main threat to the rock dove populations. This study represents an important step toward the conservation of the species, allowing the identification of genetically relevant populations on which to focus future conservation strategies to ensure the rock doveâ€™s preservation.

352 - Phylogenomics and water isotopes unveil the evolutionary history of how cacti colonized the hyperarid Atacama Desert

Pablo C. Guerrero

Pablo C. Guerrero, Jose Garcia-Beltran, Sandy Toledo

Departamento de Botanica, Universidad de Concepcion, Instituto de Ecologia y Biodiversidad , Millennium Institute Biodiversity of Antarctic and Sub-Antarctic Ecosystems, Departamento de Botanica, Universidad de Concepcion, Instituto de Ecologia y Biodiversidad, Departamento de Botanica, Universidad de Concepcion, Instituto de Ecologia y Biodiversidad

The study delves into how lineages adapt to harsh environments, focusing on the hydrological niche evolution within the diverse cactus genus Eriosyce, spread across a 2,300 km arid gradient in southwestern South America. By integrating Whole Genome sequencing and Water Isotope analyses (2H and 18O), we explored the genus's phylogenomic relationships and hydrological adaptations. This approach allowed us to reconstruct time-calibrated phylogenies, assess evolutionary shifts in hydrological niches and functional traits, and infer historical biogeographical patterns concerning the Atacama Desert occupation. Comparison with geological data illuminated the origins of these desert ecosystems. The study identified the water sources (rainfall, groundwater, dew, fog) utilized by Eriosyce, noting variations in niche breadth and isotopic signals linked to species' microhabitats and growth forms. A notable delay was observed in the colonization and subsequent adaptation to hyperarid areas by Eriosyce subclades, suggesting a migration route from milder, Mediterranean-like central Chile to harsher desert conditions. These insights underscore the gradual and intricate nature of desert adaptation, highlighting the significant survival challenges species face amid accelerating climate change.

639 - Reviving from four birds: post-bottleneck overview of temporal genomics of Mauritius kestrel

Xuejing Wang

Xuejing Wang, Jim Groombridge, Cock van Oosterhout, Ken Norris, Simon Tollington, Vikash Tatayah, HernÃ¡n Eduardo Morales

GLOBE Institute, University of Copenhagen, Division of Human and Social Sciences, Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, School of Environmental Sciences, University of East Anglia, Centre for Agri-Environmental Research, School of Agriculture, Policy and Development, University of Reading, North of England Zoological Society, Chester Zoo, Durrell Institute of Conservation and Ecology, University of Kent, Mauritian Wildlife Foundation, GLOBE Institute, University of Copenhagen

In many species, population declines can damage gene pools through the loss of genetic diversity and increased genetic load. This process, known as genomic erosion, increases the extinction risk, even after populations recover from a bottleneck. The Mauritius kestrel is one of the most extreme examples of population decline, having experienced a population bottleneck of four individuals in 1974. Through intensive conservation management, including captive-breeding programs and artificial nest boxes, the population rebounded to 350-400 individuals in the early 2000s and declined to <250 individuals. Here we analyze a whole-genome temporal dataset of 59 Mauritius kestrels from 1986 to 2021, supplemented with genomes from six outbreed and inbreed sister species. Demographic inference

indicated that the kestrel population was small even before human arrival to Mauritius. Despite marked differences in population viability, we detected no genetic divergence between the subpopulations on the eastern and western parts of Mauritius. Their remaining genetic diversity is extremely low, with two mitochondrial genotypes found in early samples (1986-1995) and one genotype remaining in recent individuals (2012-2021). Genome-wide heterozygosity has remained stable, though at a minimal level â€” 5% of that found in widely distributed kestrel species. Runs of homozygosity (ROH, â‰¥1 Mb) constitute over 40% of the genomes, and the fraction of long ROHs (>5Mb) increased in recent times, indicating ongoing inbreeding within the recovered population. Our findings underscore the dramatic impact of the bottleneck on deleterious genetic variation, highlighting the population's heightened vulnerability to future challenges.

1056 - Reconstructing the population history of the endangered Hispaniolan solenodon, one of the most evolutionary distinct mammals on earth

Maria H Zicos

Maria H Zicos, Selina Brace, Laurent Frantz, Samuel Turvey

Natural History Museum, Queen Mary University of London, Natural History Museum , Ludwig Maximilian University of Munich, Queen Mary University of London, Institute of Zoology at the Zoological Society of London

Island systems like the Caribbean have been focal in researching evolutionary patterns and processes. Islands are particularly relevant for reconstructing colonisation and radiation events, but also for preserving ancient populations and biodiversity with important conservation implications for understanding the evolutionary history of insular species. The insular Caribbean was colonized by non-volant land mammals that diversified into more than 120 different species, however the region suffered from a massive Holocene extinction event and few terrestrial mammalian species remain. The extant fauna includes two representatives of the Solenodontidae, the Hispaniolan solenodon (Solenodon paradoxus) and Cuban solenodon (Atopogale cubana). Solenodons are nocturnal insectivores and the only living mammals with dental venom delivery systems, forming an ancient lineage that diverged from other living mammals c.73 million years ago. Solenodons are considered endangered and coupled with their evolutionary distinctiveness and current threats to survival, both species have been identified as global priorities for mammal conservation.

To explore the population history of these high priority species this study has generated six new solenodon genomes using ancient DNA methodologies: five for the Hispaniolan solenodon, one of which is the first genome for the previously described south-western population, and the first genome for the Cuban solenodon. Our genome-wide data supports structure into three populations as previously suggested from mitochondrial only data, with limited gene flow between biogeographical regions of the island. Further, low genome-wide heterozygosity was identified, comparable to that of endangered eulipotyphan relatives. A more contiguous genome assembly will allow investigations into genomic erosion.

370 - MBD2/3 lost its methyl-CpG binding ability in multiple families of Holometabola

Elisa Israel

Elisa Israel, Zoe Marie LÃ¤nger, Joachim Kurtz, Sonja Prohaska

Leipzig University, University of MÃ¼nster, University of MÃ¼nster, Leipzig University

DNA methylation is sparse in insects, this reduction is even more pronounced in Holometabola. In some Holometabola, DNA methylation is lost entirely. Methyl-CpG-binding domain (MBD) proteins bind methylated DNA and are therefore important readers of these epigenetic marks. We hypothesize that evolutionary reduction of genome-wide methylation results in changes to MBD proteins. Only one MBD family member, MBD2/3, is known in insects. Two isoforms have been identified in Bombyx mori. The long isoform MBD2/3-L contains a complete MBD domain spanning the first two exons, whereas the short isoform MBD2/3-S lacks the second exon and therefore half of its

MBD domain. It has been reported that only the MBD2/3-L isoform is able to bind methyl-CpGs.
In this study, we analyzed transcriptomic and genomic sequence data from holometabolan orders to identify MBD2/3 genes and isoforms. Furthermore, we analyzed transcripts of representative species from each insect order. Overall, MBD2/3 is highly conserved in sequence and gene structure, with splice sites being identical within each order. Both isoforms are present in most insect orders. However, in Coleoptera and Hymenoptera, the longer isoform MBD2/3-L, capable of binding methylated DNA, has been lost entirely. This loss is also observed in several other holometabolan families and species.
The results suggest that MBD2/3 has lost its binding ability in Coleoptera and Hymenoptera. Furthermore, we propose that MBD2/3-L has been lost several times independently but exclusively in Holometabola. The reduced levels of DNA methylation may be linked to these losses.

624 - Molecular determinants of transgenerational epigenetic variation at transposable element sequences in plants and its impact in nature

Pierre Baduel

Pierre Baduel, Vincent Colot
IBENS / CNRS, IBENS / CNRS

Transgenerational epigenetic inheritance (TEI) has garnered increasing interest, yet this additional system of inheritance remains highly controversial. In plants, several examples of true epiallelic variants (i.e. that are inherited independently of any DNA sequence changes), have been described and all involve DNA methylation changes over transposable element (TE) sequences. However, the precise determinants of TEI are still unknown in plants, and so are its extent and functional impact in nature.

To address this issue, we have undertaken a stringent and comprehensive genome-wide analysis of DNA methylation variation at TE sequences in an experimental population where severe hypomethylation was experimentally-induced as well as in hundreds of natural strains of the plant Arabidopsis thaliana. Collectively, our results indicate that TE-associated epiallelic variation is relatively abundant in nature where it is shaped by an antagonism between RNA-directed DNA methylation limiting its inheritance on the one hand and transcriptional activity favoring it on the other, and as a result is found preferentially over divergent and low copy-number TE sequences located near stress-responsive genes. In addition, we will discuss the different mechanisms through which the environment also exerts a major influence and their consequences for understanding the adaptive role that TEI may play in plants.

867 - Exploring the evolutionary dynamics of epigenetic switches and beyond: Insights from gene regulatory circuit models

Mariana Gómez-Schiavon

Mariana GÃ³mez-Schiavon
Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano (LIIGH), UNAM

Epigenetic switches have long fascinated evolutionary biologists as key players in adaptation to fluctuating environments. However, our research now extends beyond their traditional role. We investigate the broader landscape of dynamic properties in gene regulatory circuits, utilizing mechanistic models to explore evolutionary processes shaping these circuits. Through in silico evolution, we uncover how dynamic properties emerge, propagate, and persist under natural selection, shedding light on the structural dependencies within circuits and their cellular contexts. While maintaining a focus on epigenetic switches, we delve into other emergent properties such as plasticity and oscillations, examining their selective advantages and competition both among themselves and with classical genetic

adaptation. Our findings reveal fundamental trade-offs, such as the balance between adaptation time and phenotypic robustness to biochemical noise, driving the selection of epigenetic switches in fast fluctuating environments. This research not only enriches our understanding of epigenetic mechanisms but also provides insights into broader biological phenomena. Our investigations hold implications for diverse fields, from bacterial antibiotic resistance to plant survival strategies in extreme environments. Through this work, we aim to contribute new perspectives on the complexity of life and the underlying principles that govern it.

1048 - The role of CTCF in the transcriptional activation of the hematopoietic-specific gene HBA2 is evolutionary conserved

Gustavo Tapia

Gustavo Tapia-UrzÃºa, Hober Nelson NÃºÃ±ez-MartÃ-nez, Martin Franke, Sylvia Garza-Manero, Ã•ngel JosuÃ© Cerecedo-Castillo, Carlos Alberto Peralta-Ã•lvarez, Georgina Guerrero, FÃ©lix Recillas-Targa
Institute of Cellular Physiology, UNAM, Institute of Cellular Physiology, UNAM, Andalusian Centre for Developmental Biology (CABD), Institute of Cellular Physiology, UNAM, Institute of Cellular Physiology, UNAM, Institute of Cellular Physiology, UNAM, Institute of Cellular Physiology, UNAM, Institute of Cellular Physiology, UNAM

CTCF is a DNA binding protein that regulates chromatin organization and gene expression, by interactions with transcription factors and co-factors, and by promoting long-range interactions of specific regulatory elements. While the architectural role of CTCF has been further characterized, its relationship with DNA methylation has been less explored. Here, we show that CTCF binding to tissue-specific regulatory elements is dynamic during erythroid differentiation and that a subset of bindings sites delimits differentially methylated regions. CTCF enrichment, therefore, protects against propagation of DNA methylation, promotes chromatin accessibility, and binding of the erythroid-specific transcription factor GATA-1, in particular loci. In agreement, deletion of a CTCF binding site within the promoter of the chicken globin gene HBAD, and its orthologous in humans HBA2, results in CpG methylation propagation to the promoters, which is coupled to transcription factor binding reduction and lower transcriptional levels of HBAD and HBA2, respectively. These changes were not explained by disruption of interactions between cis-regulatory elements of the chicken Î±-globin domain, since no major changes in chromatin contacts were observed. Therefore, there are relevant CTCF mechanisms beyond of its architectural role that are conserved in specific loci among vertebrates.

377 - Viral and host metagenomics of 100-year-old insect museum specimens

Alexandra H Keene

Alexandra H Keene, Mark D Stenglein
Colorado State University, Colorado State Univeristy

Recovery of virus sequences from old samples provides an opportunity to reconstruct historic virus-host interactions. Most studies of old viruses have used DNA or RNA from frozen or fixed samples. The millions of specimens in natural history museum represent a trove of old virus sequences, but itâ€™s not clear how well RNA survives in these samples. Here, we experimentally assessed the stability of RNA in dried insects over 72 weeks. Although RNA molecules from experimentally dried samples grew increasingly fragmented over time, RNA levels remained constant. RT-qPCR of host and virus RNA targets showed minimal differences between RNA in dried and frozen specimens. Next, we acquired D. melanogaster museum specimens from entomological collections across the United States to assess RNA survival in much older specimens. We recovered coding complete and partial sequences from known and novel viruses including several coding complete virus genomes from a fly collected in 1908. We found that the virome of D. melanogaster has remained relatively stable over the past century. Galbut virus, the most prevalent virus infection in contemporary D. melanogaster, was also the most

common in historic samples. Finally, we investigated the genomic and physical features of surviving RNA and found that surviving RNA was chemically damaged, and apparently protected by RNA-RNA interactions or RNA-protein interactions. Overall, we concluded that RNA is stabler than typically thought. This highlights the utility of the millions of dried museum specimens to provide a clearer understanding of historic virus-host interactions and virus evolution.

471 - The trait specific timing of accelerated genomic change in the human lineage.

Eucharist Kun

Eucharist Kun, Mashaal Sohail, Vagheesh Narasimhan
The University of Texas at Austin, Universidad Nacional AutÃ³noma de MÃ©xico, The University of Texas at Austin

Humans exhibit distinct characteristics compared to our primate and ancient hominin ancestors including bipedal locomotion and enhanced neurocognitive ability, but the timing of accelerated changes in these traits is uncertain. To investigate if specific trait-associated variation show enrichment during particular periods of human evolution, we combine genome wide association study (GWAS) data from 70 traits, spanning multiple categories including AI-based image-derived morphological phenotypes of the brain, heart, and skeletal tissues with data from 12 different evolutionary regions obtained from comparative functional genomics, multi-species alignments from long read sequencing, and ancient DNA reflecting 4 different major evolutionary divergence points. These regions cover epigenetic differences in the brain between humans and rhesus macaques, various human accelerated regions (HARs) including regions from the Zoonomia Project, ancient selective sweeps, and Neanderthal introgressed alleles. Using two complementary approaches to examine enrichment between GWAS loci and genomic regions, we found that traits across respiratory, dermatological, reproductive, and metabolic categories were enriched through multiple timepoints, along with skeletal and brain traits, consistent with striking morphological changes between humans and other primates. We also find that different brain structures associated with various functions were enriched at different evolutionary depths, including the longitudinal fasciculus in human-gained epigenetic elements since macaques, the visual cortex in HARs, and the thalamus proper in Neanderthal introgressed alleles. Lastly, we show that more phenotypes are enriched in the periods of divergence between macaques and humans or chimps and humans, and less so during the divergence with Neanderthals.

539 - Evolutionary paths to thermal resilience in placental mammals: functional genomics insights from the Frozen Zoo Biobank

Lucas Rocha Moreira

Lucas Rocha Moreira, Carla Madelaire, Kristin Crouse, Valary Forbes, Allyson Hindle, Elinor Karlsson, Diane Genereux
University of Massachusetts Chan Medical School, Broad Institute of Harvard and MIT , San Diego Wildlife Zoo Alliance, University of Minnesota, Florida Atlantic University, University of Nevada, University of Massachusetts Chan Medical School, Broad Institute of Harvard and MIT , Broad Institute of Harvard and MIT , Broad Institute of Harvard and MIT

Scattered about the mammalian phylogeny are species whose cells endure dramatic temperature variation that would harm human cells. Hibernators, such as little brown bat and 13-lined ground squirrel, drop to refrigerator-level body temperatures every winter, and desert species such as dromedary camel, undergo daily thermal cycling. Have cellular-resilience mechanisms in these species emerged independently under similar selective pressures, or do relevant genes and pathways overlap, suggesting shared evolutionary history? Powered by resources from our Zoonomia comparative genomics project and from San Diego Zooâ€™s Frozen ZooÂ®, the world's only institution to bank cells of more than 1,000 taxa, we are using a classical common-garden framework to discover the evolutionary history of cellular robustness. We cultured primary dermal fibroblasts of 12 deeply diverged mammals, including both strict thermoregulators and

heterotherms, and assessed responses across a range of temperatures. We identified several pathways universally impacted by temperature perturbation, including regulation of cholesterol biosynthesis, expression of heat shock proteins, and maintenance of the cell cycle. We found evidence for convergent shifts in gene expression in 1,054 genes in heat-flexible species under hot treatment (41Â°C) and 875 genes in hibernator species under cold treatment (32Â°C). Under hot treatment, dromedary camel transcriptome exhibited widespread post-transcriptional changes, including a large shift in isoform usage of HSF1, an early responder to heat shock, highlighting a potential role for alternative splicing. Our comparative framework holds great promise for unraveling the mechanisms of cellular robustness and discovering therapeutic avenues for diseases of compromised homeostasis.

998 - From swampy ancestors to modern maize: tracing the allopolyploid origins of Zea

Michelle C Stitzer

Michelle C Stitzer, Taylor AuBuchon-Elder, Thuy La, Robert Bukowski, Qi Sun, M. Cinta Romay, Edward S Buckler, Elizabeth A Kellogg
Cornell University, Donald Danforth Plant Science Center, Cornell University, Cornell University, Cornell University, Cornell University, Cornell University, USDA-ARS, Donald Danforth Plant Science Center

Early cytological, genetic, and genomic studies hypothesized maize originated from an allopolyploidy event, due to the presence of duplicated regions throughout the genome. Most lines of evidence suggested the parents of the allopolyploidy belonged to the highly productive Andropogoneae tribe of grasses, today containing ~1,200 species. Limited genomic sampling of these non-model, often polyploid, grasses has left considerable gaps in our understanding of the evolutionary history of the maize genome. Here, we use extensive herbarium and field collections to sample grasslands globally, where Andropogoneae species cover nearly 20% of terrestrial land. We generate whole genome short read sequence (~15x depth) of over 500 species, and phylogenetically targeted PacBio assemblies of 28 species. We reconstruct phylogenetic relationships for 9,500 conserved gene loci, and estimate divergence of transposable elements. Approximately half of loci place the maize lineage sister to subtribe Rhytachninae, which is today found in wet environments in Africa, growing in river banks and seasonally swampy environments. The other subgenome donor is more obscure, with contributions from extant Ratzeburgiinae and Elionurus, each of which has a few species present today in the Americas. We interpret preferential retention and stronger purifying selection on gene copies inherited from the wet Rhytachninae ancestor alongside biogeographic reconstructions that show the polyploid originated in Central America as evidence that the newly formed allopolyploid succeeded in a novel ecological niche. We conclude that considering the subgenome ancestry of maize can be helpful when interpreting modern maize diversity, especially for traits like drought and flooding tolerance.

70 - The nature of the Last Universal Common Ancestor and its impact on the early Earth system

Edmund R R Moody

Edmund R R Moody, Sandra Ã•lvarez-Carretero, Tara A Mahendrarajah, James W Clark, Holly C Betts, Nina Dombrowski, LÃ©nÃ¡rd L SzÃ¡nthÃ³, Richard A Boyle, Stuart Daines, Xi Chen, Nick Lane, Ziheng Yang, Graham A Shields, Gergely J SzÃ¶llÅ‘si, Anja Spang, Davide Pisani, Tom A Williams, Timothy M Lenton, Philip C J Donoghue
University of Bristol, University of Bristol, Royal Netherlands Institute for Sea Research, University of Bath, University of Bristol, Royal Netherlands Institute for Sea Research, EÃ¶tvÃ¶s University, Centre for Ecological Research, University of Exeter, NA, University of Exeter, University College London, University College London, University College London, University College London, Okinawa Institute of Science and Technology Graduate University, EÃ¶tvÃ¶s University, Centre for Ecological Research, NA, Royal Netherlands Institute for Sea Research, University of Amsterdam, NA, University of Bristol, NA, University of Bristol, University of Exeter, University of Bristol

The nature of the Last Universal Common Ancestor (LUCA), its age, and impact on the Earth system have been the subject of vigorous

debates across diverse disciplines, often based on disparate data. Age estimates for LUCA are usually based on the fossil record, varying with every reinterpretation. The nature of LUCAâ€™s metabolism has proven equally contentious, with some attributing all core metabolisms to LUCA, while others reconstruct a simpler life form dependent on geochemistry. Here we show, using a set of pre-LUCA duplications under a new cross-bracing approach implemented in PAML, calibrated with microbial fossils and isotope records of metabolisms, that LUCA existed ~4.2 Ga. We use 700 prokaryotic genomes to infer a 2.59Mb+ genome encoding at least 2500 proteins, comparable to modern prokaryotes. Our results suggest LUCA was anaerobic, acetogenic and possessed an early immune system. While LUCA is sometimes perceived as living in isolation, we infer LUCA to have been part of an established ecology. The metabolism of LUCA would have provided a niche for other community members, while hydrogen recycling by atmospheric photochemistry could have supported a modestly productive early ecosystem.

151 - Disentangling the evolutionary impacts of relatedness and facultative/obligate life cycles during the transition to multicellularity

Autumn Peterson

Autumn Peterson, Anthony Burnetti, William C Ratcliff

Georgia Institute of Technology, Georgia Institute of Technology, Georgia Insitute of Technology

Complex multicellularity evolved in five lineages, and in each case, these organisms develop clonally and are obligately multicellular. While prior work has shown that clonal development plays a critical role in the evolution of complex multicellularity, none has disentangled this from the impact of obligate vs facultative multicellular life cycles. Theory predicts that selection acting on single cells may impede group-level adaptation, but due to a lack of model systems, this has never been tested. We circumvent this constraint by engineering a strain of Saccharomyces cerevisiae to grow as single cells or clonally-multicellular groups â€˜snowflakesâ€™ upon induction with Î²-estradiol. We evolved 20 replicate populations of one ancestral genotype with four different life cycles: always unicellular, always multicellular, or alternating between uni and multicellularity (either spending 50% or 75% of their time in a multicellular state). Only the obligately multicellular treatment evolved â€˜entrenchingâ€™ traits: those which increase group-level fitness at a cost to cell-level fitness. For example, obligately multicellular snowflake yeast convergently evolved to be tetraploid, increasing group size and fitness but reducing cellular growth rates. We show that while tetraploidy should be adaptive in the facultatively multicellular treatments, it never evolved, because selection in the unicellular stage is far more effective than the multicellular stage (~100-fold higher Ne). Our results demonstrate that an obligately multicellular life cycle plays a key role in the transition to multicellularity independent from its effect on genetic relatedness, by avoiding disproportionately powerful cell-level counterselection against novel multicellular traits.

552 - Eco-evolutionary feedbacks promoted by hypermutation rate in a predator-prey microcosm

Pu Wang

Pu Wang, Michael Travisano

University of Minnesota, University of Minnesota

Ecological and evolutionary dynamics can operate at the same time. The eco-evolutionary feedback could intensify the adaptation and lead to further diversification of both communities and environments. Eco-evolutionary dynamics are dependent on genetic variation. Mutation, the source of genetic variation, could play a role in influencing eco-evolutionary dynamics. Our project used a predator-prey microcosm with the ciliate Tetrahymena thermophila consuming Pseudomonas fluorescens bacteria. Five Pseudomonas strains with distinctive pre-existing genetic variation including mutator and maladapted genetic constraints were subjected to predatory selection. We observed the intensification of adaptation in prey populations due to fitness landscape changes via eco-evolutionary feedback. The

evolutionary responses of the five strains were compared against each other to determine the roles of genetic constraints and mutation rates in eco-evolutionary interactions. The pre- and post-predation population genome sequencing data revealed that evolved populations with lower genetic diversities had higher phenotypic diversities. These results reflect the challenges of genotype-phenotype mapping. We genome-sequenced certain colonies with new phenotypes to investigate the underlying evolutionary mechanisms. We found that potential adaptations could be limited by genetic constraints, which could be overcome by higher mutation rates. The results suggested mutation rates had a role in determining the evolutionary trajectories, thus mediating the eco-evolutionary feedback.

562 - A new genome from the pangenome - experimental evolution with horizontal gene transfer.

Mike J McDonald

Michael J McDonald
Monash University

Horizontal Gene Transfer (HGT) is a defining characteristic of microbial evolution in natural microbial communities. While HGTs importance in microbial evolution has been clearly demonstrated by comparative genomics, there are still only a handful of studies that have incorporated HGT into evolution experiments. We have carried out a series of experiments capturing the dynamics and fitness effects of horizontally transferred genetic variants in experimental populations in the naturally competent bacteria H. pylori and A. baylyi. I would like to present evolution experiments, genome sequence and large-scale sequencing fitness assay data that shows how recombination can be beneficial. In particular, I will discuss how Fisher-Muller dynamics act in microbial populations, how HGT can reverse the evolution of antibiotic resistance, and present exciting new work showing how we can carry out experimental tests of pangenome evolution.

679 - Experimental evolution of synthetic phototrophic communities

José Flores-Uribe

JosÃ© Flores-Uribe, RubÃ©n Garrido-Oter
Max Planck Institute for Plant Breeding Research, Max Planck Institute for Plant Breeding Research

Microalgae release organic compounds into their surroundings that support specific bacterial communities known as the phycosphere microbiota. In exchange, these communities provide benefits to their host by supplying nutrients or enhancing resistance to environmental stressors. The recent sequencing and culturing of phycosphere bacteria from the microalga Chlamydomonas reinhardtii has enabled the design of synthetic microbial communities for long-term gnotobiotic experiments. Despite these advances, we lack a mechanistic understanding of how these and other complex communities evolve and interact over extended periods. To address this, we developed a reductionist experimental system to explore long-term ecological interactions, metabolic exchanges, and strain-specific adaptation processes in synthetic phototrophic communities under varying environmental conditions over prolonged periods of time.

Here, we report the results of a long-term artificial evolution experiment, where multiple independent synthetic communities were grown for over 500 days using photobioreactors. By combining amplicon sequencing and metagenomics with high-resolution phenotypic data, we have generated one of the most extensive longitudinal host-microbiome datasets to date, demonstrating that long-term artificial selection experiments using complex synthetic communities are feasible. Our analysis reveals distinct ecological networks of co-occurrence and co-exclusion, likely driven by metabolic interactions among community members. Furthermore, based on sequencing results and analysis of re-isolated bacteria, we identified mutations associated with nutrient uptake and antimicrobial compound production, among others, demonstrating strain-specific adaptation processes across different taxa. Our findings provide insights into the

ecological factors that contribute to the evolutionary landscape of host-associated microbial communities.

342 - The evolutionary origins of autoimmune and infectious disease risks

Evan K. Irving-Pease

Evan K. Irving-Pease, William Barrie, Eske Willerslev, Lars Fugger, Astrid K. N. Iversen
University of Copenhagen, University of Cambridge, University of Copenhagen, University of Cambridge, University of Bremen, University of Oxford, University of Oxford, University of Copenhagen

Autoimmune diseases are one of the leading causes of chronic illness, and their prevalence has been rising for decades. Our understanding of the genetic architecture of autoimmunity has been greatly advanced by large genome-wide association studies (GWAS), which have shown that the majority of risk variants are commonly occurring and many are located within or near immune genes. Many of these risk variants also confer protection against infectious diseases, suggesting that natural selection may have had to balance protection from pathogens against increased risk of autoimmunityâ€”in a process known as â€˜antagonistic pleiotropyâ€™. Previously, we have shown that genetic risk for multiple sclerosis was subject to strong polygenic selection in Eurasia at a time when human populations experienced increased exposure to pathogens. Here, we model the evolutionary history of ten autoimmune diseases (including type 1 diabetes, coeliac disease and psoriasis). Using a panel of >1,000 ancient genomes, we inferred the allele frequency trajectories and selection coefficients of genetic risk variants for each autoimmune disease, while accounting for the complex demographic history of prehistoric Eurasia. Our results reveal that many autoimmune diseases show evidence of strong polygenic selection during recent human evolution. Using a joint statistical test designed to measure the impact of selection on genetically correlated pairs of traits, we tested for pleiotropic interactions between each autoimmune disease and more than 30 infectious disease GWAS phenotypes. Our results reveal the complex evolutionary trade-offs which have shaped present-day disease risk for both autoimmune and infectious diseases.

570 - Rapid evolution of mutation rate and germline maintenance under strong pathogen selection: new insights from experimental evolution experiments

Imroze Khan

Basabi Bagchi, Dipendra Nath Basu, Joy Bose, Srijan Seal, Imroze Khan
Ashoka University, Ashoka University, Ashoka University, Ashoka University, Ashoka University

Pathogenic infections and costly inflammatory responses are major regulators of mutation susceptibility, imposing challenging threats on health, but the evolutionary and demographic causes and consequences of such impacts are unclear. Here, we took a novel integrated approach aimed at expanding our understanding of how infection and immunity play a causative role in modulating mutation rates and their transmission across generations. Using tractable insect models (e.g., Drosophila melanogaster and Tribolium castaneum), we first show that immune activation and pathogenic infections can consistently increase the rate of deleterious (dominant and recessive) mutations in subsequent generations by compromising germline maintenance. Resource limitation or poor-quality diet can further amplify the effects of such mutation transmission. However, a long-term evolution under strong pathogen selection can drastically reduce the accumulation of deleterious mutation load in immune-challenged and infected individuals by improving germline maintenance. We could also identify the possibility of increased activity of multiple DNA repair pathways (e.g., homologous recombination, mismatch repair and non-homologous end joining) in the ovary of evolved females that contributes to their improved germline maintenance, using large-scale transcriptome analyses and RNAi. Together, these results can provide a novel conceptual framework of how variation in immune responses can, in principle, be linked to mutation load and germline maintenance, generating fundamental insights into what drives

natural variation in organismal health and resistance to pathogens across generations.

870 - Epistasis and pleiotropy constrain and expand the evolution of SARS-CoV-2 Omicron lineage

Alief Moulana

Alief Moulana, Thomas Dupic, Michael Desai

Harvard University, Harvard University, Harvard Universityt, Harvard University

The Omicron lineage of SARS-CoV-2 has undergone extensive evolution, resulting in varied disease manifestations. The receptor binding domain (RBD) of the viral spike protein, has acquired numerous mutations, affecting its stability, ACE2 affinity, and immune evasion capability. These mutations' effects are often modulated by epistatic interactions and pleiotropy among the phenotypes they influence, complicating evolution prediction. Our study addresses this by systematically measuring ACE2 and antibody binding of more than 100,000 relevant variants. Particularly, we created four RBD variant libraries to explore the protein landscapes: (i) all BA.1 mutation combinations on the Wuhan Hu-1 strain; (ii) full spectra of BA.1, BA.2, BA.4, and BA.5 mutations; (iii) common Omicron mutations on divergent backgrounds; and (iv) random mutations on these backgrounds. First, we highlighted a compensatory epistasis in which specific BA.1 mutation combinations on the Wuhan Hu-1 background can mitigate the negative impact on ACE2 binding caused by immune escape mutations. Analysis of the second and third libraries illustrates how subvariant mutations that contribute to antibody escape diversity harbor independent effects on ACE2 binding, albeit with significant pleiotropic tradeoffs between ACE2 affinity and immune escape. Further, multiple mutations deletrious to ACE2 binding on the BA.1 background are well-tolerated on the BA.2 background, possibly contributing to the rise of BA.2 over BA.1. Together with more incoming data on random mutations, our findings further elucidate the tradeoffs between binding affinity and immune escape and the constraints presented by epistasis, providing valuable insights into viral evolution and potential therapeutic targets.

928 - Interactions between sickle cell and Plasmodium falciparum genotypes in asymptomatic malaria

Helena Hopson

Helena Hopson, Alejandra Herbert Mainero, Heejung Shim, Sandrine E Nsango, Ellen M Leffler

Department of Human Genetics, University of Utah School of Medicine, Department of Human Genetics, University of Utah School of Medicine, Melbourne Integrative Genomics, University of Melbourne, Malaria Research Unit, Centre Pasteur du Cameroun, Department of Biological Sciences, FacultÃ© de MÃ©decine et des Sciences Pharmaceutiques, UniversitÃ© de Douala, Department of Human Genetics, University of Utah School of Medicine

The sickle cell allele (HbS) is protective against severe malaria caused by Plasmodium falciparum, but the mechanism remains unclear. Intriguingly, multiple nonsynonymous mutations in parasite genes have recently been found to be enriched in HbS carriers who do develop severe malaria, suggesting these changes allow parasites to overcome the protective effect of HbS. However, it is also unclear why this would be beneficial to the parasite since HbS does not protect against asymptomatic infection, and the majority of infections in HbS carriers are asymptomatic. To begin to understand the functional and fitness consequences of these parasite genotypes, we investigate the impact of parasite genotype and host HbS status on gene expression and mosquito transmission in blood samples from 70 children with asymptomatic infections from Cameroon. Using dual RNA-seq, we identify host and parasite genes differentially expressed between HbS and parasite genotypes. We find differential expression of genes that transport parasite virulence proteins to the red blood cell surface, supporting their reduced export as a likely key mechanism for HbS protection. We also find evidence for epistatic interactions on

transmission quantified by membrane feeding assays, suggesting the fitness of parasite genotypes is linked to HbS genotype. We hypothesize this may have led to a balanced polymorphism in parasites in which HbS is the selective pressure. Analysis of this ongoing interaction offers a unique approach to reveal key mechanisms of HbS protection and co-evolutionary dynamics between host and parasite genetic variation.

1071 - Deep learning phenotype imputation to measure dominance of rare variant effects on rare disease in UK Biobank

Daniel M Jordan

Daniel M Jordan, Ben Omega Petrazzini, Robert Chen, Joshua K Park, Iain S Forrest, Ha My T Vy, Ã•ine Duffy, Ghislain Rocheleau, Ron Do

Icahn School of Medicine at Mount Sinai, Icahn School of Medicine at Mount Sinai, Icahn School of Medicine at Mount Sinai, Icahn School of Medicine at Mount Sinai, Icahn School of Medicine at Mount Sinai, Icahn School of Medicine at Mount Sinai, Icahn School of Medicine at Mount Sinai, Icahn School of Medicine at Mount Sinai, Icahn School of Medicine at Mount Sinai

Recent studies on the genetic architecture of human traits and diseases have suggested that traits thought to be recessive or dominant may actually have a range of intermediate dominance. However, it has been difficult to test this hypothesis on rare variants or rare diseases, as non-additive genetic association studies require very large sample sizes to be powered. We have previously developed a deep learning phenotype imputation approach that allows us to conduct association studies on rare phenotypes (prevalence < 1 in 2,000) for aggregated ultra-rare coding variants (minor allele frequency < 10^-5). We report recessive and additive ultra-rare variant burden association tests for 155 rare phenotypes conducted using REGENIE in approximately 500,000 UK Biobank participants. Even with the lower power of recessive model association tests, we find many more significant disease-associated genes with the recessive model than the additive model (143 recessive vs. 39 additive), as well as a substantial bias of sub-significant effects towards recessivity. We find no clear examples of intermediate dominance, in contrast to results in common-variant and common-disease regimes. Overall, our findings suggest that the effects of rare variants on rare disease are largely recessive, with an additional significant role for additivity and minimal intermediate dominance. Testing these hypotheses has not previously been possible, providing a powerful example of how state-of-the-art deep learning methods applied to population-scale biobanks can reveal insights about the genetic architecture of human traits and diseases.

1080 - Radiation-induced stress and signatures of selection on anti-tumor immunity in Chornobyl wolves

Cara Love

Cara Love, Stacey Lance, Brian Arnold, Nicolas Rochette, Thomas Hinton, James Beasley, Bridgett vonHoldt, Michael Byrne, Dmitry Shamovich, Brian Nadel, Sarah Webster, Shane Campbell-Staton

Princeton University, University of Georgia, Princeton University, Princeton University, Norwegian University of Life Sciences , Fukushima University, University of Georgia, Princeton University, University of Missouri, EcoTur Belarus, UCLA, University of Georgia, Princeton University

Investigating the complex interplay between contaminant exposure, physiological responses, and evolutionary changes in wild populations is critical for understanding the impacts of anthropogenic environmental change. Such investigations can shed light on the broader health implications of these changes and may provide novel models for examining adaptive responses to disease causative agents. We explore the consequences of chronic multigenerational radiation exposure on gray wolves (Canis lupus) in the Chornobyl

Exclusion Zone (CEZ) as a model for examining adaptive anti-tumor immunity. By analyzing patterns of regulatory and genetic variation, we uncover molecular signatures of immune stress and adaptive evolution. Whole blood transcriptome data reveal significantly altered PBMC populations among Chornobyl wolves and regulatory divergence at both single gene and co-regulatory levels in CEZ wolves. These results highlight radiation-induced immune modulation, cellular apoptosis, and antitumor immune response as potential physiological consequences of radiation stress and targets of selection within the CEZ population. We further examine regulatory decoherence and genomic signatures to identify signatures of selection within the CEZ wolf population. Combining these techniques, we identify several targets of selection overlapping genes influential in cancer physiology. These genes encompass crucial functions such as anti-tumor immunity, cellular invasion, and migration. Functional modeling of selective targets further suggests significant alterations in proteins crucial for regulating immune response and cancer physiology.

104 - Bayesian Inference of Admixture Graphs on Native American and Arctic Populations

Andrew Vaughn

Andrew H Vaughn, Svend V. Nielsen, Kalle LeppÃ¤lÃ¤, Michael J. Landis, Thomas Mailund, Rasmus Nielsen

UC Berkeley, Aarhus University, University of Oulu, Aarhus University, Washington University in St. Louis, Aarhus University, UC Berkeley, University of Copenhagen

Admixture graphs are a widely used tool in population genetics to infer historical relationships from samples of genetic data. Admixture graphs provide a concise description of the historical demographic relationships between genetic samples of populations as a directed acyclic graph representing population splits and mergers. Inferring graph topologies, however, involves a combinatorial search, and since the space of graphs grows super-exponentially in the number of populations and the number of admixture events, an exhaustive search is typically not possible. Instead, existing methods that search for well-fitting graphs utilize greedy algorithms, which are liable to get stuck in local maxima. Furthermore, these methods are unable to report uncertainty in these estimates as they only yield point estimates of graphs. We here improve on these approaches by developing a novel MCMC sampling method, AdmixtureBayes, that can sample from the posterior distribution of admixture graphs. This enables an effective search of the entire state space as well as the ability to report a level of confidence in the sampled graphs. We apply AdmixtureBayes to a set of Native American and Arctic genomes to reconstruct the demographic history of these populations and report posterior probabilities of specific admixture events. In particular, we explore the sources of gene flow into the modern Athabascan and Inuit populations from distinct ancient Siberian and Native American populations.

310 - Complex neutral processes drive the evolutionary fate of Neanderthal alleles in 30,780 admixed genomes with African-like and European-like ancestry

Aaron Pfennig

Aaron Pfennig, Joseph Lachance

Georgia Institute of Technology, Georgia Institute of Technology

Following introgression, Neanderthal DNA was initially purged from modern human genomes but has remained constant at ~2% in Europe during the last 45,000 years. Here, we leverage whole-genome sequences of 30,780 unrelated, recently admixed individuals with African-like and European-like ancestry in All of Us to test the evolutionary fate of Neanderthal alleles. In these admixed genomes, Neanderthal variants were recently introduced into an African genetic background, leading to a reevaluation of their present-day fitness. Surprisingly, we observe more Neanderthal ancestry in such admixed genomes than expected under neutrality and a simple demographic model. However, this enrichment is driven by segments also observed in African genomes, due in part to incomplete lineage sorting. When focusing on introgressed segments uniquely contributed by European-like ancestors, the observed enrichment disappears.

Nevertheless, we identify 405 and 359 genomic regions significantly enriched or depleted for Neanderthal ancestry, respectively. Preliminary functional and population genetic analyses of these regions reveal enrichments for GWAS hits associated with known Neanderthal phenotypes and pleiotropic eQTLs but do not support strong recent selection. Simulations under a plausible demographic model suggest that complex neutral processes can explain the initially observed enrichment and outlier regions, underlining the importance of appropriate demographic models as nulls in population genetics. We also identified four emerging introgression deserts, whose analyses may provide insights into the evolutionary dynamics leading to previously reported archaic deserts (i.e., simple negative selection vs. hybrid incompatibilities). Altogether, Neanderthal DNA appears to behave largely evolutionary neutral in extant human genomes.

411 - Pervasive gene flow despite strong and varied reproductive barriers in swordtails

Daniel Powell

Stepfanie Aquillon, Sophia Haase Cox, Quinn Langdon, Theresa Gunn, JJ Baczenas, Shreya Banerjee, Alex Donny, Ben Moran, Carla GutiÃ©rrez-RodrÃ-guez, Oscar RÃ-os-CÃ¡rdenas, Molly Morris, Daniel Powell, Molly Schumer
University of California Los Angeles, Stanford University, Centro de Investigaciones CientÃ-ficas de las Huastecas â€œAguazarcaâ€•, A.C., Stanford University, Gladstone Institute of Virology, Stanford University, Centro de Investigaciones CientÃ-ficas de las Huastecas â€œAguazarcaâ€•, A.C., Stanford University, Centro de Investigaciones CientÃ-ficas de las Huastecas â€œAguazarcaâ€•, A.C., Stanford University, University of California Davis, Stanford University, Stanford University, Stanford university, Centro de Investigaciones CientÃ-ficas de las Huastecas â€œAguazarcaâ€•, Instituto de EcologÃ-a A.C., Instituto de EcologÃ-a A.C., Ohio University, Stanford University, Centro de Investigaciones CientÃ-ficas de las Huastecas â€œAguazarcaâ€•, A.C., Stanford University, Centro de Investigaciones CientÃ-ficas de las Huastecas â€œAguazarcaâ€•, A.C., Howard Hughes Medical Institute

The formation of new species occurs through the evolution of reproductive barriers. However, recent research has demonstrated that hybridization has been pervasive across the tree of life despite the presence of strong barriers. Swordtail fish (genus Xiphophorus) are an emerging model system for studying the interface between these barriers and hybridization. We document overlapping mechanisms that act as barriers between closely related species, X. birchmanni and X. cortezi, by combining genomic sequencing from natural hybrid populations, artificial crosses, behavioral assays, sperm performance, and developmental studies. We show that strong assortative mating plays a key role in maintaining subpopulations with distinct ancestry in natural hybrid populations. Lab experiments demonstrate that artificial F1 crosses experience dysfunction: crosses with X. birchmanni females were largely inviable and crosses with X. cortezi females had a skewed sex ratio. Using F2 hybrids we identify several genomic regions that strongly impact viability. Strikingly, two of these regions were previously identified as genetic incompatibilities in hybrids between X. birchmanni and its sister species X. malinche. Our results demonstrate that ancient hybridization has played a role in the origin of this genetic incompatibility. Moreover, the incompatibility causes extraordinarily similar survival effects in X. cortezi ï‚´ X. birchmanni hybrids as in X. malinche ï‚´ X. birchmanni hybrids. Our findings highlight complex evolutionary outcomes of hybridization in the face of varied reproductive barriers and historical hybridization between species.

546 - Genomic Insights into the Recurrent Interspecific Hybridizations of Agave-Associated Saccharomyces Yeasts

Lucia Morales Reyes

J. Abraham Avelar Rivas, Erick Santiago GarduÃ±o, IvÃ¡n Eduardo SedeÃ±o JimÃ©nez, Luis Fernando GarcÃ-a Ortega, Eugenio Mancera, Alexander DeLuna, LucÃ-a Morales

Unidad de GenÃ³mica Avanzada, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Irapuato, MÃ©xico, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano (LIIGH), Universidad Nacional AutÃ³noma de MÃ©xico, QuerÃ©taro, MÃ©xico, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano (LIIGH), Universidad Nacional AutÃ³noma de MÃ©xico, QuerÃ©taro, MÃ©xico, Departamento de IngenierÃ-a GenÃ©tica, Unidad Irapuato, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Irapuato, MÃ©xico, Departamento de IngenierÃ-a GenÃ©tica, Unidad Irapuato, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Irapuato, MÃ©xico, Unidad de GenÃ³mica Avanzada, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Irapuato, MÃ©xico, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano (LIIGH), Universidad Nacional AutÃ³noma de MÃ©xico, QuerÃ©taro, MÃ©xico

Interspecific hybridization events have shaped the genomes of various organisms across Eukarya. Yeast hybrids, recognized for their weak prezygotic barriers and prevalence in human-associated environments, are valuable models for studying the effects of hybridization on speciation, adaptation, and genome evolution. In this work, we isolated and sequenced over 60 hybrids of Saccharomyces cerevisiae and its sister species, S. paradoxus, from traditional agave distilleries and their natural surroundings across Mexico. These hybrids are prevalent in this environment, accounting for up to 10% of the total sequenced isolates and rising to 40% in semi-arid regions, suggesting that climatic factors may influence hybrid formation or survival. We will present comparative analyses of the nuclear and mitochondrial hybrid genomes which revealed numerous distinct genomic architectures, indicating multiple hybridization events and varying stages of genome stabilization. This study contributes to our understanding of the evolutionary dynamics of interspecific hybridizations, providing a valuable resource to study their impact on speciation and adaptation.

732 - Alternative genomic and biochemical pathways to red carotenoid pigments in birds

Alexander Kirschel

Alexander N. G. Kirschel, Sifiso Lukhele, Matteo Sebastianelli, Sophia C. Hayes, LoÃ¯s Rancilhac, Alan Brelsford, Bridgett M. vonHoldt
University of Cyprus, University of Cyprus, Uppsala University, University of Cyprus, University of Cyprus, University of California Riverside, Princeton University

Red pigments in animals are typically shaped by sequestration of carotenoids acquired from the diet and enzymes that catalyze the conversion of dietary yellow carotenoids to red ketocarotenoids. Several candidate genes have been found to function in this pathway, including CYP2J19 and BCO2, but a large-scale study in nature of the relative effect sizes of candidate genes is lacking. Furthermore, the relative effect on phenotypes of different dietary carotenoids present in the diet is little known. To investigate these questions, using 202 whole genomes, we conducted a genome-wide association analysis to identify the genes underpinning variation between yellow and red forecrown feathers across a hybrid zone of Pogoniulus tinkerbirds. We identified a locus of large effect spanning several genes including CYP2J19, CYP2J40, HOOK1, FGGY and NF1A that explained 80% of the variation in forecrown color, with heterozygotes at this locus all converting carotenoids and exhibiting red hues. We then compared the extent of gene expression in liver tissue across the contact zone to confirm the relative effects of genes within this locus. Moreover, we investigated the pathways resulting in red forecrown feathers in disjunct populations of tinkerbirds across Africa. We determined the extent to which red forecrown color is ancestral, has evolved in parallel, or might be explained by another pathway involving a different dietary carotenoid distribution identified using Raman spectroscopy and high performance liquid chromatography. Our study reveals how red pigmentation in birds may result from alternative genomic and biochemical pathways of carotenoid metabolism.

913 - Detecting selection in admixed populations using haplotypic information

Maria Ines Fariello Rico

Maria Ines Fariello, GastÃ³n Rijo, Bertrand Servin, Hugo Naya
CICADA, Universidad de la RepÃºblica, IMERL, Facultad de IngenierÃ-a, Universidad de la RepÃºblica, Instituto Pasteur de Montevideo, Institut Pasteur ParÃ-s, INRAe, ., Instituto Pasteur de Montevideo, Facultad de AgronomÃ-a, Universidad de la RepÃºblica

Detecting selection in admixed populations can be challenging, especially when the admixture is not considered in the model explicitly. It is possible to mistake selection signatures with haplotypes that exist in high frequency due to admixture instead of selection. To untangle these selection signatures from other biological processes, we suggest extending hapFLK, a test designed to detect selection in population trees using haplotypic information from multiple populations. Estimating the variance-covariance matrix of the allele frequencies is crucial for accurately estimating the distribution under neutrality in this test. In the original test, the variance-covariance matrix was obtained from the branch lengths of the Neighbor-joining population tree, which considers the evolutionary framework.

To extend this test to admixed populations, we suggest three different strategies: use hapFLK as it is, based on trees; estimate the variance-covariance matrix using TreeMix; and estimate the variance-covariance matrix empirically. The last procedure could be thought of as a way of estimating branch lengths based on F-statistics. For the empirical estimation, we estimate the frequency of the ancestral alleles first based on the Neighbor-Joining tree and then the covariances of each pair of populations.

Our simulations showed that the best strategy is to use the empirical variance-covariance matrix to detect selection. However, we could not achieve the same power as when using the theoretical variancecovariance matrix. Further work should be focused on the matrix estimation. Finally, as previous work has shown, detecting hard sweeps is easier than detecting soft sweeps.

333 - Cross-species comparison of gene regulatory networks using single cell sequencing

Ines Hellmann

Anita TÃ©rmeg, Fiona Edenhofer, Johanna Geuder, Wolfgang Enard, Ines Hellmann
Ludwig-Maximilians University , Ludwig-Maximilians University , Ludwig-Maximilians University , Ludwig-Maximilians University , Ludwig-Maximilians University

Many phenotypic differences between primate species must be established during development. Hence, in order to understand them we need to investigate the underlying gene regulatory networks (GRNs) and their evolution. Single-cell RNA-seq methods finally allow us to reliably infer GRNs. However, most comparative analyses of GRNs have so far focused on qualitative aspects - a quantitative comparison as required for evolutionary analysis remains challenging.

We designed a computational pipeline to perform quantitative cross-species network comparisons to identify conserved and diverged gene modules and applied it to scRNA-seq data of an early neuronal differentiation experiment, with iPS cells from human and cynomolgus macaque. Surprisingly, we found the network of the otherwise conserved pluripotency regulator OCT4 among the most diverged networks. Closer inspection of the genes that show the highest degree of differential regulation included most prominently two genes under the control of LTR7 elements.

A completely different approach is to infer GRNs from perturbation screens. To this end, we established inducible KRAB-dCas9 iPSC lines for humans and macaques. In our first CRISPRi screen, we target 27 transcription factors for each of which we designed multiple guide RNAs, allowing for a quantitative comparison between the species. Using the log fold changes between the perturbed and unperturbed conditions as edge weights, we can then construct GRNs and apply our established statistics to find diverged and conserved

co-expression modules. This analysis could e.g. confirm the most diverged genes of the OCT4 network identified in unperturbed cells.

689 - The molecular bases of red-to-yellow color variation in parrots

Roberto Arbore

Roberto Arbore
CIBIO - Centro de InvestigaÃ§Ã£o em Biodiversidade e Recursos GenÃ©ticos, InBIO LaboratÃ³rio Associado, University of Porto, Washington University in St. Louis

Parrots produce stunning plumage colors by the endogenous synthesis of a unique class of pigments known as psittacofulvins. A polyketide synthase has been shown to be required for psittacofulvin biosynthesis, but the molecular mechanisms explaining variation in psittacofulvin coloration remain to be elucidated. We found that red and yellow colors in parrots result from the differential deposition during feather development of two classes of psittacofulvins differing by a simple chemical modification. Then, we integrated genetic mapping with gene expression and functional experiments to identify a housekeeping enzyme as the key regulator of a red/yellow polymorphism in wild parrot populations. Finally, through fine mapping and single-cell genomic techniques (scRNA-seq and snATAC-seq) we traced this polymorphism to a single point mutation in a conserved regulatory element exclusive to differentiating keratinocytes in developing feathers. The simplicity of the proposed genetic and enzymatic mechanisms offers an explanation for the exceptional evolutionary lability of psittacofulvin-based colors throughout parrot evolution. Our study illuminates the molecular mechanisms governing the colorful plumage of parrots, offering insights into the evolutionary origins of key innovations via the recruitment of genes with essential cellular functions.

822 - Quantifying adaptive evolution of the human immune cell landscape

Jesús Murga-Moreno

Irepan Salvador-Martinez, JesÃºs Murga-Moreno, Juan C Nieto, Clara Alsinet, David Enard, Holger Heyn
CNAG, University of Arizona, CNAG, CNAG, Arizona, CNAG

The human immune system is under constant evolutionary pressure, primarily through its role as first line of defence against pathogens. Accordingly, population genomics studies have shown that immune-related genes have a high rate of adaptive evolution. These studies, however, are mainly based on protein-coding genes without cellular context, leaving the adaptive role of cell-types states uncharted. Inferring the rate of protein-coding genes adaptation in human immune cells at cellular resolution, we found cell-types from lymphoid and myeloid compartments to harbour significantly increased adaptation rates (eg. foetal Pre-Pro B cells and adult Trm CD8+). Focusing on the Lung, we found an enrichment of cell-types with high adaptation in barrier tissues, suggesting an adaptive response to external challenges such as respiratory pathogens. We further analysed iPSC-derived macrophages responding to various challenges, including pro- and anti-inflammatory cytokines or bacterial and viral infections, the latter simulating the evolutionary arms race between humans and pathogens. Here, we found adaptation in early immune responses, suggesting host benefits to adapt to early infection stages to control pathogen spread. Together, our study reveals spatio-temporal and functional biases in human immune populations with evidence of rapid adaptive evolution, providing a retrospect of forces that shaped the complexity, architecture, and function of the human body.

832 - Characterizing neuronal evolution between recently diverged species at single-cell resolution

Taylor L. Cooper

Taylor L. Cooper, Haowen He, Kathryn N. Leatherbury, Nikesh M. Kumar, Jeffrey T. Streelman, Patrick T. McGrath

Georgia Institute of Technology, Georgia Institute of Technology, Georgia Institute of Technology, Georgia Institute of Technology, Georgia Institute of Technology, Georgia Institute of Technology

Evolution of neuronal cell types contributes to the remarkable diversity of behaviors between species. The cellular heterogeneity of vertebrate brains and complex regulation of gene expression within these cell types can now be analyzed using single nucleus RNA- and ATAC-sequencing. To understand how neural cell types evolve, we utilize simultaneous multi-omic techniques in the Lake Malawi cichlid flock, a classical example of an evolutionary radiation that produced >800 species distinguished by social and feeding behaviors which allowed the exploit of diverse ecological niches. We investigate cell type differences in three brain regions, the telencephalon, hypothalamus, and pituitary, across 15 species (one male and one female each) representing four of the seven major ecogroups in the lake. We find differences in cell-proportion in behaviorally relevant neurons, supporting a role for neurogenesis in the evolution of these species. We also identify differentially expressed genes within conserved cell types that distinguish lineages within the lake. Further, we will discuss the association of large putative â€˜supergenesâ€™ that segregate within the lake with chromatin accessibility and gene regulation. Our work provides a framework to understand the remarkable diversity of social and feeding behaviors that characterize this impressive adaptive radiation.

917 - Evolutionary impact of transcriptomic differences in X and Y carrying gametes

Meritxell Riera

Meritxell Riera, Sofia Boeg Winge, Marjoline Rouselle, Florent Murat, Noe Mbengue, Samuele Soraggi, Christina Hvilsom, Henrik Kaessmann, Kristian Almstrup, Mikkel Heide Schierup
Bioinformatics Research Centre, Aarhus University, Department of Growth and Reproduction, Rigshospitalet, University Hospital, Bioinformatics Research Centre, Aarhus University, Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, Heidelberg University,, Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, Heidelberg University,, Bioinformatics Research Centre, Aarhus University, Copenhagen Zoo, DK 2000 Frederiksberg, Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, Heidelberg University, Department of Growth and Reproduction, Rigshospitalet, Copenhagen University Hospital, Bioinformatics Research Centre, Aarhus University

Spermatogenesis is a highly complex developmental process by which germ cells divide and differentiate to generate haploid spermatozoa. Genes expressed in the testis show signals of rapid evolution in mammals and apes, most likely driven by differences in the reproductive success of spermatozoa. Recently, in a comparative study of single-nuclei RNA-seq (snRNA-seq) data in mammals we showed that this accelerated rate of molecular evolution is driven by genes expressed in late spermatogenesis, and in particular haploid spermatids. Haploid cells differ in whether they carry an X or a Y chromosome, which can lead to conflict over transmission of sex chromosomes to the next generation. In this study, we compare the transcriptomic profiles of haploid cells carrying an X or Y chromosome using snRNA-seq data, which we spatially validate using single-molecule RNA in situ hybridization (smISH). We find both lineage specific differences and conserved expression patterns across primates which we correlate with signals of molecular evolution. Notably, we identify genes specifically expressed in X-carrying spermatids that show a human specific increase in post-meiotic gene expression and strong signals of natural selection. These genes are also found in Neanderthal deserts, suggesting that they could have been associated with reproductive barriers between humans and Neanderthals when they met and interbred around 60,000 years ago. Our study uncovers the post-meiotic expression of sex-linked genes across primates and provides valuable insights into possible cases of sexual antagonistic effects that shaped human evolution.

466 - How should we report genetic matches following an Investigative Genetic Genealogy search?

Feriel Ouerghi

Feriel Ouerghi, Dan E Krane, Nathan Adams, Michael D Edge
University of Southern California, Wright State University, Forensic Bioinformatics, University of Southern California

Traditional forensic genetics practice in the United States has historically relied on short tandem repeat (STR) analysis for DNA profiling, as implemented in the Combined DNA Index System (CODIS). One widely used statistic for summarizing the evidence provided by a match between STR genotypes is the random match probability (RMP), the estimated probability that a random person would produce an STR profile matching a forensic sample. Over the past seven years, a new approach called investigative genetic genealogy (IGG) has been instrumental in solving hundreds of criminal cases. IGG differs from standard forensic-genetic practice in that it is based on inferences of genealogical relatedness from genome-wide SNPs rather than matching at STR loci. Currently, suspects identified through IGG are genotyped at CODIS (STR) markers, and a summary of the evidence from the match, such as an RMP or a likelihood ratio (LR) is subsequently reported in court. However, the calculated statistics do not account for the fact that the suspect was initially identified through an IGG match. We demonstrate that, since an IGG search potentially includes identical-by-descent (IBD) segments over the CODIS markers, the probability of a CODIS match increases with relatedness. We also study how population structure, isolation by distance, and assortative mating increase the probability of CODIS matches, even in the absence of overlap between the IBD segments used in IGG and the CODIS markers. Based on this, we argue for a conservative framework for evaluating RMPs and LRs following identification via IGG.

753 - Highly accurate predictions indicating sexual activity using microbiome-based analyses in forensic settings

Meghna Swayambhu

Meghna Swayambhu, Sarah Ahannach, Sarah Lebeer, Eirik NatÃ¥s Hanssen, Lars Gustav Snipen, Fardin Javanmard, Larissa Walser, Rolf KÃ¼mmerli, Natasha Arora
Zurich Institute of Forensic Medicine, University of Zurich, Department of Bioscience Engineering, University of Antwerp, Department of Bioscience Engineering, University of Antwerp, Oslo University Hospital, Norwegian University of Life Sciences, Zurich Institute of Forensic Medicine, University of Zurich, Zurich Institute of Forensic Medicine, University of Zurich, Department of Quantitative Biomedicine, University of ZÃ¼rich, Zurich Institute of Forensic Medicine, University of Zurich

Advances in sequencing technologies and bioinformatic tools have shown the potential of microbiome-based analyses for several applied settings like healthcare and forensics. In forensics, particularly the body-site specificity of bacterial communities has been exploited in various studies investigating body fluid identification. Typically, all these studies focus on sequencing specific variable regions of the 16S rRNA gene and use machine learning tools to report predictions. However, different studies focus on different stretches of the 16S rRNA gene, limiting the integration of datasets. In addition, the classifiers presented in recent studies lack systematic testing on forensically relevant settings like samples deposited on substrates, mixtures, and aged samples. In this study, we combined data from studies using different regions of the 16S rRNA gene with a closed-reference OTU clustering approach to train a random forest classifier. Subsequently, we systematically tested the classifier on mixtures and substrate samples generated in a controlled manner in the laboratory and on uncontrolled forensic settings. Uncontrolled forensic settings included samples from couples with extensive metadata on the body fluids to be expected in the sample, last intercourse, and showering habits among other information. We obtained high prediction probabilities particularly for high biomass samples like saliva, skin from hand and vaginal swabs. In addition, the classifier could predict whether sexual intercourse in couples was penetrative, oral, or both. Our results demonstrate the exciting potential of using a classifier trained on a heterogeneous training dataset in forensic settings especially in sexual assault cases.

1041 - The recovery of DNA from burned forensic contexts.

Anne C Stone

Cody Parker, Amber Coffman, Erin Rawls, Adam Ben Rohrlach, Katelyn Bolhofner, Giovanna Vidoli, Joanne Devlin, Anne Stone
Arizona State University, Arizona State University, Arizona State University, University of Adelaide, Max Planck Institute for Evolutionary Anthropology, Arizona State University, University of Tennessee, University of Tennessee, Arizona State University

Innovation in molecular methods provides opportunities to increase the likelihood of successful downstream DNA identification in challenging forensic contexts such as those with thermal alteration. Since ancient DNA research is heavily invested in optimizing recovery of DNA from challenging samples under similar contexts (e.g., low yields of highly fragmented/degraded DNA), we used and evaluated aDNA as well as forensic extraction methods for 36 paired burned bone and overlying tissue samples that had been classified into different levels of burning based on bone characteristics (i.e. color, cracking). For 109 bone samples, we also constructed both double stranded and single stranded DNA libraries and then used these for targeted enrichment of the mitochondrial DNA genome (H. sapiens Representative Global Diversity Panel) and genome-wide single nucleotide polymorphisms (FORCE-v2, Arbor BioSciences). We found that charred tissue samples when available consistently returned higher concentrations of both total DNA (Qubit HS DNA assay and Agilent TapeStation D5000 HS) and endogenous DNA (Quantifilier Trio). In addition, these assays were not a reliable predictor of successful DNA recovery, likely due to high levels of co-extracted inhibitors. These inhibitors also affected NGS results, reducing recovery in burn category 2.5 but lessening in categories 3 and 3.5. DNA recovery and mapped reads dropped significantly at temperatures greater than 350C (burn categories 4 and 5). Our results show that innovative molecular methods can improve DNA recovery from burned samples for human identification; however, we will also discuss the barriers that exist to implementing these in routine forensic analyses.

1097 - Construction of Epigenetic Clock using Cell-Free DNA

Maria Flores

Maria Flores, Matteo Pellegrini
University of California Los Angeles, University of California Los Angeles

In recent years, DNA methylation for the use of investigative leads in forensic cases, specifically age estimation, has become promising and is increasingly being studied. Several studies have shown that about one-third of methylated sites in the genome are affected by age, leading to the attempted usage of age-related differentially methylated regions for the prediction of chronological age or otherwise known as â€œepigenetic age clocksâ€•. While many of these studies have developed their own mathematical models for epigenetic age prediction, most methods have only considered optimal coverage samples particularly in whole blood and saliva. Here, we present an epigenetic clock using a binomial estimator approach from whole genome bisulphite sequencing data of plasma (cell-free DNA) retrieved from pregnant women that will be used as a proxy for samples that can be found in forensic casework due to its varying degrees of coverage and sparseness. Preliminary results show that of the eight samples that were randomly selected, six of the samples were accurately assessed for a fair accuracy rate of 75%. The model presented here serves to provide a proof of principle so that a similar model may be considered for sparse forensic samples as degraded DNA is becoming more frequently analyzed among law enforcement.

485 - Heterogeneous genetic architectures and evolutionary genomics of prostate cancer in Sub-Saharan Africa

Rohini Janivara

Rohini Janivara, Wenlong Carl Chen, Ujani Hazra, Timothy Rebbeck, Joseph Lachance
Georgia Institute of Technology, Georgia Institute of Technology, University of the Witwatersrand, Georgia Institute of Technology, Dana

Farber Cancer Institute, Harvard T.H. Chan School of Public Health, Georgia Institute of Technology

Men of African descent face disproportionately high rates of prostate cancer (CaP) incidence and mortality, yet research into its evolutionary underpinnings in African populations remains limited. Additionally, CaP risk predictors built using inherited variants often exhibit poor performance across populations due to the inherent genetic differences shaped by human evolutionary history. Here, we explore the evolutionary foundations of CaP associations within sub-Saharan African populations. Leveraging genomic data from 3,963 CaP cases and 3,509 controls across Ghana, Nigeria, Senegal, South Africa, and Uganda, we unravel ancestry-specific genetic architectures and fine-map disease associations. Our analysis reveals fifteen independent associations, including four novel ones, at 8q24.21, 6q22.1, and 11q13.3, all reaching genome-wide significance. Strikingly, many lead disease associations are effectively monomorphic outside of Africa. Furthermore, we ascertain that the genetic architecture of CaP varies significantly across Africa, with effect size differences playing a more prominent role than allele frequency disparities. Integrated haplotype score statistics indicate no discernible selection acting on CaP within African populations. Comparing population branch statistics of CaP variants between Europe and Africa, we find that heterogeneity in CaP's genetic architecture predominantly arises from genetic drift rather than natural selection. Importantly, allele age estimates highlight the contributions of both African polymorphisms that are due to recent mutations and older variants that predate the out-of-Africa migration. Collectively, our findings underscore how evolutionary history has contributed to differences in the genetic architectures and hereditary disease burdens of a complex trait across human populations.

536 - Genomic Insights into the population history of southern-central Africa

Jeremy Choin

Jeremy Choin, Maggie Katongo, Kendra Sirak, Elizabeth Sawchuk, George Mudenda, Potiphar Kaliba, Lovemore Mazibuko, Menno Welling, Alan Morris, Kathryn de Luna, Nadin Rohland, Mary Prendergast, Jessica Thompson, David Reich
Department of Human Evolutionary Biology, Harvard University, Livingstone Museum, Department of Human Evolutionary Biology, Harvard University, Department of Anthropology, Stony Brook University, Livingstone Museum, Malawi Department of Museums and Monuments, Malawi Department of Museums and Monuments, Amsterdam University of the Arts, Department of Human Biology, the University of Cape Town, Department of History, Georgetown University, Department of Genetics, Harvard Medical School, Department of Anthropology, Rice University, Department of Anthropology, Yale University, Department of Genetics, Harvard Medical School, Department of Human Evolutionary Biology, Harvard University

The present-day genomic, cultural, and linguistic landscapes of southern-central Africa (Malawi and Zambia) were drastically changed by a rapid population expansion from western to southern Africa that started after ~3500 years ago. These Bantu speakers brought new technologies and new forms of subsistence to regions that were previously occupied by hunter-gatherers. The details of these interactions are especially poorly resolved in southern-central Africa, which was also historically involved in long distance trading networks. Here, we generated ~70 new genome-wide data from the south-central Africans who lived within the last ~1500 years. We investigated admixture history using an extended version of Chromopainter (PHCP) adapted to work on unphased pseudo haploid data. First, we demonstrated that PHCP could serve as a viable alternative/complement to the commonly employed qpAdm software. Utilizing PHCP, we found that the genomes of southern-central Africans were mostly composed of Bantu-related ancestry and varying proportions of hunter-gatherer-(5%-15%), Nilotic pastoralist-(0%-6%) and Afroasiatic-related(0%-2%) ancestries. One historical individual buried at the fishing site of Kalala Island in the Kafue River, Zambia, presents significantly more hunter-gatherer-related ancestry (~45%)â€”most likely from Later Stone Age foragers with ancestry similar to that found in ancient individuals from Malawi/Zambia (16,000 - 2,000BP) and to a lesser extent from Central Africansâ€”showing that this foragers ancestry profile, ubiquitous until ~2000 BP, persisted in the region. This supports local oral history that details encounters between Bantu-speaking peoples and resident foragers.

1099 - Sudanic, Central, and Western African ancestry contributions to ancient Ugandan Populations

Esther Brielle

Esther Brielle, Peter Schmidt, John Krigbaum, Jeremy Choin, Kendra Sirak, Shamam Waldman, Daniel Tabin, Jonathan Walz, Jackline Besigye, David Reich

Harvard, University of Florida, University of Florida, Harvard, Harvard, Harvard, Harvard, School for International Training, Uganda Museum and University of Pretoria, Harvard

Our analysis of ancient DNA from Uganda's Ndali Crater Lakes region uncovers a complex admixture history and genetic diversity shaped by millennia of migration and interaction. Examining 15 ancient individuals from 400 to 1600 CE revealed a dynamic interplay of Sudanic African, Central African, West African, and eastern African pastoral ancestries. Early admixture between Sudanic and Central African groups was evident in the Ugandan western Kansyore population from approximately 1500 BP, indicating admixture occurred in the first millennium BCE, predating the individuals examined. Subsequent significant gene flow from West African populations, likely Bantu speakers, in the first half of the first millennium CE, contributed to a genetic makeup of approximately one-third western Kansyore-associated and two-thirds West African-associated ancestry in several ancient Ugandans.

Additionally, modern Ngumba and Tikar South populations from Cameroon, despite dense rainforest barriers, share genetic links with ancient Ugandans, suggesting historical interactions that contributed to the broader genetic mosaic of Central Africa. Using principal component analysis, f-statistics, qpWave, qpAdm modeling, and modified ChromoPainter, we reconstruct the admixture sequence, providing insights into the genetic history of the Western Rift and Lake Victoria to elucidate the relationships between ancient genetic ancestries from across Africa and their legacy in both ancient and present-day populations in Uganda and beyond.

1113 - Adaptation to desert living in diverse pastoralist communities of East Africa.

Julien F Ayroles

Amanda Lea

Vanderbilt

Humans have explored every corner of this planet and have shown an amazing capacity to adapt to extreme conditions. TheÂ TurkanaÂ of Northwest Kenya inhabit one of the hottest and driest regions in sub-Saharan Africa, where water scarcity has been a driving force shaping peopleâ€™s lifestyles and physiologies for thousands of years. To uncover loci associated with arid living and pastoralism, we collected whole genome sequencing, genome-wide gene expression, and biomarker data from several hundredÂ Turkana. We identified 8 outlier regions with strong signatures of positive selection. One of the strongest outlier regions contains a putative enhancer of a kidney-expressed gene involved the response to dehydration, and the elimination of waste produced during the breakdown of purine-rich foods such as red meat, a key component of their diet. In addition to evidence for hard sweeps, we also identified signatures of polygenic selection at loci involved in renal and several metabolic traits. We used our genome-wide gene expression data (n=459) to identify genetic effects on gene expression in PBMCs and to overlap these functional variants with selected loci. Finally, we contrasted these results from the Turkana (Nilotic branch) to the signal derived from sequencing 162 individuals from the Cushitic branch (with whom NilotesÂ share a recent common ancestor). We identified several convergent signatures, highlighting the contribution of standing genetic variation to adaptation. This dataset integrating genomic methods with a wide range of phenotypic and health metrics is part of theÂ TurkanaÂ Genomic and Health Project.

251 - An ancient genetic insight into pre-colonial Trinidad.

Eleni Seferidou

Eleni Seferidou, Zara Ali, Ashleigh Morris, Kathrin Naegele
Max Planck Institute for Evolutionary Anthropology, Leiden University, The National Trust of Trinidad and Tobago, Max Planck Institute for Evolutionary Anthropology

Recent ancient genomic research into the Caribbean region has shed light on the complexity of past Caribbean populations, possible settlement routes, and human movement into and between the islands. Successive dispersals from the mainland to the islands have been recognised, with one of these dispersals originating from northern South America, following a northward route into the islands, possibly via Trinidad, due to its geographic location.
Genetic research thus far has taken a much-needed pan-Caribbean approach to get a grasp of population structure and admixture in the region. Given the fact that diversity in pre-colonial Caribbean populations has become apparent from multiple lines of evidence, this study is aimed at a localised approach, by focusing on a specific island, with the aim of reconstructing individual genomic histories and local interactions. This more localized research framework has led to a collaborative project, together with local archaeologists, researchers, and stakeholders to provide a valuable contribution to the local history through genetic data.
The two sites under investigation - Manzanilla and Red House - are found on opposite sides of the island, with the former located on the eastern shore and the latter near the western shore in the capital, Port of Spain. Thus, they provide a valuable opportunity to study genetic connections between their inhabitants, and among them and neighboring, contemporaneous populations.

627 - Conversations in community-based genomic research: Showcasing a platform from Chile to integrate Indigenous engagement and representation

Constanza de la Fuente Castro

Constanza de la Fuente Castro, TomÃ¡s GonzÃ¡lez Zarzar, Felipe MartÃ-nez, NicolÃ¡s Montalva, Maanasa Raghavan, Constanza Silva-Gallardo, AyelÃ©n Tonko-Huenucoy
Programa de GenÃ©tica Humana, ICBM, Facultad de Medicina. Universidad de Chile, Independent researcher, Escuela de AntropologÃ-a, Facultad de Ciencias Sociales. Pontificia Universidad CatÃ³lica de Chile, Society and Health Research Center, Universidad Mayor, Millennium Nucleus on Sociomedicine, Department of Human Genetics, University of Chicago, Society and Health Research Center, Universidad Mayor, Museo Nacional de Historia Natural

Human genomic research has greatly increased our understanding of biological diversity. Yet, the underrepresentation of certain regions limits its global impact. In Latin America, initiatives to diversify sampling, particularly focused on minoritized and marginalized groups aim to democratize genetic research. However, sampling strategies do not always consider a key problem in research: the level of engagement by communities and their involvement in research design, implementation, and result interpretation. Overcoming this issue requires reconsidering how research is formulated and performed within a region-specific context as well as boosting strategies that level the field and establish long-term relationships between academia and Indigenous researchers, representatives, and communities. With this goal in mind, we designed and implemented two 4-day workshops in 2022 and 2023 in Chile titled â€˜Genomics and Identities in Chileâ€™ focusing primarily on the interplay between genetic histories and self-identities. Both workshops were attended by 30 participants from various Indigenous communities and 12 genetic and anthropological researchers from Chile and abroad. These workshops, first of their kind in Chile, were not designed around an ongoing or new research project, but as a way to facilitate and promote spaces for dialogue on genetic research in the country, especially when involving Indigenous participation. In this presentation, we will discuss these workshops' rationale, content, and format, our experience implementing them, and the limitations and difficulties we faced as organizers. We also discuss participantsâ€™ feedback to motivate similar initiatives for working towards more equitable and

collaborative relationships between academia and Indigenous peoples.

692 - Paleogenomic analyses of archaeological remains reveal sex roles and mobility of ancient families in central Mexico

Daniela Orozco-Perez

Daniela Orozco-Perez, SofÃ-a Ivonne Vieyra Sanchez , Viridiana Villas-Islas, Maria A. Nieves-ColÃ³n, Maria C. Ã•vila-Arcos , Rafael Montiel Duarte, Gabriela Zepeda-GarcÃ-a Moreno , Alberto Aveleyra Talamantes, Karla Sandoval-Mendoza, Andres Moreno-Estrada
National Laboratory of Genomics for Biodiversity (UGA-LANGEBIO), CINVESTAV , National Laboratory of Genomics for Biodiversity (UGA-LANGEBIO), CINVESTAV, International Laboratory for Human Genome Research (LIIGH), UNAM City: QuerÃ©taro , Department of Anthropology, University of Minnesota Twin Cities , International Laboratory for Human Genome Research (LIIGH), UNAM , National Laboratory of Genomics for Biodiversity (UGA-LANGEBIO), CINVESTAV , National Institute of Anthropology and History (INAH), Amigos del Museo de San Miguel A.C./Proyecto cultural Artesanos del Tiempo , Equality and Gender Office of the Centre for Research and Advanced Studies (CODIGO-C), CINVESTAV, National Laboratory of Genomics for Biodiversity (UGA-LANGEBIO), CINVESTAV

Ancient DNA recovery remains a challenge in temperate environments across Latin America. A recently excavated archaeological site in Central Mexico, CaÃ±ada de la Virgen, has been hypothesized to be a multicultural complex inhabited from 540 to 1050 CE., but its origins remain elusive. Using aDNA we inferred the genomic origins, kinship dynamics and biological sex of the individuals buried at the site. We sequenced 19 human skeletal remains, with six samples having >1% endogenous DNA and undergoing enrichment and deep sequencing. Chromosomal sex showed that the most notable burial -â€• The Hierarchâ€•-, previously assigned as male by osteology, was genetically inferred as female and, according to C14, is surprisingly ~1000 years older than the complex. We also uncovered a mother-son relationship involving two skeletons buried next to each other and assessed their genetic origins by combining our data with reference panels of modern Indigenous Mexican populations. Using PCA and F-statistics we found that the mother exhibits genetic proximity to southern Oaxaca and Guerrero Indigenous groups, while her son showed genetic contribution from northern groups traced back to Sonora (near the US border), which is not observed in his mother. Overall, our results suggest that this complex had multiple genetic contributions from places as distant as Mexico's northern and southern tips, converging around 800 CE and bringing burials from before the settlement. Our work illustrates a unique perspective on sex roles in prehispanic societies, challenging the gender narrative of leading characters and mobility dynamics in ancient Mexico.

1019 - Distinct positions of genetic and oral histories: Perspectives from India

Esha Bandyopadhyay

Esha Bandyopadhyay, Arjun Biddanda, Constanza de la Fuente Castro, David Witonsky, JosÃ© Antonio Urban Aragon, Nagarjuna Pasupuleti, Hannah M. Moots, RenÃ©e Fonseca, Suzanne Freilich, Jovan Stanisavic, Tabitha Willis, Anoushka Menon, Mohammed S. Mustak, Chinnappa Dilip Kodira, Anjaparavanda P. Naren, Mithun Sikdar, Niraj Rai, Maanasa Raghavan
Department of Human Genetics, University of Chicago, Department of Human Genetics, University of Chicago, Programa de GenÃ©tica Humana, Instituto de Ciencias BiomÃ©dicas, Facultad de Medicina, Universidad de Chile, Department of Human Genetics, University of Chicago, Department of Human Genetics, University of Chicago, Department of Applied Zoology, Mangalore University, Department of Human Genetics, University of Chicago, Department of Human Genetics, University of Chicago, Department of Human Genetics, University of Chicago, Department of Anthropology, University of Vienna, Department of Human Genetics, University of Chicago, Department of Human Genetics, University of Chicago, Department of Archaeology, University of Cambridge, Department of Applied Zoology, Mangalore University, PureTech Health, Division of Pulmonary Medicine, Cystic Fibrosis Research Center, Department of

Pediatrics, Cincinnati Childrenâ€™s Hospital Medical Center, Anthropological Survey of India, Birbal Sahni Institute of Palaeosciences,Uttar Pradesh, Department of Human Genetics, University of Chicago

Several studies over the past decade have demonstrated the power of human genomic datasets to provide insights into population histories. Concurrently, there is potential for tension arising from genetic histories being prioritized over oral histories or used to confirm community-based knowledge and ethnography, especially if they differ. To investigate the interplay between oral and genetic histories, we focused on the southwestern region of India. We generated whole-genome sequence data from 158 individuals identifying as Bunt, Kodava, Nair, and Kapla. We additionally considered historical records and oral histories, focusing on references to non-local origins such as ancient Scythians for Kodava, Nair, and Bunt, members of Alexander the Greatâ€™s army for the Kodava, and an African-related source of ancestry for Kapla.
Our results showed that the Bunt, Kodava, and Nair exhibit strong genetic similarity to other Indian populations, with Kapla being more similar to tribal groups from South India that harbor a greater proportion of genetic ancestry related to Ancient Ancestral South Indians. We did not find evidence of additional genetic sources in the study populations other than those identified in other present-day South Asians. Our results demonstrate that oral and genetic histories may not always provide a consistent account of population origins. Moreover, our study highlights the importance of community-engaged, multi-disciplinary investigations to motivate avenues for ethnographic and anthropological follow-up of the origin stories in these communitiesâ€™ oral histories, while also cautioning against the dangers of conflating self-identities and genetic histories.

75 - A lethal mitonuclear incompatibility in complex I of natural hybrids

Benjamin M Moran

Benjamin M Moran, Molly Schumer
Stanford University, Stanford University

The evolution of reproductive barriers is the first step in the formation of new species and can help us understand the diversification of life on Earth. These reproductive barriers often take the form of â€œhybrid incompatibilities,â€• where genes derived from two different species have diverged such that they no longer interact properly. Theory predicts that hybrid incompatibilities may be more likely to arise at rapidly evolving genes and in multi-protein complexes, both of which implicate the mitochondria as potential hotspots for incompatibilities. However, there has been sparse empirical data to evaluate these predictions, especially outside experimentally manipulated contexts. Here, we describe a mitonuclear incompatibility involving three genes within respiratory Complex I in naturally hybridizing swordtail fishes. Through a combination of morphometrics, histology, and respirometry, we develop an integrative view of the phenotypes which underly severe selection against this incompatibility in the wild. Ancestry incompatible with mtDNA at either of the two nuclear genes creates a distinct lethal phenotype of either stalled embryonic development or postnatal organ dysfunction, but both genes modify the otherâ€™s effects. Finally, we document a history of accelerated evolution of the genes involved, and provide evidence for introgression of the incompatible genes into a congener. This work thus provides the first glimpse into the genetic architecture, physiological impacts, and evolutionary origin of a multi-gene incompatibility impacting naturally hybridizing species. Our results are consistent with the idea that strong pairwise interactions may overshadow more complex mitonuclear interactions, which will be revealed with increasing experimental power.

133 - The provenance of Proto-North Dravidian ancestry in the Indus valley

Jaison Jeevan Sequeira

Jaison Jeevan Sequeira, George van Driem, Ranajit Das, Mohammed S Mustak

Department of Applied Zoology, Mangalore University, Institut fÃ¼r Sprachwissenschaft, UniversitÃ¤t Bern, Yenepoya Research Centre, Yenepoya (deemed-to-be) University, Department of Applied Zoology, Mangalore University

Research has shown that the present-day population on the Indian subcontinent derives its ancestry from at least three components: Iranian plateau farmer-related, Pontic-Caspian steppe pastoralist-related and Andamanese hunter-gatherer related. However, with more sequences of ancient and modern genomes and fine structure analyses, we can expect more complex ancestries. In this study, we focus on the North-Dravidian linguistic groups to propose a fourth putative source which may have branched out from the basal Middle Eastern component that gave rise to the Iranian plateau farmer-related ancestry. The Elamo-Dravidian theory and the linguistic phylogeny of the Dravidian family tree provide a number of possible chronological fits for the genetic findings presented here but also leave a number of open questions. Our findings show a correlation between the linguistic and genetic lineages in the North Dravidian language speakers when they are modeled together. We suggest that this source, which we call 'Proto-North Dravidian' ancestry, emerged around the time of the Indus Valley Civilisation. This ancestry is distinct from all other sources described so far, and its plausible origin around 4,500 years ago on the Iranian plateau bordering the Indus valley strengthens the idea of a Greater Dravidian heartland before the arrival of Indo-European languages on the Indian subcontinent. Admixture analysis shows that Proto-North Dravidian ancestry is carried by most of the present day non-tribal populations. Our research highlights the importance of population-specific fine structure studies to avoid oversimplification of ancestral reconstruction.

152 - Human sex ratio at birth has virtually no detectable genetic variation: Was Fisher wrong about the evolution of sex ratio?

Siliang Song

Siliang Son, jianzhi Zhang
University of Michigan, University of Michigan

The human sex ratio (fraction of males) at birth is close to 0.5 at the population level, an observation commonly explained by Fisher's principle that, when both sexes require equal parental investment, natural selection favors the production of the less common sex until the sex ratio reaches 0.5. However, past human studies yielded conflicting results regarding the existence of sex ratio-influencing mutationsâ€”a prerequisite to Fisherâ€™s principle, questioning whether the nearly even sex ratio is instead dictated by the random X/Y chromosome segregation in male meiosis. Here we show that, because a personâ€™s offspring sex ratio has an enormous measurement error, a gigantic sample is required to detect sex ratio-influencing genetic variants. Conducting a genome-wide association study that is more powerful than previous studies, we detect a rare allele at rs144724107 upstream of ADAMTS14, a constituent of the ADAMTS proteinase gene family with roles in fertilization, that is associated with a 18% sex ratio reduction. Given the difficulty in precisely measuring offspring sex ratio, it is unsurprising that its estimated heritability is effectively zero. In fact, the estimated heritability would be virtually zero even if offspring sex ratio is as genetically variable as the highly heritable human standing height. These analyses, along with simulations of human sex ratio evolution respectively under purifying and positive selection, demonstrate the compatibility of a nearly zero estimated heritability with Fisherâ€™s principle, suggesting the potential presence of small-effect and/or rare, large-effect genetic variants influencing the offspring sex ratio in humans.

155 - Evolutionary shift of a tipping point forestalls collapse in a microbial community.

Christopher Blake

Christopher Blake, Michael J McDonald
Monash University, Monash University

Climate change is driving global ecosystems rapidly towards tipping points, critical thresholds where small environmental changes can lead to drastic ecological shifts. While progress has been made towards identifying these tipping points, one crucial question remains unanswered: Can evolution modify a tipping point and delay a potential ecosystem collapse? Here, we investigate the power of evolutionary processes to alter tipping points and potentially safeguard communities from collapse.
We evolved a two-species microbial system over 4000 generations to map the changes in ecological stability resulting from coevolution. After that, we adapt our two-species community to an acute environmental stress and reassess the tipping point. Our findings reveal that through adaptation, communities can not only delay their tipping points but also alter their overall stability landscape, expanding the range of environmental conditions under which they can flourish. By developing a mathematical model, we identify how specific traits, such as growth rate, carrying capacity, interspecific interactions, and resistance to environmental change, can enhance ecological resilience.
In this study, we demonstrate that the adaptation of key species within an ecosystem can significantly impact its tipping point, potentially postponing or preventing collapse. These findings open doors for novel strategies like microbiome engineering and directed evolution, which could be employed to bolster the resilience of ecosystems facing the challenges of climate change.

456 - MAYEX is an ancient long non-coding RNA recruited for X chromosome dosage compensation in lizards

Mariela Tenorio

Mariela Tenorio, Diego Cortez, Selene Fernandez-Valverde, Katarzyna Oktaba, Samantha Cruz-Ruiz, Sinai LÃ³pez, Fania Santiago, Joanna Serwatowska, Jose Corona-Gomez, Sergio Encarnacion, Magdalena HernÃ¡ndez , Robert Ossiboff, Fausto MÃ©ndez de la Cruz , Mario Zurita, Cynthia Flores - Aguirre , Diego Arenas-Moreno
CCG, UNAM, CCG, UNAM, BABS, UNSW, LANGEBIO, IPN, CINVESTAV, IPN, IBT, UNAM, CCG, UNAM, CCG, UNAM, CINVESTAV, IPN, LANGEBIO, IPN, CCG, UNAM, CCG, UNAM, CDPM, UF, IB, UNAM, IBT, UNAM, IB, UNAM, IB, UNAM

Long non-coding RNAs (lncRNAs) are essential regulatory elements of sex chromosomes that act to equalize gene expression levels between males and females. XIST, RSX, and roX2 regulate X chromosomes in placentals, marsupials, and Drosophila, respectively. Given that the green anole (Anolis carolinensis) shows complete dosage compensation of its X chromosome, we tested whether a lncRNA was involved. We found an ancient lncRNA, MAYEX, transcribed in multiple tissues of both males and females that gained male-specific expression and sequence conservation more than 89 million years ago, likely after the origin of the XY chromosomes in pleurodonts. Using RNA-seq, ChIP-seq, Hi-C, ChAR-seq, CHART, MS, and CRISPRi data, we found that, specifically in males, MAYEX binds to histone 4, an acetyltransferase, and other transcription activators, to modify the epigenetic landscape of the X chromosome and then significantly increase the expression levels of the X. The MAYEX system operates in all tissues during development and in adults. MAYEX is the first lncRNA in reptiles linked to a dosage compensation mechanism that balances the expression of full sex chromosomes. The similarities between mammals, the fruit fly, and now lizards, indicate that lncRNAs have been concurrently recruited in distant groups during evolution to establish chromosome-wide cis-regulatory mechanisms that control the expression levels of entire sex chromosomes and restore balanced expression ratios between males and females, constituting a remarkable instance of convergent regulatory molecular evolution.

474 - Revisiting the evolution of lactase persistence: insights from South Asian genomes

Meaghan Marohn

Meaghan Marohn, Elise Kerdoncuff, Claire McEvoy, Alka Mohan Chutani, Joyita Prakash, A.B. Dey, Sharon Kardia, Jinkook Lee, Priya Moorjani
University of California, Berkeley, University of California, Berkeley, Queens University Belfast, All India Institute of Medical Sciences, All

India Institute of Medical Sciences, All India Institute of Medical Sciences, University of Michigan, University of Southern California, University of California, Berkeley

Lactase persistence (LP), the ability to digest milk into adulthood, is a striking example of natural selection in humans attributed to mutations upstream of the LCT gene. Within the past 5,000 years, LP-associated alleles have rapidly increased in frequency and are present in nearly 80% of Europeans. Yet there is limited understanding of LP in non-Europeans. India is the world's largest producer of milk, from both cattle and water buffalo, and it is consumed across the subcontinent. Puzzlingly, a previous study found that LP-associated alleles identified in Europeans are on average seen in ~10% of Indians. We assembled a genome-wide dataset of 7,568 present-day South Asians and 1,223 ancient DNA samples (7500 BCE â€“ 1650 CE) from South and Central Asia. For a subset of contemporary Indians, we had access to diet questionnaires showing 30â€“70% of Indians routinely consume milk. Across India, dairy consumption is correlated to LP-associated allele frequency, with highest rates in the north, decreasing south and east. Turning to ancient DNA, we find LP-associated alleles were first seen during the historical period in South Asia. Notably, like in Europeans, LP-associated allele frequency is associated with steppe pastoralist-related gene flow in India that occurred ~1900â€“1500 BCE. This indicates strong selection pressure to retain the LP-associated haplotype, which is at over 65% frequency in some Indian pastoralist groups. Our phenotypic data combined with contemporary and ancient genomic data aid in understanding the origins and evolution of LP outside Europe.

678 - Investigating the ecological suicide (“ecocide”) theory in Rapa Nui with ancient DNA data

Bárbara Sousa da Mota

BÃ¡rbara Sousa da Mota, J. VÃ-ctor Moreno-Mayar, Tom Higham, Signe Klemm, Moana Gorman Edmunds, Francisco Torres Hochstetter, Martin Friess, Jesper Stenderup, Miren Iraeta-Orbegozo, VÃ©ronique Laborde, Evelyne Heyer, Morten E. Allentoft, Hannes Schroeder, Olivier Delaneau, Anna-Sapfo Malaspinas
University of Lausanne, Swiss Institute of Bioinformatics, University of Copenhagen, University of Vienna, University of Copenhagen, Rapanui archaeologist, Rapa Nui Museum (previous affiliation), MusÃ©e de l'Homme, University of Copenhagen, University of Copenhagen, MusÃ©um national dâ€™Histoire naturelle, MusÃ©e de l'Homme, Curtin University, University of Copenhagen, University of Copenhagen, Regeneron Genetics Center, University of Lausanne, Swiss Institute of Bioinformatics

Rapa Nui (Easter Island), in East Polynesia, is one of the most remote places in the world. After the peopling of the island a thousand years ago, it underwent profound environmental changes that have been associated with human presence, including the extinction of endemic birds and, more infamously, the complete disappearance of trees around the 1600s. It has been hypothesized that deforestation triggered famine, warfare and cannibalism, eradicating most of the Rapanui at the time - a so-called ecocide. However, there is a growing body of evidence that does not support that the Rapanui underwent a population collapse in the 1600s. Here, we addressed this question by analyzing 15 newly sequenced ancient Rapanui genomes (0.4x-26x) dated to the early 1800s. Using pseudohaploid and imputed data, we examined whether there was increased relatedness and consanguinity among the individuals by estimating kinship, inbreeding and runs of homozygosity (ROH). We found no close relatives up to 3rd degree and found very low inbreeding levels. Although effective population size was low (maximum ~2000 individuals), we show that the population grew steadily after the peopling of the island. Furthermore, effective population size and ROH estimates for simulated genomes under strong population-collapse scenarios did not match the observed data. In other words, the ancient Rapanui genetic diversity is inconsistent with a population collapse taking place in the 1600s. This is the first contribution from genetics to debunk the most famous example of ecological suicide.

939 - Rapid coevolution preserves the epigenetic establishment of telomere protection

Sung-Ya Lin

Sung-Ya Lin, Hannah Futeran, Mia T. Levine
University of Pennsylvania, University of Pennsylvania, University of Pennsylvania

During sperm maturation, paternal chromosomes undergo extreme compaction. Immediately following fertilization, maternally deposited proteins reverse this paternal chromosome compaction and reestablish canonical protein packaging. This process is conserved from flies to humans, yet many of these maternally deposited proteins evolve adaptively. The causes and consequences of this adaptive evolution remain mysterious. In Drosophila, one such adaptively evolving protein, HipHop, protects chromosome ends from lethal fusions. Maternal HipHop is loaded de novo onto zygotic paternal telomeres, where it recruits other essential subunits of the telomere protection complex. We hypothesized that the de novo establishment of paternal telomere protection requires recurrent innovation of HipHop. To test this, we swapped hiphop from Drosophila yakuba (hiphop[yak]) into D. melanogaster. We found that most embryos of hiphop[yak]/+ females failed to hatch, suggesting that HipHop[yak] is toxic. These embryos arrest at the first mitosis and show delayed paternal chromosome condensation, chromatin bridging, and mitotic catastrophe â€“ phenotypes associated with paternal telomeres that fail to assemble the end-protection complex. Importantly, one other member of this complex both physically interacts with HipHop and evolves adaptively. We hypothesized that HipHop[yak] interferes with paternal telomere recruitment of this D. melanogaster protein, called HOAP. Consistent with this hypothesis, swapping in HOAP[yak], the conspecific interaction partner of HipHop[yak], rescues embryonic lethality. Furthermore, interspecific chimeric HipHop proteins revealed that evolution of HipHopâ€™s HOAP-interaction domain alone is necessary and sufficient for end-protection. Our study demonstrates that essential protein coevolution shapes the epigenetic establishment of vital chromosomal landmarks on zygotic paternal DNA.

1045 - Beyond Continental Groups: Unveiling Dynamic Human Genetic Communities with a Novel Network Analysis Pipeline

María José Palma Martínez

MarÃ-a J Palma MartÃ-nez, Yuridia Posadas GarcÃ-a, Brenda E LÃ³pez Ã•ngeles, Claudia Quiroz LÃ³pez, Anna C F Lewis, Kevin A Bird, Tina Lasisi, Arslan Zaidi, Mashaal Sohail
Centro de Ciencias GenÃ³micas, UNAM, Centro de Ciencias GenÃ³micas, UNAM, Centro de Ciencias GenÃ³micas, UNAM, Escuela Nacional de AntropologÃ-a e Historia, Brigham and Women's Hospital, and Harvard Medical School, Univ. of California, Davis, Univ. of Michigan, Department of Genetics, cell biology, and Development, University of Minnesota, Institute of Health Informatics, University of Minnesota, Centro de Ciencias GenÃ³micas, UNAM

Traditional clustering and visualization approaches in human genetics often operate under statistical and conceptual frameworks that assume inherent, discrete groupings. These methods can inadvertently simplify the multifaceted relationships among individuals, functioning to unveil assumed underlying differences between static groups. Recently there has been a growing embrace of more sophisticated and dynamic methodologies that better reflect the interconnected and fluid nature of human populations. Here we introduce a network-based computational pipeline and visualization tool grounded in relational thinking as a departure from typological models of analysis and interpretation that strive to categorize individuals into a predefined number of sets.

Our pipeline offers the flexibility of using different metrics, such as genetic relationship matrices (GRM), principal components, identical-by-descent (IBD) segments to construct networks. Through the application of community detection (Louvain), we infer communities at multiple resolutions. These aspects facilitate more dynamic data-driven, assumption-free analyses and visualizations of population structure that can identify patterns at different temporal and spatial resolutions suited to the particular research questions.

We applied our pipeline to a dataset merged from the 1000 Genomes and Human Genome Diversity Project. We also infer communities

for trait-specific variation to capture co-occuring patterns relevant for specific traits. Our analysis not only reveals the limitations of traditional groupings but also captures the complexities introduced by recent and distant demographic events and evolutionary processes.

To enable broader engagement with these intricate genetic landscapes, we provide a user-friendly web application (https://sohail-lab.shinyapps.io/GG-NC/) for interactive visualization.

175 - Green Computing in Molecular Evolution and Phylogenetics for the Global South

Beatriz Mello

Beatriz Mello

Federal University of Rio de Janeiro

The computational demands imposed by modern science come with a significant environmental cost. The development and adoption of energy-efficient computing infrastructure, algorithms, and data storage are crucial for sustainable computing in molecular biology and evolution. Whenever computational tools developed for the same purpose provide similar accuracy, choosing greener options should be a commitment. For the Global South, where access to computational resources is frequently a limiting factor, green algorithms have a more profound impact on research development. This aspect of green computing promotes equity and diversity in our field. However, benchmarking efforts to determine the most environmentally efficient approaches are still needed. Studies establishing green methods will lead to more environmentally friendly bioinformatic analyses, besides prompting the Global South to play a bigger role in science. To illustrate this, I will provide a comparative example of the impact of green methods on molecular dating, which has been ranked among the top carbon footprint analyses in phylogenetics. Computationally efficient methods provide divergence time estimates equivalent to those obtained by time-demanding approaches. Among them, the relative rate framework implemented in RelTime was significantly faster and, consequently, more environmentally friendly than the penalized likelihood implemented in treePL. RelTime was orders of magnitude faster than statistically sophisticated Bayesian software. Thus, molecular dating, a mandatory step in many evolutionary studies, demonstrates how the adoption of green solutions not only addresses environmental concerns but also contributes to the advancement of research within scientific communities in low-resource settings.

182 - CMAPLE: efficient phylogenetic inference in the pandemic era

Bui Quang Minh

Bui Quang Minh, Nhan Ly-Trong, Chris Bielow, Nicola De Maio

Australian National University, Australian National University, Freie UniversitÃ¤t Berlin, European Bioinformatics Institute (EMBL-EBI)

We have recently introduced MAPLE (MAximum Parsimonious Likelihood Estimation), a new pandemic-scale phylogenetic inference method exclusively designed for genomic epidemiology. In response to the need for enhancing MAPLEâ€™s performance and scalability, here we present two key components: (1) CMAPLE software, a highly optimized C++ reimplementation of MAPLE with many new features and advancements; and (2) CMAPLE library, a suite of Application Programming Interfaces for easy integration of the CMAPLE algorithm into existing phylogenetic inference packages. Notably, we have successfully integrated CMAPLE into the widely used IQ-TREE2 software, enabling its rapid adoption in the scientific community. These advancements serve as a vital step towards better preparedness for future pandemics, offering researchers powerful tools for large-scale pathogen genomic analysis.

198 - MEGA 12: Advancing Green Computing and Phylogenomics

Sudhir Kumar

Sudhir Kumar, Glen Stecher, Sudip Sharma, Koichiro Tamura
Temple University, Temple University, Temple University, Tokyo Metropolitan University

Molecular evolutionary and phylogenetic analyses play a pivotal role in evolutionary biology and functional genomics studies. However, the computational demands of maximum likelihood methods, frequently used in these studies, become high with increasing lengths of contemporary sequence alignments. The computational burden often exceeds the processing capabilities and memory of personal computers. This limitation is particularly acute in resource-constrained environments, impeding broad participation in evolutionary research. The 12th version of the Molecular Evolutionary Genetics Analysis, MEGA, software suite introduces novel statistical and computational advancements that dramatically enhance computational efficiencyâ€”achieving large reductions in time and memory required for selecting the best-fit substitution model, conducting bootstrap tests of phylogeny, and estimating evolutionary parameters such as branch lengths and divergence times. These advances will make phylogenomic analyses feasible on personal computers. They will also promote wider research participation, accelerate discoveries, and minimize energy consumption, resulting in accessible, sustainable, rigorous, and reproducible phylogenomics. MEGA 12 software suite will be made available for download at www.megasoftware.net.

325 - Shifting Focus from Gene-centric to Evolution-centric Analyses in Cancer Genomics

Li Liu

Hai Chen, Li Liu
Arizona State University, Arizona State University

Introduction: The absence of known driver genes in many tumors hampers precision treatment. We propose that a tumor's somatic evolutionary characteristics, derived from aggregated mutations, can reveal hidden risk factors.

Methods: We developed a novel method, I-SEE-IT to Identify Significant Evolutionary Events In a Tumor. Given whole-exome or whole-genome sequencing data, this method estimates the tumor-specific mutation rate and a set of subclone-specific parameters, including time at emergence, selection coefficient, duration of expansion, cell population fraction, and dN/dS. Using this method, we analyzed 517 multiple myeloma samples.

Results: We discovered that early emergence and prolonged expansion of an advantageous subclone were associated with poor patient survival. We further identified a germline genetic variant in the TBKBP1 promoter region that predisposed a tumor to these evolutionary risk factors. RNA-seq analysis confirmed that this genetic variant was associated with differential expression of the TBKBP1 gene, and high expression of TBKBP1 was linked to poor patient survival. Given that TBKBP1 facilitates tumor-mediated immunosuppression, and treating multiple myeloma cells with Bortezomib induced anti-tumor immune responses, our results collectively imply that the TBKBP1 gene and its related signaling pathways are promising prognostic biomarkers and potential drug targets.

Discussion: I-SEE-IT offers a comprehensive view of tumor evolutionary dynamics, uncovering novel genetic risk factors and potential therapeutic targets overlooked by gene-centric methods.

578 - Causes and impacts of the widespread lack of topological convergence in Bayesian phylodynamic inference on large viral datasets

Jiansi Gao

Jiansi Gao, Guy Baele, Frederick A Matsen
Fred Hutchinson Cancer Center, KU Leuven, Fred Hutchinson Cancer Center, Howard Hughes Medical Institute, Seattle, University of Washington

Bayesian phylodynamic analysis of genomic datasets has been key for elucidating evolutionary and transmission dynamics of pathogens. However, topological convergence of these analyses on large viral datasets has not been comprehensively assessed. By carefully rerunning and analyzing 15 classic large phylodynamic analyses we show that: 1) convergence and mixing issues are widespread in tree topology sampling, which makes phylodynamic inference more computationally challenging; 2) despite apparent failure in topological convergence, for most datasets only a small fraction of clades and nucleotide sites exhibit poor mixing, suggesting that the inefficient exploration of tree space may frequently stem from a small number of viral sequences that possibly have undergone recombination or convergent evolution; 3) conflicting information in the substitution and branching processes may further exacerbate the difficulty in sampling viral time trees, reflected by the negative correlation between phylogenetic and coalescent likelihoods observed in most datasets, and; 4) the inferred treewise molecular and demographic processes appear to be minimally affected by poor exploration of tree space, whereas impacts on the estimated origin time and introduction history of particular clades are more pronounced. We identify biological properties of the viral datasets that may impede tree exploration and suggest new sampling mechanisms targeting these properties to improve the computational performance of Bayesian phylodynamic inference. Outputs from our long analyses (over one trillion generations across datasets) may serve as a comprehensive training dataset for machine-learning-based phylogenetic methods and provide valuable benchmarks for evaluating phylogenetic inference mechanisms, facilitating phylodynamic inference of large viral datasets.

1122 - A spectral step on the divide and conquer path to energy efficient phylogenetic reconstruction

Gavin Huttley

Robert McArthur, Michale Charleston, Ahad Zehmakan, Gavin Huttley
Australian National University, University of Tasmania, Australian National University, Australian National University

The algorithms for phylogenetic reconstruction are central to computational molecular evolution. The relentless pace of data acquisition has exposed their poor scalability and the conclusion that the conventional application of these methods is impractical and not justifiable from an energy usage perspective. This is true even for state-of-the-art tools such as IQ-TREE 2. Furthermore, the drive to improve the statistical performance of phylogenetic methods produces increasingly parameter-rich models of sequence evolution, which worsens the computational performance.

Is the statistical robustness versus computational performance tradeoff inescapable? Decades-old results, both theoretical and algorithmic, suggest we may be able to divide and conquer an escape. The combination of the Disk Coverage Method (DCM, Huson et al. 1999) and Min-Cut supertree (MCS, Semple and Steel 2000) provides a formal basis for expecting that large phylogenetic problems can be split up (DCM) and the results merged (MCS) to obtain a final result. MCS explicitly accommodates rooted topologies, which can arise from the more biologically plausible non-stationary models of sequence evolution.

Here we present Spectral Cluster Supertree (SCS), a novel supertree method for merging overlapping rooted phylogenetic trees. It offers significant improvements over previous methods in terms of time complexity and accuracy, particularly for large problems. While Bad Clade Deletion (BCD, Fleischauer and BÃ¶cker 2017) generates more correct clades, SCS's generated tree is generally topologically closer to the true tree. For datasets containing 10000 taxa and ~500 source trees, BCD takes ~2 hours, while SCS takes ~15 seconds.

187 - Modeling recent positive selection using identity-by-descent segments

Seth D Temple

Sharon R Browning, Ryan K Waples

University of Washington, University of Washington

Recent positive selection can result in an excess of long identity-by-descent (IBD) haplotype segments. The statistical methods that we propose here address three major objectives in studying selective sweeps: scanning for regions of interest, identifying possible sweeping alleles, and estimating a selection coefficient s. First, we implement a selection scan to locate regions of excess IBD rate. Second, we develop a statistic to rank alleles that are in strong linkage disequilibrium with a putative sweeping allele. We aggregate these scores to estimate the allele frequency of the sweeping allele, even if it is not genotyped. Third, we propose an estimator for the selection coefficient and quantify uncertainty using the parametric bootstrap. Comparing against state-of-the-art methods in extensive simulations, we show that our methods are better at identifying sweeping alleles that are at low frequency and at estimating s when s >= 0.015. We apply these methods to study positive selection in European ancestry samples from the TOPMed project. We analyze eight loci where the IBD rate is more than four standard deviations above the population median. The IBD rate at LCT is thirty-five standard deviations above the population median, and our estimates of its selection coefficient imply strong selection within the past two hundred generations. Overall, we present robust and accurate approaches to study very recent adaptive evolution without knowing the identity of the causal allele or using time series data.

550 - Reconstructing parent genotypes at genotyping array accuracy using siblings and other relatives

Amy L. Williams

Ying Qiao, Ethan M Jewett, Kimberly F McManus, William A Freyman, John Blangero, The 23andMe Research Team, Amy L Williams

Cornell University, 23andMe, Inc., 23andMe, Inc., 23andMe, Inc., The University of Texas Rio Grande Valley, 23andMe, Inc., 23andMe, Inc.

Children inherit two chromosome copies, one from each parent, with both formed via recombination. In expectation, n siblings inherit a proportion of 1-1/2n of both parents' genomes, providing rich information about their parentsâ€™ DNA. Reconstructing ungenotyped parents can increase power in genome-wide association studies, reveal parentsâ€™ ancestral origins, and improve relatedness inference.

We developed a method to reconstruct the genotypes of parents from full siblings using a combination of family-based phasing, sex-specific genetic maps to infer parent sexes, and identity by descent (IBD) sharing to close relatives. Specifically, HAPI2 jointly phases siblings, building a one-generation ancestral recombination graph (ARG), and implicitly reconstructs the parentsâ€™ DNA. Given data for â‰¥7 siblings, HAPI2 often provides chromosome-scale haplotypes for the parents. For fewer siblings, we combine phase information with IBD between the reconstructed parent haplotypes and genotyped relatives. This disambiguates parental assignment of the DNA within and between chromosomes since these IBD segments are typically inherited through only one parent.

We tested this method on research participant families in the 23andMe, Inc. database and the San Antonio Mexican American Family Studies (SAMAFS). With â‰¥7 children, we reconstructed 73.7% (mean; 96.6% maximum) of the two parents' genotypes with mean error rate of <10-4. Using 4 children and leveraging IBD sharing with close relatives, we reconstructed 62.7% (mean; 82.9% maximum), with mean error rate <10-4. As datasets grow in size, more families will be implicitly collected; our work holds promise to enable high quality reconstruction of parent genotypes.

700 - How to paint a dog: limits of local ancestry inference in ancient genomes

Katia Bougiouri

Katia Bougiouri, Evan K Irving-Pease, Laurent AF Frantz, Fernando Racimo, Martin Petr
University of Copenhagen, Globe Institute, University of Copenhagen, Globe Institute, School of Biological and Behavioural Sciences, Queen Mary University of London, Palaeogenomics Group, Department of Veterinary Sciences, Ludwig Maximilian University, University of Copenhagen, Globe Institute, University of Copenhagen, Globe Institute

Recent advances in imputation of ancient genomes have created the opportunity to apply state-of-the-art statistical genetics methods to ancient genomes of different species to study various aspects of evolutionary history across large timescales. One popular class of such methods performs what is known as local ancestry inference (LAI), which allows to â€œpaintâ€• an individualâ€™s genome as a mosaic of segments inherited from different ancestral source populations. Despite their power to reveal patterns of genetic ancestry across the genome, most popular LAI methods have been designed to study relatively recent admixture events in human history using large panels of present-day genomes. It is currently unknown how reliable is LAI in ancient genomes, which are characterized by scarce geographic and temporal sampling, potentially large divergence between ancient source proxies and admixing populations, and complex ancient demographic histories. In this study, we evaluated the performance of two popular LAI methods, FLARE and MOSAIC, under different demographic scenarios, various sampling schemes and sample sizes, as well as the timing of admixture events, and assess their accuracy as a function of these parameters. Furthermore, we tested to what extent the inclusion of source proxies pre- or postdating the admixture event or grouping putative sources from different time points, affect the inferences. Our study establishes the limits of LAI in ancient genomes and provides general guidelines for LAI in future studies, which we demonstrate by a proof-of-concept LAI in a set of imputed Mesolithic dog genomes from northwestern Siberia.

720 - Archaic ancestry fragment calling using wavelet decomposition

Eduardo Martín González Orozco

Eduardo MartÃ-n GonzÃ¡lez-Orozco, Leonardo Iasi, Divyaratan Popli, Benjamin Peter
Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology

Detecting introgressed fragments from Neandertals and Denisovans in modern human genomes is an important step in the study of human evolution. A common strategy, implemented e.g. in Admixfrog and HMMix, is to use a Hidden Markov model that can detect archaic ancestry in small windows along the genome (bins). These methods estimate a posterior probability of putative archaic ancestry for small bins along a genome. Nearby bins with high probability can then be merged together into introgressed segments, the specifics and parameters of these procedures can introduce biases that may influence further downstream analyses such as admixture dating. Here, we develop a new method to call archaic ancestry fragments from posterior probabilities using a wavelet decomposition., This method allows us to perform a multiresolution fragment calling using the posterior probabilities of each bin, meaning that we can identify different fragment sizes across the genome by merging bins. We tested this method on simulation of Neandertals introgression into Denisovans. We then compared the results to the penalty score method performed by Admixfrog for fragment calling. We were able to identify segments with higher sensitivity and specificity without a fragment size cutoff in comparison to the Admixfrog penalty score.This method presents a promising alternative for accurately identifying the size of fragments of archaic ancestry in modern humans.

258 - The impact of spontaneous mutation accumulation in a mammalian species

Peter Keightley

Peter Keightley, Jobran Chebib, Eugenio Lopez-Cortegano, Anika Jonas, Sven Kuenzel, Diethard Tautz
University of Edinburgh, University of Edinburgh, University of Edinburgh, Max Planck Institute for Evolutionary Biology, Max Planck Institute for Evolutionary Biology, Max Planck Institute for Evolutionary Biology

In 1950, H. J. Muller argued that the accumulation of newly arising spontaneous mutations could reduce fitness of human populations over a measurable timescale. This concern has been echoed more recently, since improvements in living conditions and health care have undoubtedly reduced the strength of natural selection in many human societies. However the impacts of new mutations on fitness of human populations is difficult (or perhaps impossible) to measure directly. In this study, we combine molecular and phenotypic assays of spontaneous mutation accumulation in the house mouse to predict the impact of mutation accumulation in humans. We employed the mutation accumulation (MA) experimental design, where replicated full-sib mated lines are bred for many generations in conditions where natural selection is largely removed. We sequenced individuals from the MA lines using both Illumina and PacBio technologies, and present results on rates of single nucleotide and small and large insertion mutations, mutation spectra, and the new variation arising from transposable element movements. We integrate these molecular estimates of the mutation rate with estimates of the new heritability for quantitative traits and the rate of erosion of fitness per generation from spontaneous mutation accumulation. We predict the rate of fitness loss from spontaneous mutation accumulation in humans, and determine whether this should be of concern in the foreseeable future.

606 - Why do we get sick? Unraveling Genetic Trade-offs between Fertility, Longevity, and Complex Diseases

Eva Brigos Barril

Eva Brigos, Claudia Vasallo, Elena Bosch FustÃ©, Hafid Laayouni, NatÃ lia Vilor-Tejedor, Carlos Morcillo-SuÃ¡rez, Gerard MuntanÃ©, Arcadi Navarro
Department of Medicine and Life Sciences, Institute Of Evolutionary Biology (UPF-CSIC), Centre for Genomic Regulation, The Barcelona Institute of Science and Technology, Department of Medicine and Life Sciences, Institute Of Evolutionary Biology (UPF-CSIC), Centro de InvestigaciÃ³n BiomÃ©dica en Red de Salud Mental (CIBERSAM), Department of Medicine and Life Sciences, Institute Of Evolutionary Biology (UPF-CSIC), Bioinformatics Studies, ESCI-UPF, Barcelonabeta Brain Research Center, Pasqual Maragall Foundation, Radboud University, Centre for Genomic Regulation (CRG), Department of Medicine and Life Sciences, Institute Of Evolutionary Biology (UPF-CSIC) , Department of Medicine and Life Sciences, Institute Of Evolutionary Biology (UPF-CSIC), Hospital Universitari Institut Pere Mata, IISPV, Universitat Rovira i Virgili, Biomedical Network Research Centre on Mental Health (CIBERSAM), Department of Medicine and Life Sciences, Institute Of Evolutionary Biology (UPF-CSIC), Centre for Genomic Regulation, The Barcelona Institute of Science and Technology, Barcelonabeta Brain Research Center, Pasqual Maragall Foundation, InstituciÃ³ Catalana de Recerca i Estudis AvanÃ§ats (ICREA)

Modern humans exhibit lower fertility rates, extended longevity, and a higher incidence of complex diseases relative to other mammals. Life history theory suggests that varying levels of investment in reproductive function can impact the likelihood of developing diseases and longevity, suggesting intricate interactions between life-history traits and disease. However, despite a few identified mutations acting as genetic switches, we lack unambiguous and systematic genomic evidence to understand these trade-offs. We use genomic data to evaluate the shared pleiotropic landscape between fertility, longevity, and 61 complex diseases across 12 disease domains. We report a negative genetic correlation and pleiotropy between fertility and longevity. This trade-off is also evident when evaluating their shared genetic architecture with disease. We show that positive genetic correlations are common between fertility and disease. The demographic analysis using the UK Biobank supports the latter genetic observations, where individuals diagnosed with a disease tend to

have more offspring than controls. In contrast, there is a predominant proportion of negative genetic correlations and pleiotropies with longevity. Specifically, we identified 297 pleiotropic SNPs shared between diseases and fertility and 569 with longevity. The analysis of the signature of selection shows that risk-allele frequencies are higher for pleiotropic variants that are linked to increased fertility. Also, these have undergone recent positive selection despite increasing disease risk. Overall, we highlight fertilityâ€™s role in shaping current disease patterns and disentangle the complex network of interactions between fertility, longevity, and disease.

650 - De novo mutations, gene conversions and recombination in humans, chimpanzees and gorillas inferred from direct Pacbio HIFI sequencing of sperm samples and testis tissue

Mikkel Heide Schierup

Peter S Porsborg, Anders P Charmouh, Vinod Singh, Sofia Boegh Winge, Meritxell Riera, Christina Hvilsom, Marta Pelizzola, Asger Hobolth, SÃ¸ren Besenbacher, Kristian Almstrup, Thomas Bataillon, Mikkel H Schierup
Aarhus University, Aarhus University, Aarhus University Hospital, Copenhagen University Hospital, Aarhus University, Copenhagen Zoo, Aarhus University, Aarhus University, Aarhus University Hospital, Copenhagen University Hospital, Aarhus University, Aarhus University

De novo mutations are typically identified by analysis of deep sequencing of father, mother, and offspring trios. Each child is born with 40-100 de novo SNVs in great apes, including humans. About 75% of these are transmitted from the father, increasing in number with age. Thus, mutation accumulation during spermatogenesis is the major source of de novo mutations. We performed highly accurate and 30-70X Pacbio HIFI sequencing of testis biopsies from humans, chimps and gorillas, and human sperm samples. Each sample is then de novo assembled into haplotypes and remapping the reads we can accurately call de novo mutations, gene conversion events and recombination events observed in just a single read. This allows many thousands of de novo mutations, hundreds of gene conversions and hundreds of recombination events to be called from each individual sample. We recover age dependence and species differences in mutation accumulation and study the differences between testis and sperm samples and differences in single-strand damage and true de novo mutations. We are also able to accurately estimate gene conversion tract lengths to be similar and very short (mean length less than 100 bp) in each species and find that the gene conversion events are distributed more uniformly along the genome than recombination events.

826 - Discovering the cis-regulatory basis of archaic human-derived phenotypes

Ryder Easterlin

Ryder Easterlin, Simon Fishilevich, Katharina Lange, Nadav Ahituv, David Gokhman
UC San Francisco, Weizmann Institute of Science, Weizmann Institute of Science, Weizmann Institute of Science, UC San Francisco, Weizmann Institute of Science

Our understanding dictates that the phenotypic differences between modern humans and our closest evolutionary relatives, Neanderthals and Denisovans, is largely attributable to genetic variation in noncoding regions of the genome. These regions, which contain enhancers and promoters, play crucial roles in gene regulation. Despite this, the precise impact of lineage-specific changes in noncoding regions on human evolution remains poorly understood. In this study, we utilized massively parallel reporter assays to investigate the influence of genetic variants in putative cis-regulatory elements on the phenotypic divergence between modern humans and archaic humans (Neanderthals and Denisovans). Leveraging high-coverage archaic human genomes from three Neanderthals and one Denisovan, along with catalogs of modern human genetic variation, we identified 57,403 fixed single-nucleotide variants that originated in the archaic lineage. For each variant, we synthesized two sequences: one with the archaic-derived allele and another with the ancestral allele. By testing the effects of each sequence version on gene expression in key cell types relevant to human divergenceâ€”osteoblasts, skin

fibroblasts, adipocytes, and excitatory neuronsâ€”we identified thousands of active regulatory sequences containing archaic-derived variants. Furthermore, we found that hundreds of these variants drive differential gene expression. We pinpointed the genes likely affected by these cis-regulatory changes, shedding light on the organs, biological processes, and phenotypes impacted by these regulatory alterations. Overall, our study offers a comprehensive catalog of the gene regulatory effects of archaic-derived variants in critical cell types, providing valuable and functionally supported insights into recent human evolution.

911 - Machine learning reveals the diversity of human 3D chromatin contact patterns

Erin Nicole Gilbertson

Erin N Gilbertson, Colin M Brand, Evonne McArthur, David C Rinker, Shuzhen Kuang, Katherine S Pollard, John A Capra
University of California, San Francisco, UCSF, University of Washington, Vanderbilt University, Gladstone Institutes, Gladstone Institutes, University of California, San Francisco, University of California, San Francisco

Understanding variation in chromatin contact patterns across human populations is critical for interpreting non-coding variants and their effects on gene expression and phenotypic variation. However, experimental determination of chromatin contacts at a population-scale is prohibitively expensive.

To overcome this challenge, we developed a machine learning method to quantify the diversity 3D chromatin contacts at 2 kilobase resolution from genome sequence alone. We validated that it accurately generalized on contact data from diverse individuals. We then applied this approach to thousands of diverse modern humans and the inferred human-archaic hominin ancestral genome. Patterns of 3D contact divergence genome-wide are qualitatively similar to patterns of sequence divergence, but we find that 3D divergence in local 1-megabase genomic windows does not follow sequence divergence. In particular, we identify 392 windows with significantly greater 3D divergence than expected from sequence. Moreover, 26% of genomic windows have rare 3D contact variation observed in a small number of individuals. Using in silico mutagenesis, we find that most sequence changes to do not result in changes to 3D chromatin contacts. However, in windows with substantial 3D divergence, just one variant can lead to divergent 3D chromatin contacts without high overall sequence divergence.

In summary, inferring 3D chromatin contact maps across human populations reveals diverse contact patterns. We anticipate that these genetically diverse maps of 3D chromatin contact will provide a reference for future work on the function and evolution of 3D chromatin contact variation across human populations.

925 - Archaic introgression in modern humans modelled by ancestral components

Moisès Coll Macià

MoisÃ¨s Coll MaciÃ , Mikkel Heide Schierup
Bioinformatics Research Centre (BiRC), Aarhus University, Bioinformatics Research Centre (BiRC), Aarhus University

The geographical distribution of archaic DNA in modern humans provides crucial insights into our history of admixture with Neanderthals and Denisovans. However, grouping individuals by ethnicity or location does not always reflect a shared evolutionary history. Take South Americans, their genomes are the result of recent admixtures between divergent Native American, European and African lineages and so to understand their diversity it is crucial to account for the individual-level ancestry proportions. This applies to archaic components as well, since admixture with archaics occurred with the ancestral populations of extant groups. This project proposes a novel approach that models the archaic content of modern individuals with all their ancestral components jointly in a Bayesian framework. Here we analyse

individuals from 1000GP and the HGDP (both called by gnomAD) by first obtaining the ancestral components and the archaic fragments with ADMIXTURE and hmmix, respectively. We then incorporate ancestral fractions as predictors in a multiple linear regression to model the Neanderthal and Denisova components of each individual. One intriguing finding that we detect is that the ancestry component which maximizes in South Asians is estimated to harbour similar Neanderthal total sequence than other ancesries, but distributed in much shorter fragments. This potentially hints at shorter generation intervals in their ancestral population over the past 50,000 years than any other Eurasian group. This work demonstrates that describing the archaic legacy in the light of ancestral components reveals a deeper understanding of past admixture events and the demographic history of our ancestral populations.

355 - Maximum-likelihood inferences from ancient environmental DNA using panmitogenomes

Gabriel Renaud

Gabriel Renaud, Nicola Alexandra Vogel, Joshua Daniel Rubin

Technical University of Denmark, Technical University of Denmark, Technical University of Denmark

Due to its small size and high abundance relative to nuclear DNA, mitochondrial DNA continues to play a central role in ancient DNA and ancient environmental DNA studies. Some special challenges plaguing such studies include high divergence to all known modern references and high rates of ancient DNA damage. Especially for ancient environmental DNA, the scarcity of data becomes problematic when one seeks to make inferences based on such scant data. We present two different applications of panmitogenomics, i.e. pangenomics applied to mitochondrial DNA. These applications target ancient DNA from a single source and ancient environmental DNA. The first involves the robust inference of human mitochondrial haplogroups from ancient DNA. We show that our method outperforms existing tools by offering greater resilience against low coverage or incomplete consensus sequences via the mitigation of reference-related biases. Additionally, it generates Bayesian confidence scores informed by phylogeny that are impartial to any specific population. The second application concerns the sensitive detection of animal DNA traces in ancient environmental DNA samples. The method can performs taxonomic assignment as well as quantify the respective abundance of detected taxa. In addition, we developed a method that can use our previous taxa assessment and perform their respective species identification. Both of these applications show that panmitogenomics is a groundbreaking novel framework to improve phylogenetic inferences for ancient DNA as well as ancient and modern environmental DNA.

817 - Pangenome assemblies reveal the evolution and recent in vivo activity of human LINE-1 retrotransposons

Rick McLaughlin

Lei Yang, Amanda Norseen, Rick McLaughlin

Pacific Northwest Research Institute, Pacific Northwest Research Institute, Pacific Northwest Research Institute, University of Washington

LINE-1 retrotransposons dominate the human genome. As the most abundant and only active autonomous transposable elements in extant human genomes, LINE-1s affect human physiology and disease risk. Because LINE-1s are long and highly repetitive, our understanding of LINE-1 variation and recent evolution within diverse humans has been limited by the scarcity of haploid-resolved and long read-based assemblies. Here, we provide the first analysis of LINE-1 insertions, sequence variation, and recent activity using diverse, long read-resolved human genomes from the Human Pangenome Reference Consortium (HPRC). We found 13,045 LINE-1s with both intact open reading frames required for retrotransposition located at 671 unique insertion sites, and we infer ~3,000 such insertion sites in the human population. We constructed networks of these sequences which show evolutionary transitions from ancient, active groups of LINE-1s to many younger, currently active LINE-1s in modern humans. Using sequence-distances computed between these 671 LINE-1s

and the 3â€™ transduction between LINE-1 insertion sites, we quantified the recent in vivo activity of each LINE-1, revealing the specific LINE-1s that have replicated at high rates in humans. Finally, our evolutionary analysis shows the emergence of new adaptive changes which we test in vitro to identify putative drivers of LINE-1 evolution in humans. These data constitute a novel analysis of the diversity and activity of the youngest human LINE-1s, made possible by an increasing number of diverse, long read-based human genome assemblies.

988 - Global diversity, recurrent evolution, and recent selection on amylase structural haplotypes in humans

Runyang Nicolas Lou

Davide Bolognini, Alma Halgren, Runyang Nicolas Lou, Alessandro Raveane, Joana L. Rocha, Andrea Guarracino, Nicole Soranzo, Jason Chin, Erik Garrison, Peter H. Sudmant

Human Technopole, Department of Integrative Biology, University of California Berkeley, Department of Integrative Biology, University of California Berkeley, Human Technopole, Department of Integrative Biology, University of California Berkeley, Human Technopole, Department of Genetics, Genomics, and Informatics, University of Tennessee Health Science Center, Human Technopole, Foundation for Biological Data Science, Foundation for Biological Data Science, Department of Genetics, Genomics, and Informatics, University of Tennessee Health Science Center, Department of Integrative Biology, University of California Berkeley, Center for Computational Biology, University of California Berkeley

The adoption of agriculture, first documented ~12,000 years ago in the Fertile Crescent, triggered a rapid shift toward starch-rich diets in human populations. Amylase genes facilitate starch digestion and increased salivary amylase copy number has been observed in some modern human populations with high starch intake, though evidence of recent selection is lacking. Here, using 52 long-read diploid assemblies and short read data from ~5,600 contemporary and ancient humans, we resolve the diversity, evolutionary history, and selective impact of structural variation at the amylase locus. We find that both salivary and pancreatic amylase genes have higher copy numbers in populations with agricultural subsistence compared to fishing, hunting, and pastoral groups. We identify 28 distinct amylase structural architectures and demonstrate that identical structures have arisen independently multiple times throughout recent human history. Using a pangenome graph-based approach to infer structural haplotypes across thousands of humans, we identify extensively duplicated haplotypes present at higher frequencies in modern day populations with traditionally agricultural diets. Leveraging 534 ancient human genomes we find that duplication-containing haplotypes (with 5 or more AMY copies in contrast to the ancestral 3 copies) have increased in frequency more than seven-fold over the last 12,000 years providing evidence for recent selection in Europeans at this locus comparable in magnitude to that at lactase. Together, our study highlights the strong impact of the agricultural revolution on human genomes and the importance of long-read sequencing in identifying signatures of selection at structurally complex loci.

1038 - A Pan-pangenome captures the full spectrum of genetic variation in humans, chimpanzees and bonobos

Joana L. Rocha

Joana L Rocha, Runyang Nicolas Lou, Juan M Vasquez, Alison Killilea, Stacy Li, Andrea Guarracino, Kendra Hoekzema, Evan E Eichler, Erik Garison, Matt W Mitchell, Peter H Sudmant

UC Berkeley, Department of Integrative Biology, University of California Berkeley, UC Berkeley, Department of Integrative Biology, UC Berkeley, Department of Integrative Biology, UC Berkeley, Department of Molecular and Cell Biology, UC Berkeley, Department of Integrative Biology, Department of Genetics, Genomics, and Informatics, University of Tennessee Health Science Center, Memphis, USA, Human Technopole, University of Washington School of Medicine, Department of Genome Sciences, University of Washington School of

Medicine, Department of Genome Sciences, University of Washington, HowardHughes Medical Institute, Department of Genetics, Genomics, and Informatics, University of Tennessee Health Science Center, , Coriell Institute for Medical Research, UC Berkeley, Department of Integrative Biology

Chimpanzees and bonobos (genus Pan) are our species closest living relatives. The genetic background of these species provides evolutionary context for capturing and representing the extensive structural diversity of human haplotypes. However, while long-read sequencing consortium projects such as T2T and HPRC are providing an unprecedented opportunity to decipher some of the most intractable and complex regions of the human genome, we have yet to fully understand the extent of genome structural diversity in the Pan genus. Here, we have assembled 60 haplotype-resolved genomes from a comprehensive sampling of a diverse set of chimpanzees and bonobos throughout their geographic range in Africa, to create a graph-based, telomere-to-telomere representation of Pan genomic diversity. This Pan-pangenome effort comprises some of the most contiguous assemblies to-date for non-human primates at a population scale. Our work complements and extends the recent human focused efforts to obtain a near-complete representation of the full spectrum of global genetic variation. Using these resources, we are characterizing the structure, composition, function, and evolutionary trajectories of genomic variation that is shared between humans, chimpanzees, and bonobos. We found that chimpanzees and bonobos have a higher length distribution of SVA retrotransposons and show extensive structural variation in regions associated with resistance to malaria resistance and sleeping sickness in humans. This will allow us to interrogate and test hypotheses on the adaptive significance of complex genomic structures that give way to shared genetic diversity across the primate lineage.

130 - A paleogenomic study of ancient island inhabitants of the Chonos archipelago, in northern Patagonia, Chile

Roberta Davidson

Roberta Davidson, Xavier Roca Rada, EpifanÃ-a Arango Isaza, Omar Reyes, Chiara Barbieri, Bastien Llamas, Lars Fehren-Schmitz Australian Centre for Ancent DNA and The Envronment Institute, School of Biological Sciences, The University of Adelaide, Australian Centre for Ancent DNA and The Envronment Institute, School of Biological Sciences, The University of Adelaide, Universidade de Coimbra, Department of Evolutionary Biology and Environmental Studies, University of Zurich, Center for the Interdisciplinary Study of Language Evolution, University of Zurich, Universidad de Magallanes, Department of Life and Environmental Sciences, University of Cagliari, Australian Centre for Ancent DNA and The Envronment Institute, School of Biological Sciences, The University of Adelaide, UCSC Paleogenomics, Department of Anthropology, University of California

The Chonos archipelago is a region of over 150 fjordic islands in northern Patagonia, Chile. Once inhabited by a collection of seafaring ethnic groups broadly known as the Chono people, the impacts of forced transportation and slavery resulted in total loss of Indigenous culture by the end of the 18th century. Limited ethnohistorical records describe the Chonos and â€˜canoe peopleâ€™ with a marine subsistence. Recent studies of carbon and nitrogen isotopes in pre-European contact human remains have revealed signatures consistent with a marine diet, and archaeological findings of shell middens corroborate a marine lifestyle. However, the human population history of this region of northern Patagonia has never been studied with ancient DNA. In this project, twenty ancient genomes, from eight locations, spanning from 300â€“2000 years old, are analysed to gain insights into the population history of the ancient peoples. Preliminary results reveal a population that is related to, yet genetically distinct from other ancient Patagonian populations. Signals of homogeneity suggest inter-island interactions with some degree of isolation from the mainland and southern Patagonia.

188 - Colombian Arawak genomes shed light on population demography of South America

Marisol Naydu Espitia Fajardo

Marisol N Espitia Fajardo, Daniela Orozco Perez, Juan P PeÃ±a Sanchez, Angela M Zamudio Jacome, Yuberlis Gonzalez Ipuana, MarÃ-a D Ortega, Guillermo Barreto, Andres Moreno Estrada
Universidad del Valle, Centro de investigaciÃ³n y de estudios avanzados del IPN (Cinvestav) , Universidad del Valle, Universidad del Valle, Universidad del Valle, Centro de investigaciÃ³n y de estudios avanzados del IPN (Cinvestav) , Universidad del Valle, Centro de investigaciÃ³n y de estudios avanzados del IPN (Cinvestav)

Population genetic studies in South American indigenous groups are essential to better understand their complex demographic history. However, genomic data available for these populations remain scarce, which makes it difficult to reconstruct their population dynamics, dispersal patterns, ancestral origins and migratory routes within the continent. As the main gate to South America, presentday Colombia plays a key role in understanding such evolutionary dynamics. To fill this gap, we genotyped 120 Colombian indigenous individuals belonging to the Arawak linguistic family, the largest in South America, together with individuals of the ChibchaPaezen family. Genomewide data for 1.8 million SNPs were generated using MEGA arrays and combined with previously published data from Native American individuals. We performed analyses of global ancestry and population structure, including PCA, ADMIXTURE, as well as IBD sharing and F statistics to infer the genetic history of these populations. Arawak populations showed genetic affinities with both Andean and other Amazonian populations, which could be explained by the extended range of ecoregions inhabited by Arawak speaking groups across South America. This suggests that the Arawak are not strictly substructured into genetic clusters from the Andes and the Amazon, but rather show patterns of shared ancestry throughout their expansion range. This also implies that the Arawak language did not spread exclusively through cultural transmission, as we found genetic connections between Arawak speaking communities over long geographic distances across South America. Our study contributes to the population demography reconstruction of overlooked indigenous cultures that have shaped present-day genetIc diversity of SouthAmerica.

527 - The contribution of Neanderthal introgression to the heritability of human complex traits in Mexico

Valeria Anorve-Garibay

Valeria Anorve-Garibay, Sofia Vieyra-Sanchez, Lourdes Garcia-Garcia, Andres Moreno-Estrada, Diego Ortega-Del Vechyo, Mashaal Sohail, Emilia Huerta-Sanchez
Laboratorio Internacional de Investigacion sobre el Genoma Humano, Universidad Nacional Autonoma de Mexico, Center for Computational Molecular Biology, Brown University, National Laboratory of Genomics for Biodiversity (LANGEBIO), CINVESTAV, Instituto Nacional de Salud Publica, National Laboratory of Genomics for Biodiversity (LANGEBIO), CINVESTAV, Laboratorio Internacional de Investigacion sobre el Genoma Humano, Universidad Nacional Autonoma de Mexico, Centro de Ciencias Genomicas, Universidad Nacional Autonoma de Mexico, Department of Ecology and Evolutionary Biology, Brown University, Center for Computational Molecular Biology, Brown University

The analysis of high-coverage sequences of two archaic humans: Neanderthal and Denisovan, have revealed that they interbred with anatomically modern humans (AMH). Recent studies in European and East Asian cohorts have shown that Neanderthal and Denisovan introgression facilitated adaptation and contributed to complex traits associated with pigmentation, the resistance to pathogens and drug metabolism. However, the impact of archaic variation on complex traits remains under-explored in admixed populations from the Americas. Here, we leverage genome-wide genotype data from the Mexican Biobank Project (MXB) to investigate the genomic landscape of archaic variation in Mexico. We inferred a map of 113,542 Neanderthal introgressed variants (NiVs) and 8,374 Denisovan introgressed variants (DiVs) identified in Mexican individuals. We conducted and validated a rigorous pipeline to estimate the contribution of Neanderthal introgressed variants to the heritability of 31 distinct phenotypes measured in 5,753 individuals in MXB. We find that

Neanderthal introgressed variants do not significantly contribute to trait variation (on average, NiVs explain ~ 0.23 % of trait variation, pval > 0.05). Consistent with previous studies (Wei et al., 2023), we find that Neanderthal introgressed variants are depleted of heritability for all traits considered. Our study provides novel evidence on the impact of Neanderthal introgression on complex trait variation in Mexicans and represents a valuable resource for future studies characterizing archaic introgression in underrepresented populations in genomic studies.

647 - Picuris Pueblo oral history and genomics reveal genetic continuity in North American Southwest

Thomaz Pinotti

Thomaz Pinotti, Michael A Adler, Richard Mermejo, Julie Bitz-Thoren, Hugh McColl, Gabriele Scorrano, Motahare Feizabadi Farahani, Devlin Gandy, Charleen Gaunitz, Abigail RamsÃ¸e, Lasse Vinner, Fabrice Demeter, FabrÃ-cio R Santos, Martin Sikora, David J Meltzer, J. VÃ-ctor Moreno-Mayar, Craig Quanchello, Picuris Pueblo Tribal Council, Eske Willerslev

Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, LaboratÃ³rio de Biodiversidade e EvoluÃ§Ã£o Molecular (LBEM), Universidade Federal de Minas Gerais, Southern Methodist University, Picuris Pueblo, Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, Department of Archaeology, University of Cambridge, Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, MusÃ©e de l'Homme, UniversitÃ© de Paris, LaboratÃ³rio de Biodiversidade e EvoluÃ§Ã£o Molecular (LBEM), Universidade Federal de Minas Gerais, Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, Southern Methodist University, Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, Picuris Pueblo, Picuris Pueblo, Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, Department of Zoology, University of Cambridge

Archaeological, ethnographic, and anthropological evidence, along with oral tradition, establishes a direct connection between Ancestral Puebloans and present-day Indigenous Puebloans. However, policymakers often overlook or downplay this evidence when making decisions about Ancestral Puebloan archaeological sites and heritage. Despite the potential of ancient DNA to support these claims, decades of exploitative research ethics have led to mistrust among Indigenous peoples of North America, and the few genetic studies conducted so far have been driven by the necessity of repatriation claims rather than a desire to understand tribal histories. Here we report the first study in which an Indigenous group, the sovereign nation of Picuris Pueblo in New Mexico, initiated collaboration with geneticists and archaeologists to fill in gaps in traditional knowledge and further their understanding of their population history and ancestry. We generated 29 genomes from 16 ancient individuals spanning the last millennium and 13 present-day members of Picuris Pueblo. We show genetic continuity between the ancient and present-day Picuris, and more broadly between Picuris and other ancient individuals from Ancestral Puebloan sites in the region. Haplotype-based methods show no indication of a population decline before European contact, indicating that the depopulation of some Pueblo sites represented a reshuffling of peoples in the region rather than a demographic collapse. Our collaboration illustrates an approach that equally privileges and complements oral traditional, archaeological, ethnographic, and genetic lines of evidence while also ensuring the Indigenous communityâ€™s control of their genetic data.

944 - Adaptation in Ancient Maya

Esther Brielle

Esther S Brielle, Vera Tiesler, Keith M Prufer, Gabriel D Wrobel, Douglas J Kennett, Takeshi Inomata, Travis Stanton, Richard George, Aubree Marshall, David Reich

Harvard, Universidad Autonoma de Yucatan, University of New Mexico, Michigan State University, University of California, Santa Barbara, University of Arizona, University of California, Riverside, University of California, Santa Barbara, Michigan State University, Harvard

Our study analyzes ancient DNA from 421 Maya individuals, spanning 200 to 9500 years before present, to identify genetic adaptations to environmental and societal shifts. The consistent genetics of the Maya over millennia allow a general linear model to analyze genotypes over time. Preliminary results reveal significant selection in genes related to body weight, muscle mass, height, and metabolic traits, including glucose regulation and blood flow. Notably, chromosomes 6, 7, 9, and 18 exhibited strong selection signals, with chromosome 6 being particularly significant for genes such as E2F transcription factor 3, associated with body fat and muscle, CDK5 regulatory subunit associated protein 1-like 1, linked to diabetes and metabolic diseases, and neurensin 1, related to growth and various protein and blood-related measurements. Among the identified loci, cytochrome P450, family 3, subfamily A on chromosome 7, recognized for its role in hormone processing that results in increased strength and muscle mass, and nuclear factor, erythroid 2-like 3 pseudogene 1 on chromosome 18, related to heart size and ethanol metabolism, were also highlighted. These findings suggest adaptations to changes in diet and lifestyle associated with the transition from foraging to farming, urbanization, and long-distance trade, without reliance on domesticated pack animals. We are further investigating differences between elite and nonelite groups, highland vs. lowland adaptations, and the timing of these genetic adaptations, potentially in response to dietary shifts and microbial changes. This study will enhance our understanding of the Maya's evolutionary responses to their changing environment.

967 - Inferring fine-scale demography of indigenous lineages from 500 whole genomes across Mexico

Santiago G Medina-Muñoz

Santiago G Medina-MuÃ±oz, Amit Gourav Ghosh, Magnus Pharao Hansen, Gustavo Valadares Barroso, Carmina Barberena Jonas, Elena S Gusareva, Stephan C Schuster, Hie Lim Kim, Aaron P Ragsdale, Andres Moreno-Estrada
National Laboratory of Genomics for Biodiversity (LANGEBIO), Advanced Genomics Unit (UGA), CINVESTAV, Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, GenomeAsia 100K consortium, University of Copenhagen, Department of Cross-Cultural and Regional Studies, Department of Integrative Biology, University of Wisconsin-Madison, National Laboratory of Genomics for Biodiversity (LANGEBIO), Advanced Genomics Unit (UGA), CINVESTAV, Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, The Asian School of the Environment, Nanyang Technological University, GenomeAsia 100K consortium, Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, GenomeAsia 100K consortium, Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, The Asian School of the Environment, Nanyang Technological University, GenomeAsia 100K consortium, Department of Integrative Biology, University of Wisconsin-Madison, National Laboratory of Genomics for Biodiversity (LANGEBIO), Advanced Genomics Unit (UGA), CINVESTAV

Genetic studies of population history in the Americas have traditionally focused on early settlement and dispersal, often overlooking the local dynamics in areas with a rich Prehispanic heritage, like Mexico. This oversight misses critical insights into the genetic landscape shaped by these regional movements. To deepen our understanding of Mexico's demographic history, we sequenced 500 present-day genomes at 30x coverage, capturing a broad spectrum of indigenous ancestries nationwide. This effort represents the most comprehensive sequencing project of indigenous ancestries by leveraging fine-scale sampling of the MX Biobank.
By integrating these genomes with those from indigenous populations in South America and additional diverse global ancestries, we assessed a series of demographic hypotheses. We use gene-genealogy based methods for exploratory analysis and linkage

disequilibrium statistics to infer demographic models. Our demographic reconstruction is further enriched by insights drawn from linguistics, as well as historical and archaeological records.

We found previously unreported genetic relationships among Mesoamerican populations, particularly among the Nahuatl-speaking peoples who led one of the major expansions within Mesoamerica. We identified a shared genetic component among central Nahua populations, while also uncovering more distant connections: West Nahua groups exhibit a mixture of central and northern Mexican ancestries, whereas East Nahua groups blend central Nahua with Southeast Mayan ancestries. Additionally, our findings highlight genetic connections with South American populations, emphasizing the intricate demographic interplays shaping modern Mexico. This research provides new insights into Mexico's genetic landscape, with significant implications for historical narratives and genomic studies in indigenous populations.

145 - Population genomics of tequila and mezcal (Agave spp.): Unexpected findings and lessons for these Mexican industries in the face of climate change.

Luis E. Eguiarte

Luis E. Eguiarte
Universidad Nacional AutÃ³noma de MÃ©xico

Agave is a diverse genus, with more than 280 species and reports of use by humans for at least 140 species, usually as fibers, as food, and in particular for making alcoholic drinks as pulque (just fermented), mezcal or tequila (both distilled). Recently, there has been an impressive boom in the interest and production of both mezcales and tequila. Here we present the results of our populations genomics studies with tequila agave, A. tequilana in the state of Jalisco, espadÃ-n agave of Oaxaca (A. angustifolia), bacanora from Sonora (also A. angustifolia), tobalÃ¡ agave (A. potatorum) from Oaxaca and Puebla, tobasiche, cirial, cuije, etc. agave, all different names for A. karwiskii, from Oaxaca and Puebla, agave tepextate or pizometl (A. marmorata) also from Oaxaca and Puebla, and agave papalomÃ©tl or ancho (A. cupreata) from Guerrero and MichoacÃ¡n. We found very high levels of genetic variation and low geographic differentiation, and in some cases more inbreeding than expected, but diversity decreases with the level of management (the more intensively managed populations usually have less overall variation). In tequila and espadÃ-n, we detected that the plants are clonally propagated, but in these selected clones each plant is highly heterozygous. We are using these genomic data and detailed studies on the potential distribution of the species in the future along their genetic variation (i.e., genomic offset analyses) to propose areas where plants could be grown and cultivated in the future, which we believe will be important results for the producers of tequila and mezcal.

230 - Disentangling microbiome effects on plant adaptation

Alejandra Hernandez-Teran

Alejandra Hernandez-Teran, Maria Rebolleda-Gomez
University of California - Irvine, University of California - Irvine

Itâ€™s increasingly clear that microbes influence host health, fitness, and even important developmental transitions like metamorphosis in mosquitoes and flowering time. The timing of these transitions is central for organisms responses in changing environments. However, we donâ€™t know which microbes alter such traits and how they do it, whether itâ€™s adaptive for the microbes or a consequence of their metabolism. Understanding the role of microbial communities in developmental timing of other organisms can help us better predict, and even enhance adaptive responses to changing signals. Using life-history models and community selection, our goal is to understand microbial effects on plant developmental timing. One intriguing hypothesis is that since microbes are indicators of environmental and ecological conditions, they can provide information about optimal phenology to the plant hosts. We are using Arabidopsis thaliana to

determine how microbial communities can affect developmental timing. We sampled microbial communities from several environments and inoculated A. thaliana seeds to see the effect of those communities on development. Using directed evolution, we serially propagated microbial communities in planta and selected only those communities that promoted extreme phenotypes on developmental timing. After four transfers, we stabilized the community to obtain microbes that reliably accelerate, delay, or maintain flowering time with respect to uninoculated controls. Using omics approaches, we are characterizing the community and the metabolic functions that are associated with either late or early flowering and incorporating our insights into predictive life-history models to evaluate possible effects of microbes on plant adaptation to climate change.

602 - A boaring story of hybridization and adaptation in wild boars, domestic pigs and feral swine

Mirte Bosse

Mirte Bosse, Niek Barmentlo, Timothy J Smyser
VU University, Wageningen University, VU University, National Wildlife Research Center

The ecological threat posed by invasive species is well established, however evolutionary processes that contribute to the invasion potential of a given species are not as well understood. Wild pigs in the US serve as an ideal model to evaluate genomic and evolutionary processes associated with invasiveness given the unique molecular resources available for Sus scrofa, profound ecological and economic destruction caused by the species, and the global extent of their invasive spread. Within the United States, wild pigs overwhelmingly descend from the historical hybridization of wild boar and domestic pigs and have expanded rapidly over the past 40 years to invade as many as 38 states. We sought to evaluate the importance of genetic enrichment through hybridization to the perceived heighted invasiveness of the species, using thousands of genotyped feral swine. The vast majority of feral swine were of mixed ancestry, with dominant genetic associations to Western heritage breeds of domestic pig and European populations of wild boar. Mixed ancestry from wild and domestic lineages clearly contributes to the rapid expansion of invasive feral swine. populations. We discovered several genomic regions under selections that encode for phenotypic traits likely to impact fitness in the wild such as coat coloration and craniofacial development. These results elucidate genetic processes that facilitate establishment and subsequent evolutionary responses that contribute to invasive spread.

1000 - Strategies for mining useful alleles for climate change adaptation

Ayelet Salman-Minkov

Ayelet Salman-Minkov, Daniel Runcie, Jeffery Ross-Ibarra
UC Davis, UC Davis, UC Davis

Climate change and global warming pose significant challenges to future food security, as elite lines of major crop species are susceptible to heat and drought. A potential solution may lie within traditional varieties. Cultivated by local farmers worldwide for hundreds to thousands of years, traditional varieties possess adaptations to diverse climatic conditions. We would like to leverage these adaptations, identify alleles beneficial in future environments, and integrate them into the elite lines of crop species.
Our objective is to determine the most effective strategy for identifying and incorporating valuable alleles into elite lines, utilizing both geo-referenced and genomic data. To represent diverse, realistic genetic architectures, we utilize simulated genomic data representing populations evolving across various demographies and landscapes. We evaluate three breeding approaches: 1) genotype-environment association (GEA) followed by marker-assisted selection (MAS), 2) genomic prediction of local adaptation combined with QTL mapping in a simulated future environment followed by MAS, and 3) genomic prediction of local adaptation combined with genomic selection in a simulated future environment. We anticipated that the GEA approach would be effective when there are common alleles with large-effect

sizes, but the QTL mapping approach might be more effective when the population has many large-effect alleles, but all are rare. We found that the GEA approach was rarely better than the alternatives, while the QTL approach was best in nearly 50% of scenarios, regardless of genetic architecture. This suggests that the QTL approach may be the preferred choice for mining beneficial alleles.

213 - No passive sidekick: a substantial role for mitochondrial compensatory evolution during adaptation of ETC-deficient strains of Caenorhabditis elegans

Vaishali Katju

Vaishali Katju, Alberto Corral-LÃ³pez, Zachary P. Dietz, Suzanne Estes, Ulfar Bergthorsson
Uppsala University, Uppsala University, Portland State University, Portland State University, Uppsala University

The intergenomic coadaptation of the mitochondrial (mtDNA) and nuclear (nDNA) genomes is reliant on the dual genetic control of both nDNA- and mtDNA-encoded protein subunits comprising the mitochondrial electron transport chain (ETC). The molecular basis of these mitonuclear epistatic interactions requisite for maintaining key metabolic function remain poorly understood. We conducted a laboratory evolution experiment to investigate the functional and genomic patterns of mitonuclear adaptation in independent populations of Caenorhabditis elegans bearing four focal ETC mutants. Although a large number of both nDNA and mtDNA mutations were identified following 60 generations of laboratory adaptation, mutations in genes with assigned mitochondrial function were restricted to the mtDNA genome. A total of 77 mtDNA variants with intra-individual frequencies ranging from 5-100% were identified across 124 adaptation lines. In direct contrast to the patterns observed in C. elegans mutation accumulation lines under minimal selection, the adapted lines were enriched in base substitutions relative to small indels. The vast majority of base subsitutions in protein-coding genes were nonsynonymous changes (82.5%) with significantly higher heteroplasmic frequencies, suggesting potential targets for positive selection. Notably, novel mutations in the same ETC complex as the original background mutation reached significantly higher frequency than mutations in a different complex, suggestive of compensatory evolution. While the nuclear genome has been hypothesized as the primary driver of mitonuclear coevolution given its larger mutational target, our results argue for a disproportionately outsized role for mitochondria in mitonuclear adaptation despite their diminutive genome size and protein-coding capacity.

660 - Phylonumtomics uncovers diverse evolutionary trajectories of mitogenomic fossils buried in mammalian and avian genomes

Toni Gossmann

Toni Gossmann
TU Dortmund University

Sporadically genetic material that originates from an organelle genome integrates into the nuclear genome. However it is unclear what processes maintain such an integration over longer evolutionary time. Recently it was shown that nuclear DNA of mitochondrial origin (NUMTs) may harbour genes with intact mitochondrial reading frames despite the fact that they are highly divergent to the host's mitochondrial genome. Two major hypotheses have been put forward to explain the existence of such mitocoding nuclear genes: (A) recent introgression from another species and (B) long-term selection. To address whether these intriguing possibilities do play a role we scanned the genomes of more than 1,000 avian and mammalian species for NUMTs. We indeed identified that the subclass of divergent NUMTs harbouring mitogenes with intact reading frames are widespread across mammals and birds. We can show that for these NUMTs signatures of cross-species introgression are widespread in birds, but not mammals with the exception of ungulates. We can also show that a substantial fraction of deeply divergent NUMTs are maintained by selection. For a small number of NUMT genes we identify an evolutionary signature that is consistent with adaptive evolution including one human NUMT that is shared among seven ape species.

This highlights the intriguing possibility that NUMT insertions occasionally may be functional.

749 - Evolution of mitochondrial mRNA editing and linear multipartite mt-genome architecture in calcaronean sponges (invited talk)

Dennis V. Lavrov

Dennis V. Lavrov, Noah Gettle

Iowa State University, Wellcome Sanger Institute

Transcripts of mitochondrial genes in most calcaronean sponges undergo an unusual process of insertional mRNA editing in which 1-3 uridylates are inserted at specific sequence motifs. Although, phenotypically, the editing found in calcaronean sponges is similar to that in some other eukaryotes, its mechanism appears to be unique as both the occurrence and the pattern of editing (the number of inserted Us) can be predicted from the primary DNA sequence. In a sense, editing sites in calcaronean mtDNA are â€œgenetic abbreviationsâ€• that are expanded to the full sequence at the RNA level. Surprisingly, we found that editing in the group persists despite high rate of mt-sequence evolution and that the editing sites are gained and lost frequently both within and between species even though such events should lead to a frameshift in the coding sequence. Here we provide evidence that evolution of editing sites in calcaronean sponges is enabled by highly unusual mt-genome organization in these species, where most of the genes are located on separate linear chromosomes and multiple copies of each chromosome are transmitted from generation to generation. At the same time, there appear to be selection against nucleotide composition that promotes appearance of new editing sites. Overall, calcaronean sponges provides an interesting model for studying co-evolution between DNA- and RNA-level processes in animal mtDNA.

777 - Decoding the Puzzle of Human Somatic mtDNA Mutagenesis: Bridging Species Life-History Traits and Organ-Specific Cellular Properties

Konstantin Popadin

Konstantin Popadin, Bogdan Efimenko, Alina Mikhailova

Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University

Mitochondrial DNA (mtDNA) mutagenesis in human cancers is still enigmatic: external mutagens like UV light or tobacco smoke, known to impact nuclear DNA, don't seem to affect mtDNA. This discrepancy suggests a dominant, mitochondria-specific internal mutagen. To understand this, we analyzed mtDNA mutational spectra across a wide range of vertebrates, uncovering that some mutations (A>G on the heavy chain) are linked to internal, age-related chemical damage, likely oxidative (DOI: 10.1093/nar/gkac779). Others (C>T on the heavy chain) stem from internal replication processes (DOI: 10.1101/2023.12.08.570826). Leveraging our insights, we hypothesized that in somatic mtDNA mutagenesis, slow-dividing cells would exhibit more damage-related mutations, while fast-dividing cells would show a predominance of replication-associated mutations, i.e. we posited a variable damage-to-replication ratio (A>G/C>T) across tissues with different replication rates. To test this, we reanalyzed the two most comprehensive human mtDNA datasets: the TCGA, with 7,611 mutations across 37 cancer types and 21 tissues, and the GTEx, with 2,762 mutations in 25 healthy tissues. In both datasets, a higher damage-to-replication ratio was found in organs with slow-dividing cells (neurons, ovary, heart, skeletal muscles) and a lower ratio in those with rapid cell division (colon, intestine, stomach, vagina). Furthermore, a decrease in this ratio in late-stage somatic mutations across all cancer types (TCGA data) aligns with the anticipated increase in cell turnover during carcinogenesis. Our research shows that mtDNA mutations can reveal cell traits, unveiling their potential importance in biomedical applications.

927 - The contribution of within- and between-species mitochondrial variation to adaptation in

experimental Drosophila populations

Leah Darwin

Leah Darwin, Yevgeniy Raynes, Faye Lemieux , David Rand

Brown University, Brown University , Brown University, Brown University, Brown University

The oxidative phosphorylation system (OXPHOS) relies on both the mitochondrial and nuclear genomes to function. Due to its importance to cellular function, it has been hypothesized that strong purifying selection prevents functional genetic variation from accumulating in the mitochondria. This theory has been challenged by evidence of the contribution of mitochondrial variation to fitness differences. We use an evolve and resequence approach to quantify the role of within-species and between-species mitochondrial variation in adaptation to a novel environment. We constructed 24 outcrossed experimental populations (12 control, 12 experimental) from a founder population of Drosophila melanogaster with introgressed mitochondria (mt) DNA from 24 different strains of D. melanogaster from Zimbabwe or Beijing, 3 strains of D. simulans, and 1 strain of D. yakuba. The design of this study provided a replicate source of extensive nuclear and mtDNA variation. We subjected experimental populations to selection on rotenone, an insecticide that inhibits mitochondrial complex I in the OXPHOS system to specifically target a mitochondrial response. Phenotypic assays were performed on the founder lines and demonstrated fitness differences between mitochondrial haplotypes in the rotenone environment. After 20 generations of selection, pooled sequencing analyses of the evolved populations were used to quantify haplotype frequency dynamics of nuclear and mtDNA loci. We find the evolutionary trajectories to be contingent on the history of the experimental population. We report the mitochondrial and nuclear variants that have contributed to the differentiation between experimental populations.

1109 - Population-Genetic Environment of Genome Retention and Reduction in Mitochondria

Wei-Chin Ho

Wei-Chin Ho, Jeremy G. Wideman

University of Texas at Tyler, Arizona State University, Arizona State University

Originated from an endosymbiotic alphaproteobacterium during the eukaryotic evolution over two billion years ago, mitochondria play an important role in cellular physiology, such as oxidative phosphorylation. Based on comparative genomics studies, it has been inferred that the mitochondrial genome of Last Eukaryotic Common Ancestor (LECA) encoded proteins related to electron transport chain complexes, ribosomes, protein translocation and heme maturation. After LECA, the gene contents in mitochondrial genomes varies much across different eukaryotic taxa. While some cases of mitochondria-to-nucleus transposition were identified, many of these genes remain in the mitochondrial genomes in many lineages. Why there are very mixed outcomes of retention versus reduction in mitochondrial genome evolution remains controversial. To study this question, we ask what kind of population-genetic environment facilitates or constrain the mitochondria-to-nucleus transposition. Specifically, we used SLiM to simulate a population of haploid cells with intercompartmental duplication in mitochondrial and nuclear genomes and tracked which copy is lost earlier. The simulation results show the transposition can be facilitated by beneficial effects. Interestingly, even when transposition is deleterious, it can still be preferred with a much higher mutation rate in the mitochondrial genome compared to the nuclear genome. These findings have implications for our understanding of the mode and tempo of endosymbiont gene transfer.

52 - Host adaptation and genomic change in intestinal eukaryotes revealed from sequencing microbial mixtures

Abigail Lind

Abigail Lind, Nathan McDonald, Elias Gerrick, Ami Bhatt, Katie Pollard
Georgia Institute of Technology, Stanford University, Stanford University School of Medicine, Stanford University School of Medicine, Gladstone Institutes

The most prevalent microbial eukaryote in the human gut is Blastocystis, an obligate commensal protist also common in other vertebrates. Yet it is unclear how Blastocystis adapted from a free-living stramenopile ancestor to thrive within its wide range of hosts. To enable comparative genomics interrogation of changes throughout Blastocystis evolution, we cultured and sequenced six strains spanning the phylogenetic and host diversity of the genus. As establishing axenic cultures of Blastocystis is failure-prone and laborious, three strains were grown xenically alongside other microbes. To overcome the challenge of generating contiguous and contamination-free genomes from xenic cultures, we established a metagenomic assembly pipeline utilizing long-read, short-read, and Hi-C DNA sequencing, plus genome annotation leveraging RNA sequencing. Investigating differences in gene content and genome organization, we observe a transition within human Blastocystis towards minimal introns and intergenic regions, loss of canonical stop codons, gain of genes related to host temperature, and diversification of subtype-specific host-interfacing genes. In contrast, Blastocystis isolated from an herbivorous tortoise contained plant carbohydrate metabolizing enzymes, some horizontally acquired from bacteria, likely reflecting fermentation within the host gut. Finally, our analyses indicate that the common ancestor of all Blastocystis lost nearly all ancestral genes for flagella, resulting in its spherical morphology. Our findings reveal adaptations in different Blastocystis lineages for life in the gut of diverse vertebrates. The long-read hybrid assembly methods used here are effective tools for establishing genomics of eukaryotes recalcitrant to axenic culture.

111 - What's the Rate: Molecular Clock Calibration Using Sedimentary Ancient DNA

Maya Lemmon-Kishi

Maya Lemmon-Kishi, Rasmus Nielsen
University of California, Berkeley, University of California, Berkeley, University of Copenhagen

In many species there is a great deal of uncertainty regarding the rate of substitution, implying that the molecular clock is poorly calibrated. Historically, molecular clock calibration has required species divergence times to be known from the fossil record or from geologically dated events, limiting the use of molecular clocks for many species, including bacteria. Sedimentary ancient DNA (sedaDNA) is a rich source of genetic data that captures broad taxonomic diversity over long time transects. These transects of ancient DNA across time can, in theory, be used to calibrate the molecular clock for a large variety of species, including bacteria. However, such analyses are challenged by the fragmented and damaged nature of short read ancient DNA. Here, we introduce ratePlacer, a phylogeny-based method for analyzing sedaDNA that can combine the information from many short reads in a sample while accounting for DNA damages to provide maximum likelihood estimates of sample ages. Furthermore, the method can be used to calibrate the molecular clock when sample ages are known. Using a birth-death process to simulate ancient DNA, we show that ratePlacer is able to accurately estimate the sample age across time. Applying ratePlacer to eighteen well-dated sedaDNA samples spanning eighteen thousand years and comprehensive COI and 16S databases, we calibrate the COI and 16S molecular clock across various taxonomic groups. Our new calibrations of the molecular clock will be useful for dating split times in phylogenies in many taxonomic groups that previously were lacking information about molecular clock rates.

278 - Nitrogen cycling in symbiotic bacterial communities associated with ancient American plants

Edder Daniel Bustos-Diaz

Edder Daniel Bustos-Diaz, Francisco Barona-Gomez, Angelica Cibrian-Jaramillo

Leiden Univeristy, CINVESTAV Irapuato, Leiden University, Naturalis Biodiversity Center

Both the angiosperm family of Gunneraceae and the gymnosperm order of Cycadales can form nitrogen fixing symbiotic relationships with Nostocales cyanobacteria using specialized organs. Multiple studies, mostly on cycads, have shown that other bacterial taxa can also be found in the symbiotic organ, but not much information is known about the functional role of these associated bacterial communities. To address this, we sampled multiple cycads and Gunnera plants at different locations in order to inoculate co-cultures which could sustain the bacterial communities. These co-cultures where then sequenced through shotgun metagenomes in order to obtain metagenome assembled genomes (MAGs). Using these we were able to identify Hyphomicrobiales in the bacterial communities from different plants and locations are capable of partial denitrification, similarly to what has been found in the bacterial communities of the fern Azolla, another cyano-centric symbiotic system. Moreover, the phylogenetic analysis of these MAGs suggest that these Hyphomicrobiales could be under co-selection along with the Nostocales for the symbiosis, which might help to explain why some of these bacterial cyano-centric bacterial communities can jump from host to host with relative ease, even though each systems have their own particular symbiotic organs. The shared taxonomic and functional signals found in these communities suggest that even though Nostocales are key players in the symbiosis, the communities also play a role in the success of these symbioses.

702 - De novo assembly of ancient human metagenomes improves community diversity estimates and reveals varied microbial taxon-specific evolutionary trajectories

Irina M Velsko

Irina M Velsko, Keri Burge, Alexander HÃ¼bner, I-Ting Huang, Colleen Cavanaugh, Christina Warinner
Max Planck Institute for Evolutionary Anthropology, Harvard University, Max Planck Institute for Evolutionary Anthropology, Harvard University, Harvard University, Harvard

Ecological diversity estimates for ancient oral microbiomes have appeared highly similar to those of modern oral microbiomes when profiled with large and diverse taxonomic databases. However, artifacts in the profiling results have hinted at the presence of microbes that remain undetectable because they have no representative genome sequences in the databases used for taxonomic profiling. Hence, diversity estimates for ancient human microbiomes may be artificially low, potentially masking signals of evolutionary change and muddying comparisons to their modern counterparts. To improve our ability to detect and investigate oral microbiome evolution, we directly reconstructed genomes from modern (n=42) and ancient (n=455) oral metagenomes with de novo assembly to generate metagenome-assembled genomes (MAGs, n=8696 medium/high-quality) spanning 100,000 years of human history and from globally diverse populations on six continents. These MAGs reveal substantial undescribed microbial diversity present across ancient and modern human microbiomes (n=562 unnamed taxa), and enable us to better resolve microbiome community composition and to correlate community-level changes with human cultural, social, and demographic history. In addition, specific investigation of individual taxa reveals that phylogenetic patterns mirroring human migratory patterns are taxon-specific, and are stronger in novel taxa than clinical taxa that are well-described, as well as stronger in taxa that are rarely detected in heavily studied populations in Europe and the US. These results highlight the utility of incorporating ancient metagenomic data into human microbiome studies for revealing evolutionary signals at the levels of individual taxa as well as the entire microbiome.

752 - Genomic characterization of novel Treponema species from the oral microbiome of Aboriginal Australians

Davide Bozzi

Davide Bozzi, Samuel Neuenschwander, Gabriel Yaxal Ponce-Soto, David M Lambert, Eske Willerslev, Nicolas Rascovan, Anna-Sapfo

Malaspinas
University of Lausanne, Swiss Institute of Bioinformatics, University of Lausanne, Swiss Institute of Bioinformatics, Institut Pasteur, Griffith University, University of Copenhagen, University of Cambridge, University of Bremen, Institut Pasteur, University of Lausanne, Swiss Institute of Bioinformatics

Metagenomic analyses, over the last decade, have yielded a wealth of genomic information regarding microbial communities. The assembly of novel genomes from metagenomic data is rapidly outpacing the culture-based approaches in generating new genetic knowledge about unknown microbial species. In this study, we focused on Treponema, a genus of bacteria encompassing various environmental and host-associated taxa, including human symbionts and pathogens. We investigated novel high-quality metagenome-assembled genomes (MAGs) from the genus Treponema identified within the oral microbiome of Aboriginal Australian populations. Some of these MAGs represent species not found in previously studied human populations. We characterized their functional potential and we compared them from a phylogenetic and functional perspective to other species of the same genus to gain insights into their evolution and adaptation processes.

893 - Time-lapse of the microbial composition during agave fermentation using Meta-HiC

Mariana G Guerrero-Osornio

Mariana G Guerrero-Osornio, Maritrini ColÃ³n-GonzÃ¡lez, AgnÃ¨s Thierry, Alexander de Luna, Eugenio Mancera, Martial Marbouty, LucÃ-a Morales
Instituto de EcologÃ-a, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Posgrado en Ciencias BiolÃ³gicas, UNAM, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Institut Pasteur, Unidad de GenÃ³mica Avanzada (Langebio), Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Institut Pasteur, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano

In traditional Mexican distilled spirit production, the fermentation step is conducted in open tanks, promoting microbial interchange with the surrounding environment. This creates a dynamic succession of microbial communities, which, until now, have not been characterized at the genome-wide, especially with species-level resolution. Previous research indicates dynamic shifts in microbial diversity throughout the agave juice fermentation process-starting with a higher diversity and ending with a lower one as fermentation concludes.
The objective of this study is to map a genome-wide and species-level resolution of the succession of microbial communities during agave juice fermentation by utilizing the Meta-HiC technique, at seven consecutive time points throughout fermentation. We detail the relative abundance of Metagenome Assembled Genomes (MAGs) across different life domains. Predominantly, MAGs are linked to bacteria and yeast, with a notable prevalence of lactic acid bacterial groups. Intriguingly, we noted an uptick in the relative abundance of MAGs possibly originating from viruses in the fermentation's latter stages. Consistent with metagenomic studies, some data remain still unclassified, though this portion decreased over time. The Meta-HiC technique has facilitated the assembly of bacterial and yeast genomes at each phase of fermentation, allowing for an analysis of species' relative coverage and genomic diversity. These findings enhance our understanding of the microbial dynamics involving bacteria, yeast, and viruses in this unique fermentation process, providing essential insights into the ecological interactions driving this microbial succession

62 - A genealogy-based framework to estimate population structure and demographic history

Charleston W.K. Chiang

Charleston W.K. Chiang
University of Southern California

Many methods infer population structure or demographic history from genetic data using relatively low-dimensional summaries, such as the allele frequency spectra (AFS). In principle, the complete genomic history is encoded in the sequence of gene-genealogical trees, known as the ancestral recombination graph (ARG). As a step toward utilizing all available genomic information, we introduce two methods that leverage the ARG to infer population structure and demographic histories. First, we describe a framework to infer the expected relatedness between pairs of individuals given an ARG of the sample, which we call the eGRM. We show that the eGRM captures the structure of a population better than the canonical Genetic Relationship Matrix (GRM), even when only genotyping array data is available. Second, we devised a method called gLike, which employs a graph-based structure to summarize the relationships among all lineages in a tree with all possible trajectories of population memberships through time and to efficiently compute the probability under any parameterized demographic model. Through extensive simulations of multiple admixtures, we showed that gLike accurately estimates dozens of parameters such as ancestral population sizes and admixture timing, and outperforms conventional approaches based on the AFS. We applied both methods to real-world human genomic data from Finnish, Latino American, and Native Hawaiian cohorts to gain further insights into the patterns of population structure and to estimate parameters of the admixture histories. Taken together, our studies demonstrate the power of leveraging the genealogical trees for downstream population genetic inferences.

142 - Inference of ultra recent demography, recurrent mutation, and natural selection from biobank-scale site frequency spectra

Joshua G Schraiber

Joshua G Schraiber, Michael D Edge

University of Southern California, University of Southern California

Modern population genetic datasets have sample sizes in the tens or even hundreds of thousands of individuals. With such large sample sizes, many mutations will have multiple independent origins and the standard infinite sites assumption of population genetics becomes untenable. In addition, the vast majority of variants are extremely rare, and therefore sensitive to recent demographic history. Moreover, discovery of ultra rare variants provides an unprecedented window into the action of natural selection. Because of this, biobank-scale data presents an unprecedented opportunity to learn about recurrent mutation, recent demography and natural selection. We developed a new approach to modeling the site frequency spectrum of ultra rare variants (frequency less than 0.1%) in biobank scale datasets (>10000 samples). Our approach leverages the branching process approximation to the Wright-Fisher process. We show that finite sample probabilities follow an inhomogeneous birth death process with immigration and present an interpretable analytical solution that allows for efficient computation. When applied to data from gnomAD, we show that recurrent mutation common, and estimate rates of population growth in the last ~100 generations. We then estimate a distribution of fitness effects across different functional categories. As examples of downstream applications, we present a method for estimating the number of independent origins of a mutation subject to natural selection, and find that many rare, deleterious mutations have arisen multiple times independently in humans.

261 - Hidden Diffusion: Accurate Joint Inference of Selection and Demographic History from Time-series Data for Multiple Populations

Ekaterina Noskova

Ekaterina Noskova, Madleina Caduff, Andreas Fueglistaler, Anna Parker, Christoph Leuenberger, Daniel Wegmann

University, of Fribourg, Swiss Institute of Bioinformatics, University, of Fribourg, Swiss Institute of Bioinformatics, University, of Fribourg, Swiss Institute of Bioinformatics, University, of Fribourg, University, of Fribourg, University, of Fribourg, Swiss Institute of Bioinformatics

Recent advancements in sequencing technologies and ancient DNA preparation have enabled the generation of extensive time-series

dataâ€”genetic snapshots at various time pointsâ€”that promise to identify loci affected by selection.

Time-series allele frequency data is effectively modeled with Hidden Markov Models (HMM). Inference under Wright-Fisher HMMs involves calculating transition probabilities between frequency states and data likelihood given those probabilities. However, due to the large number of possible allele frequency states, analytical inference is computationally challenging. Several methods address this challenge: ApproxWF uses a hidden Markov model on a coarse allele frequency grid, while TimeSweeper is machine learning method trained on simulated data under explicit demographic histories.

We introduce a novel method based on hidden diffusion, employing a single partial differential equation solver to calculate both transition probabilities under Wright-Fisher diffusion and likelihood. This approach is faster than ApproxWF and more accurate than TimeSweeper in jointly inferring demographic parameters and locus-specific selection in a single population. Moreover, hidden diffusion identifies loci under selection in models involving multiple populations connected by migration and with potentially different selection coefficients. It can identify loci under divergent selection between populations, where single-population time-series data may be uninformative.

Implemented in user-friendly software, our method also models auto-correlation in selection coefficients along the genome using a second-layer HMM to account for linkage between neighboring loci. We demonstrate its utility by identifying selection targets from a time-series of ancient human samples from southern Germany spanning from 2000 BC to modern times.

293 - Understanding the interactions between demography and selection landscapes in human populations: insights from Ancestral Recombination Graphs (ARGs)

Yun Deng

Yun Deng, Vince Buffalo, Rasmus Nielsen
UC Berkeley, UC Berkeley, UC Berkeley

Both demography and Background Selection (BGS) have profound influence on the diversity patterns along the genome. Demography has a consistent genome-wide effect while the effect of BGS on diversity depends on the local strength of selection and linkage patterns. However, the joint effects of demography and BGS are not well understood: the traditional theory of BGS leading to local diversity reduction maps (B-maps), assumes a constant population size. Also, almost all coalescent-based demographic inference methods (e.g., standard PSMC and MSMC) cannot incorporate BGS effects. Here we investigate a new theory of a distorted coalescence process that incorporates BGS and we implement an inference method for ARGs that accommodate known genomic maps of BGS. We show that this method can lead to accurate coalescent-based demography inference in the presence of BGS. We also show that ARG inference allows accurate fine-scale (kb scale) estimation of diversity levels, which otherwise is difficult to obtain from genotype data. Fine-scale diversity also facilitates solutions to the reverse problem: learning the full BGS landscape from observed diversity patterns, rather than summarizing them as Distribution of Fitness Effects (DFE). Joint modeling of demography and BGS has the potential to provide estimates of demography under BGS, obtain more accurate fine-scale B-maps, and learn region-specific BGS coefficients. We illustrate the utility of the method by exploring the effect of BGS on inference of population size changes in different populations and investigate their differential selection landscapes, using published human genomic data.

420 - Linkage equilibrium between rare alleles

Anastasia Lyulina

Anastasia Lyulina, Zhiru Liu, Benjamin Good
Stanford University, Stanford University, Stanford University , Chan Zuckerberg Biohub

Patterns of linkage disequilibrium contain a wealth of information about the evolutionary forces acting within a population. Much of our existing understanding has focused on pairwise correlations between alleles, and how they decay as a function of the recombination rate. However, the magnitudes of these correlations are strongly influenced by other evolutionary forces like natural selection and genetic drift, making it difficult to tease out the effects of recombination (or vice versa). Here we introduce a theoretical framework for studying an alternative class of summary statistics that explicitly quantify the homoplasy produced by recombination. We derive analytical expressions that predict how these homoplasy metrics depend on the rates of recombination and recurrent mutation, the strength of selection and genetic drift, and the present-day frequencies of the two alleles. We find that the degree of homoplasy can strongly depend on this frequency scale, which reflects the underlying timescales over which these mutations occurred. We show how these scaling properties can be used to disentangle the rates of recombination, and discuss their implications for understanding the rates of horizontal gene transfer in bacteria.

851 - Bayesian mediation models reveal patterns of host genomic disease adaptation driven by environment

M. Elise Lauterbur

M. Elise Lauterbur, David Enard
University of Arizona, University of Arizona

Global patterns of pathogen prevalence are shifting in response to climate change and other human activity. Traditional epidemiological and disease ecological approaches used to examine these impacts are typically restricted to a modern timescale and narrow scope of host, pathogen, or environment diversity. By using host genomes as a record of past interactions between hosts and pathogens, we can extend disease ecological analyses to investigate these relationships over evolutionary time in the context of the hostâ€™s environment. We develop this eco-evolutionary framework applied to a dataset of genomes from over 200 host mammal species occupying a wide range of habitats. We quantify the impact of virus exposure across species and environments using genomic adaptation at virus-interacting proteins (VIPs) with a novel application of a Bayesian mediation model structure. This allows linking ecological and environmental factors with viral impact across evolutionary time, controlling for genome-wide adaptation. We show that there are broad, general effects of ecological factors including precipitation seasonality, host diet, and viral transmission mode on the impact of virus exposure as measured through host genomic disease adaptation. This work has important implications for understanding the distribution of viruses across host and virus types worldwide, and for predicting future changes in virus exposure both to humans and other organisms in response to changing environments. In addition, this framework can be generalized to other questions for which genomic adaptation is an expected response that can be correlated to ecological conditions.

272 - Accurate inference of population history in the presence of background selection

Arun Durvasula

Trevor Cousins, Daniel Tabin, Nick Patterson, David Reich, Arun Durvasula
Cambridge University , Harvard University, Harvard University , Harvard University , Harvard University, USC

Most published methods for learning about demographic history make the simplifying assumption that the genome evolves neutrally, and do not seek to account for the effects of natural selection on patterns of variation. This is a major concern, as ample work has

demonstrated the pervasive effects of natural selection and in particular background selection (BGS) on patterns of genetic variation in diverse species. Simulations and theoretical work have shown that methods to infer changes in effective population size over time (Ne(t)) become increasingly inaccurate as the strength of linked selection increases. Here, we introduce an extension to the Pairwise Sequentially Markovian Coalescent (PSMC) algorithm, PSMC+, which explicitly co-models demographic history and natural selection. We benchmark our method using forward-in-time simulations with BGS and find that our approach improves the accuracy of effective population size inference. Leveraging a high resolution map of BGS in humans, we infer considerable changes in the magnitude of inferred effective population size relative to previous reports. Finally, we separately infer Ne(t) on the X chromosome and on the autosomes in diverse great apes without making a correction for selection, and find that the inferred ratio fluctuates substantially through time in a way that differs across species, showing that uncorrected selection may be an important driver of signals of genetic difference on the X chromosome and autosomes.

540 - Drivers of historical population declines in Australian Marsupials: are humans behind the wheel?

Toby G. L. Kovacs

Toby G. L. Kovacs, Simon Y. W. Ho, Carolyn J. Hogg
The University of Sydney, The University of Sydney, The University of Sydney

Characterising the historical context of extant biodiversity can help to explain the vulnerability of certain groups to extinction and provide insight into how populations will respond to future environmental change. Owing to the historic geographical isolation of Australia, the island harbors a unique and outstanding mammal biota. Nevertheless, Australia has the worst mammal extinction rate in the world, with many remaining species now restricted to small, isolated populations. Previous demographic estimates have suggested that populations of Australian marsupials declined in the late Pleistocene, coinciding with the arrival of humans on the continent. However, the reliability of these demographic inferences is uncertain, given that they need to be scaled using estimates of species-specific mutation rates. To address this, we sequenced parent-offspring trios for two marsupial species of high conservation concern, and modeled mutation rates across marsupials. Using these new mutation rate estimates, we inferred population sizes across marsupials over the last million years. Our results suggest that many marsupial species experienced large population declines before the arrival of humans, while highlighting the impact of historical climate change on past population health. These findings could explain the vulnerability of Australian marsupials to extinction and raises concern for already struggling marsupial populations under a currently changing climate.

747 - AI and demographic inference in structured populations

Alba Nieto Heredia

Alba Nieto Heredia, Stefano Mona, Oscar Lao
ISYEB, EPHE, UniversitÃ© PSL, Sorbonne UniversitÃ©, MNHN, IRD, CNRS, 75005 , Institut de SystÃ©matique, Evolution, BiodiversitÃ©, ISYEB (UMR 7205), MusÃ©um National d'Histoire Naturelle, CNRS, Sorbonne UniversitÃ©, EPHE, UniversitÃ© des Antilles, Paris, France, Laboratoire dâ€™Excellence CORAIL, Papetoai, French Polynesia., EPHE, PSL Research University, Paris, France, Institute of Evolutionary Biology (UPF-CSIC), PRBB, Barcelona, Spain

Gaining insights into the historical demographic dynamics helps to identify potential threats to species survival and to implement targeted conservation measures. Typically, the main metrics to assess the conservation status from genetic data is the variation in effective population size through time. This is related to the inverse of the coalescence rate (CR) which ultimately depends on the true demographic history of a species.

SMC-based algorithms predict the variation of CR through time employing population genetics summary statistics obtained from genomic data. However, their performance in structured meta-populations remains poorly investigated.
Moreover, under panmixia, variations in CR correspond to changes in population size over time, which has a simple biological interpretation. Conversely, the interpretation of CR variation in structured population is challenging, because it depends on the joint effect of all evolutionary forces acting on the meta-population. Fully understanding of the CR variation is of paramount importance, claiming for novel approaches to interpret the CR and integrate it with other population genetics statistics.
Here we evaluated the accuracy of SMC-based algorithms by testing their performance in complex structured demographic scenarios for which the true CR was generated by coalescent simulations. We observed that SMC-based algorithms performance varies depending on the demographic parameters. We then used a Deep Learning (DL) model to correct SMC-predictions to be used in an AI model integrating information with other summary statistics, such as site frequency spectrum (SFS) to infer the demographic history of a population from its observed genetic variation.

993 - Detecting structural variation in reconstructed genealogies

Anastasia Ignatieva

Anastasia Ignatieva, Martina Favero, Jere Koskela, Jaromir Sant, Simon R Myers
University of Glasgow, University of Warwick, University of Warwick, University of Oxford, University of Oxford

In the presence of recombination, the genealogy of a sample of sequences can be fully captured in the form of an ancestral recombination graph (ARG). Following recent breakthroughs, it is now possible to reconstruct genome-wide ARGs from large-scale sequencing datasets, and new ARG-based statistical inference methods promise powerful insights into evolutionary events and parameters. Genomic structural variants (SVs) are a ubiquitous and evolutionarily important type of mutation, however their presence is ignored by all currently available ARG reconstruction tools. Through developing new theoretical results on the distribution of the genomic span of a haplotype block, using the powerful framework of the SMCâ€™ model, we present an accurate ARG-based tool for detecting SVs which result in the suppression of recombination in heterozygotes (such as chromosomal inversions). By applying this tool to ARGs reconstructed using Relate for the 1000 Genomes Project data, we detect several known and putative new inversions with high confidence, allowing for the genealogy-based analysis of their age and evolutionary history. We also detect other SVs, in particular copy number variants, by identifying and scanning for the characteristic artefacts they induce in reconstructed ARGs. Our results definitively demonstrate that reconstructed ARGs capture the signal of SVs, even though their presence is not explicitly modelled by the ARG reconstruction method. Identifying and analysing the evolution of SVs at all scales is an incredibly important goal, and our results provide the first evidence that reconstructed ARGs can be leveraged for this purpose.

54 - Deciphering Sorghum Domestication: The Pivotal Role of Condensed Tannins in Co-evolution Among Plants, Humans, and Birds

Xianran Li

Xianran Li
USDA-ARS WHGQ

Among major cereal crops, sorghum is the only one with a high proportion of cultivars containing condensed tannins, potent secondary metabolites inducing bitterness perception, in seeds. This presentation explores a case of co-evolution among plants (domesticate), humans (domesticator), and birds (biotic environmental conditions) to unravel this evolutionary mystery.

An examination of the geographic distribution of 11,557 sorghum varieties revealed that tannin sorghum are predominantly found in East and South Africa, while non-tannin sorghum predominates in West Africa. Analysis of functional polymorphisms of two key genes suggested that smallholders in different African regions independently domesticated non-tannin sorghum.

Further investigation unveiled a parallel geographic distribution between sorghum and the pest Quelea quelea in Africa. Q. quelea, the worldâ€™s most abundant wild birds, mainly inhabits East and South Africa, but not West Africa. Smallholders in East and South Africa selected tannins to combat severe herbivore threats.

To address how African smallholders could embrace "bad taste" tannin sorghum as a staple food, a parallel geographic distribution between bitter taste receptor (TAS2Rs) alleles in humans and sorghum was uncovered. Ancestral non-taste alleles are predominantly observed in East and South Africa, while derived taste alleles sensitive to tannins mainly originate from West Africa.

Integration of these coevolutionary relationships suggested that in East and South Africa, smallholders (perceiving less bitterness due to torelant TAS2Rs) selected tannins to combat agricultural pests, while in West Africa with mild threats from Q. quelea, smallholders (carrying sensitive TAS2Rs) selected non-tannin sorghum for better food.

117 - Meta-analysis of wild and domesticated crop phenotypic spaces

Arthur Wojcik

Arthur Wojcik, Stella Huynh, Ewen Burban, Karine Alix, Pierre GÃ©rard, Amandine Cornille, Harry Belcram, AgnÃ¨s Rousselet, Martine Le Guilloux, Rebecca Stevens, Marie Brault, Sylvain GlÃ©min, Yves Vigouroux, Domenica Manicacci, Catherine Dogimont, Maud Tenaillon
GQE - Le Moulon, IRD, EcoBio, GQE - Le Moulon, GQE - Le Moulon, GQE - Le Moulon, GQE - Le Moulon, GQE - Le Moulon, GQE - Le Moulon, GAFL, GAFL, EcoBio, IRD, GQE - Le Moulon, GAFL, GQE - Le Moulon

Plant domestication has led to a divergence between domestic (D) and wild (W) forms, at phenotypic and genomic levels. While the phenotypic domestication syndrome has been extensively documented, comparative analysis of phenotypic spaces between species have yet to be undertaken. The underlying pattern of phenotypical evolution have also to be linked to biological traits. Here,
we explore the phenotypic spaces of 14 crop-wild pairs of diploid
plant species with various mating systems and life-history traits, allowing comparative analyses. We measured traits on 20 W and 20 D populations in a randomized-block design. We performed variant calling on 40 W/D short-read sequenced genomes. We observed a marked disjunction of W/D phenotypic spaces in all systems. However, the strength of the phenotypic domestication syndrome depended neither on the genomic divergence nor the time to domestication. Second, the relative breadth of W/D phenotypic spaces followed a variety of trend -from smaller D than W in African rice to larger one in grape- that did not relate to mating system or life-history traits. The D/W genomic diversity ratio was <1 in all systems except for perennials, as chosen species did not suffer from domestication bottlenecks. It also does not correlate to the ratio of domestic to wild phenotypic space. Altogether, our results show uncorrelated patterns between genomic and phenotypic domestication syndromes. However, analysis of traits variance-covariance matrices reveals an uncoupling between W/D matrices as time to domestication increases, suggesting that phenotypic evolution has impacted genetic correlations.

709 - Wheat domestication and adaptation to new environments: the role of wild emmer.

Laura Botigué

Laura BotiguÃ©, Alice Iob, Nathan Wales, Ferran AntolÃ-n, Jana Helmbold-Doye, Marta Ortigas
Centre for Research in Agricultural Genomics, Centre for Research in Agricultural Genomics, NA, University of York, German Institute of Archaeology, Egyptian Museum of Berlin, Centre for Research in Agricultural Genomics

Wild emmer wheat (Triticum turgidum subsp. dicccoides) stands among the earliest plant species to give rise to domesticated forms in the Fertile Crescent, approximately 11,000 years ago. Today, one of its descendants, bread wheat, is a staple crop of immense socio-economic importance, contributing to approximately 20% of the world's caloric intake. The domestication of wheat has long captivated interdisciplinary scholars, and the prevailing theory suggests a protracted and geographically diffused process spanning millennia. In this study, we delve into the population structure of wild emmer populations and their impact on the domestication process and subsequent dispersal of the first domestics. Our investigation involves quantifying, for the first time, the genomic proportion of ancestry from wild emmer populations in two distinct germplasms, encompassing landraces from Europe, Africa, and Asia. We also provide estimates of the time since admixture of these two wild populations, aligning with the domestication and dispersal patterns observed, particularly towards Africa. Furthermore, we explore the potential adaptive role of wild emmer from the Southern Levant in domestication and dispersal, revealing an enrichment of genes associated with resistance to biotic stress and drought. Finally, we analyze the exome of 2,000 year-old wheat specimens from Egypt to trace the presence of these putative adaptive haplotypes. Our findings deepen our understanding of the origins of domestic wheat and underscore the significance of modern domestic landraces of emmer wheat in elucidating the genetic basis of resilience in cultivated crops.

1026 - Multipurpose species domestication in Mayan homegardens

Miriam Monserrat Ferrer

Miriam M Ferrer, Hector Estrada-Medina, Juan J JimÃ©nez-Osornio, Patricia Irene MontaÃ±ez-Escalante, RocÃ-o Ruenes-Morales
Universidad AutÃ³noma de YucatÃ¡n, Universidad AutÃ³noma de YucatÃ¡n, Universidad AutÃ³noma de YucatÃ¡n, Universidad AutÃ³noma de YucatÃ¡n, Universidad AutÃ³noma de YucatÃ¡n

During domestication, plants are expected to show a reduction in genetic diversity and in the microbiota associated with them in the agroecosystems in which they are managed, compared to that of forest populations. However, initial evaluations of genetic variability in Brosimum alicastrum (ox in Maya), Cordia dodecandra (k'oopte' in Maya), Selenicereus undatus (woob in Maya), and Spondias purpurea (abal in Maya) in homegardens in Yucatan suggest that the species have similar genetic diversity to sympatric forest populations due to gene flow among wild and managed individuals. We have also found that the most abundant bacterial and fungal endophytes of leaves in C. dodecandra are similar in individuals from homegardens and forests, even when they are grown in distant regions. These results suggest that native species have in homegardens an agroecosystem that functions as a hologenomic reservoir. On the other hand, we observed different cytotypes in B. alicastrum individuals from homegarden and forest and a ploidy series in S. purpurea (2n, 3n and 4n) varieties growing in homegardens. Also, in soils and rhizosphere associated with C. dodecandra, bacterial and fungal communities have a higher turnover, which may be associated with specific management practices of each agroecosystem. We believe that omics studies can help us to understand how management by the Mayan people promotes agrobiodiversity by maintaining the microbiome and increasing genomic variability of species with biocultural importance.

46 - Microbial dynamics in mangrove sediments during the anthropogenic era

Mirna Vázquez-Rosas-Landa

Mirna VÃ¡zquez-Rosas-Landa, Rosela PÃ©rez-Ceballos, Leonardo PÃ©rez-GonzÃ¡lez, Julio Canales-Delgadillo
Instituto de Ciencias del Mar y LimnologÃ-a, Instituto de Ciencias del Mar y LimnologÃ-a, Instituto de Ciencias del Mar y LimnologÃ-a,

Instituto de Ciencias del Mar y LimnologÃ-a

Mangrove ecosystems, situated in the tropics, play a critical role in global weather patterns and serve as essential carbon reservoirs. However, they face imminent threats from human activities like altered hydrological regimes and extensive deforestation. The symbiotic relationship between mangrove trees and microbiota within these ecosystems is vital for mutual survival. This study delves into interaction dynamics within degraded mangrove ecosystems due to hydrological connectivity loss. Using 16S rRNA sequencing, we analyzed microbial communities across various degradation stages (conserved, moderate-degradation, and degraded) during dry and rainy seasons at three sediment depths interacting with mangrove roots. Our analysis revealed 11,469 Operational Taxonomic Units (OTUs), indicating that microbial community structure is primarily influenced by degradation degree, with seasonal variations also playing a role. Significant diversity shifts were observed within the upper sediment layers, with conserved sites dominated by Vibrionaceae and moderate-degradation sites by Halomonas and Marinomonas, known halophytes. The presence of Vibrionaceae in conserved sites suggests potential urban contamination, while microbial diversity in moderate-degradation sites correlated with the dry season. Interestingly, a core community of Firmicutes persisted in deeper sediment layers across all degradation scenarios, hinting at its potential as a seed community for ecosystem restoration. These findings provide insights into microbial community responses to human-induced stressors and underscore the role of core microbial communities in guiding restoration strategies for degraded mangrove ecosystems.

395 - Within-patient evolution of ciprofloxacin resistance in Pseudomonas aeruginosa across a large-scale clinical trial

Matthew J Shepherd

Matthew J Shepherd, Niamh E Harrington, Anastasia Kottara, Kendall Cagney, Taoran Fu, Steve Paterson, Joanne L Fothergill, James D Chalmers, Dylan Z Childs, Michael A Brockhurst

University of Manchester, University of Liverpool, University of Manchester, University of Liverpool, University of Manchester, University of Liverpool, University of Liverpool, University of Dundee, University of Sheffield, University of Manchester

The antimicrobial resistance (AMR) crisis threatens to endanger modern medicine within the next 20-30 years. A key part of the problem is the emergence of AMR within patients themselves, during courses of antibiotic therapy. Whilst within-patient AMR emergence is increasingly receiving attention in the literature, investigations commonly involve a small number of patients and study single-patient cases which may poorly represent the more general patient population. Here we report finding on the evolution of ciprofloxacin resistance by Pseudomonas aeruginosa across a large-scale clinical trial. This trial utilised inhaled liposomal ciprofloxacin as therapy for patients with the lung condition non-cystic fibrosis bronchiectasis (NCFB), who suffered with P. aeruginosa infection. We isolated a total of ~25,000 independent bacterial colonies, prior to treatment and during the year-long duration of the trial, and measured the ciprofloxacin MIC and growth rates in KB broth for each of these isolates. Across the trial, ciprofloxacin MICs increased during periods of treatment and decreased during treatment withdrawal, in a pattern suggestive of resistance fitness trade-offs operating within patients. We also find that distinct phenotypic adaptive trajectories are followed in different patient cases, indicating diverse evolutionary dynamics driving resistance emergence. These findings allow us to start characterising the mechanisms through with AMR can emerge within-patients during treatment, and to ask questions on how best we could predict resistance emergence in the future.

455 - Evolution of the human gut resistome across diverse lifestyles and environments

Rahgavi Poopalarajah

Rahgavi Poopalarajah, Muhammad Arshad, Aashish R Jha

The New York University, The New York University, The New York University

Antimicrobial resistance is an evolutionary adaptation that is modulated by the gut microbiome. The gut microbiome harbours the gut resistome. The gut resistome refers to the antimicrobial resistance genes from pathogenic and non-pathogenic microbes in the gut. Microbes can acquire antimicrobial resistance genes on mobile genetic elements via horizontal gene transfer. The gut flora is a known reservoir of antimicrobial resistance genes, and the transfer of antibiotic resistance genes can occur from commensal gut microbes to opportunistic pathogens in the gut. There are several reports of lifestyle and environmental factors that influence gut microbiome composition and function, yet changes in the gut resistome due to lifestyle and environment are less known. Here, we evaluate the gut resistomes of global populations across lifestyle gradients and environments. Initially, by metagenomic sequencing of 97 fecal samples from healthy and prediabetic individuals, aged 18-70 years, we characterized the gut resistomes of the Western community. Our preliminary findings show that in spite of health status, Western adults constitute a high prevalence of glycopeptide resistance genes (~68% of total antimicrobial resistance genes) in the gut. Surprisingly, we discovered many vancomycin resistance genes. Vancomycin is last resort antibiotic against gram-positive bacteria. Notably, Eubacteriales and Clostridiales contained four or more vancomycin resistance genes, of which vanT was the most abundant gene type. Future work will include identifying gut resistome signatures linked to diverse lifestyles and environments. To conclude, our results may provide more insights into the evolution of drug resistance in the host gut.

518 - Feather microbiota compositional and functional landscapes across phenology and biogeography in Magellanic penguins in the Magellan Strait, Chile

Manuel Ochoa-Sánchez

Manuel Ochoa-SÃ¡nchez, Rosalinda Tapia-LÃ³pez, Ulises E. RodrÃ-guez Cruz, Jorge Acevedo, Claudio A. Moraga, Eliana Paola AcuÃ±a Gomez, Katherine Gaete, Luis E. Eguiarte, Valeria Souza

Instituto de EcologÃ-a, UNAM , Centro de Estudios del Cuaternario de Fuego, Patagonia y AntÃ¡rtica (CEQUA), Posgrado en Ciencias BiolÃ³gicas, UNAM, Instituto de EcologÃ-a, UNAM, Instituto de EcologÃ-a, UNAM, Posgrado en Ciencias BiomÃ©dicas, UNAM, Centro de Estudios del Cuaternario de Fuego, Patagonia y AntÃ¡rtica (CEQUA), Centro de Estudios del Cuaternario de Fuego, Patagonia y AntÃ¡rtica (CEQUA), Centro de Estudios del Cuaternario de Fuego, Patagonia y AntÃ¡rtica (CEQUA), Centro de Estudios del Cuaternario de Fuego, Patagonia y AntÃ¡rtica (CEQUA), Instituto de EcologÃ-a, UNAM, Instituto de EcologÃ-a, UNAM, Centro de Estudios del Cuaternario de Fuego, Patagonia y AntÃ¡rtica (CEQUA)

Phenology and environmental characteristics are important factors driving wildlife microbiomes. However, studies addressing the effect of both phenology and environmental variation in wildlife microbiomes are scarce. Here, we sequenced the hypervariable region V3-V4 of the 16S rRNA gene to profile the nest soil and feather microbiota of Magellanic penguins across three phenological times (i.e., courtship, egg incubation, and chick rearing) and five islands with contrasting environmental characteristics in the Magellan Strait, Chile. In addition, we sequenced the global metatranscriptome of penguin feathers (one metatranscriptome per island/time). Preliminary analysis indicates that Moraxellacea bacteria are the most abundant taxa in the feather microbiota across all conditions. Moreover, we detected seven core bacterial taxa associated with penguin bodies across all islands. Pyschrobacter was the core taxa with the highest abundance across all conditions. Furthermore, we also detected island-specific core bacterial taxa in four islands. Among these core taxa, Tomitella, Paeniglutamicibacter, Gelidibacter, Chryseobacterium, and Flavobacterium were shared between these. In turn, feather microbiota compositional patterns differed across islands. It consistently differed from the nest soil and across phenological stages in Contramaestre and Rupert Islands, whereas it had more variable patterns in the other islands. Ongoing analyses are being carried out to test the processes influencing feather microbiota composition, and the global function of penguin feather microbiota through metatranscriptomics.

108 - Compensatory mutations potentiate constructive neutral evolution by gene duplication

Christian R Landry

Christian R Landry, Philippe C DesprÃ©s, Alexandre DubÃ©, Rong Shi, Marie-Ãˆve Picard, Jordan Grenier
UniversitÃ© Laval, UniversitÃ© Laval, UniversitÃ© Laval, UniversitÃ© Laval, UniversitÃ© Laval, UniversitÃ© Laval

Protein functions generally depend on their assembly into complexes. During evolution, some complexes have transitioned from homomers encoded by a single gene to heteromers encoded by duplicate genes. This transition could occur without adaptive evolution through intermolecular compensatory mutations. Here, we experimentally duplicate and evolve an homodimeric enzyme to examine if and how this could happen. We identify hundreds of deleterious mutations that inactivate individual homodimers but produce functional enzymes when co-expressed as duplicated proteins that heterodimerize. The structure of one such heteromer reveals how both losses of function are buffered through the introduction of asymmetry in the complex that allows them to subfunctionalize. Constructive neutral evolution can thus occur by gene duplication followed by only one deleterious mutation per duplicate.

171 - How repeats rearrange chromosomes in deer mice

Landen Gozashti

Landen Gozashti, Olivia Harringmeyer, Hopi E Hoekstra
Harvard University, Harvard University, Harvard University

Large genomic rearrangements, such as chromosomal inversions, often underlie karyotype variation, but the mechanisms by which these rearrangements arise remain poorly understood. To study the origins of inversions, we generated chromosome-level de novo genome assemblies for four subspecies of deer mice (Peromyscus maniculatus) with known inversion polymorphisms and identified ~8,000 inversions, including 47 mega-base scale inversions. Analysis of inversion breakpoints suggests that while most small (<1 Mb) inversions arise via ectopic recombination between retrotransposons, large (>1 Mb) inversions are primarily associated with segmental duplications (SDs). Large inversion breakpoints frequently occur near centromeres, which may be explained by an accumulation of transposable elements in pericentromeric regions driving SD formation. Additionally, multiple massive inversions likely arose from ectopic recombination between near-identical centromeric satellite arrays located megabases apart, a previously uncharacterized mechanism of inversion formation. Together, our results illuminate how repeats give rise to chromosomal rearrangements, shaping the propensity for massive shifts in chromosome architecture.

215 - Inferring the characteristics of ancient polyploidization events from modern plant genomes

Tamsen Dunn

Tamsen Dunn, Arun Sethuraman
San Diego State University, San Diego State University

Multiple rounds of whole genome duplication (WGD) have occurred throughout angiosperm evolution. The majority of flowering plants appear to have evolved from allopolyploids (hybrids with multiple subgenomes derived from different species), while only a few genera are ancient autopolyploids (multiple subgenomes derived from the same species). Ancient WGDs are commonly detected in a modern genome by analyzing the levels of divergence (Ks) between pairs of duplicate genes. Since Ks is a proxy for time passed, modes in the histogram of these data indicate a spike in contemporaneous gene duplication. WGDs may be detected by fitting mixture models to identify peaks in Ks distributions. However, there are several limitations to this technique, one of which is that the Ks peak gives the divergence time between the duplicated subgenomes, not the actual time the subgenomes coalesce in an ancestor, which for

allopolyploids can be a difference of several million years. Furthermore, allo and autopolyploids follow distinct molecular evolutionary trajectories. Thus surveys which seek to generalize the consequences of polyploidy must appropriately take into account the time and mode of speciation. To address this concern, I propose and test a novel, scalable analysis method to determine the mode of ancient speciation from the shape of the Ks histogram of a modern genome. Preliminary work so far has demonstrated that our prototype method provides a better fit to Ks curves from the 1000 plant transcriptome dataset as compared with previous methods, and works well for both allo and autopolyploid-derived species.

227 - Investigating Functional Convergence Across Millions Of Years Of Evolution

Navya Shukla

Navya Shukla, Laura E Cook, Maddy L Comerford, Charles Y Feigin, Andrew J Pask, Irene Gallego Romero
School of BioSciences and Melbourne Integrative Genomics, University of Melbourne, St. Vincent's Institute of Medical Research, Lawrence Berekely National Laboratory, School of BioSciences and Melbourne Integrative Genomics, University of Melbourne, St. Vincent's Institute of Medical Research, LaTrobe University, School of BioSciences and Melbourne Integrative Genomics, University of Melbourne, St. Vincent's Institute of Medical Research, School of BioSciences and Melbourne Integrative Genomics, University of Melbourne, Center for Genomics, Evolution and Medicine, Institute of Genomics, University of Tartu

Understanding the molecular mechanisms underpinning convergent traits can help characterize forces that drive adaptation and provide insight into the predictability of the evolutionary process. The extinct thylacine, or Tasmanian tiger, has been long known to have a very similar morphology to that of the eutherian canids, particularly in the craniofacial region, despite having shared a common ancestor ~160 million years ago. Previous comparative genomics studies between the thylacine and a representative canid, the gray wolf, have found that much of underlying molecular convergence between these species is found in non-coding, regulatory regions of the genome. These analyses identified a set of 339 thylacine-wolf accelerated regions (TWARs), putative craniofacial cis-regulatory elements that have undergone accelerated evolution in both thylacine and wolf lineages.

Using a Massively Parallel Reporter Assay, we functionally characterize the activity of orthologous TWAR sequences from 7 different speciesâ€”the species of interest (thylacine and wolf), ancestral reconstruction from their clades (Dasyuromorphia and Carnivora), two related outgroup species with different craniofacial phenotypes (Tasmanian devil and panda) and a control species (mouse). We test their activity in two mouse cell populations involved in craniofacial development. By comparing changes in regulatory activity of orthologous elements, we assess which regions showing a convergent shift in molecular activity also underlie convergent phenotypes. Dense 10bp tiling of our sequence library allows the resolution to detect specific transcription factor motif changes driving the observed phenotypic changes.

291 - Genomic architecture in social insects is more strongly associated with phylogeny than social behavior

Sara E Miller

Sara E Miller, Garrett Behrends, Thomas Hagan, Samir Kuinkel
University of Missouri - St. Louis, University of Missouri - St. Louis, University of Missouri - St. Louis, University of Missouri - St. Louis

The evolution of sociality in insects has been predicted to reduce effective population sizes, in turn leading to changes in genome architecture, including higher recombination rates, increased GC-biased gene conversion (gBGC) and greater intragenomic variation in GC content to maintain castes through differential methylation. As the number of sequenced insect genomes continues to grow, it remains

an open question which of these features of genome are common across all social insect genomes. A major challenge to determining such commonalities has been the lack of phylogenetically controlled analyses. Here we analyze genome assemblies for 352 species of Hymenoptera and seven species of Blattodea to test if GC content, genome size, distribution of CpG sites or codon bias repeatedly differed between social and non-social species. Overall, we found little support for predictable changes in genome architecture associated with sociality. Social species have slight differences in GC content, genome size and CpG site distributions compared to non-social species, however these differences disappeared after accounting for phylogenetic relationships. Similarly, Hymenoptera genomes had an overall AT-bias at four-fold degenerate nucleotides that were similar for social and non-social species. We performed phylogenetically controlled analyses within two well represented families, Apidae and Halictidae, but only found a significant negative relationship between sociality and GC content within Apidae. In all, these results suggest that unique origins of social behavior may produce unique trends in genomic architecture. Our study highlights the need to examine genome architecture across independent origins of social behavior.

556 - Firefly toxin lucibufagins evolved after the origin of bioluminescence

Ying Zhen

Ying Zhen
Westlake University

Fireflies were believed to originally evolve their bioluminescence as warning signals to advertise their toxicity to predators, which was later adopted in adult mating. Although evolution of bioluminescence has been investigated extensively, the warning signal hypothesis of its origin has not been tested. We test this hypothesis by systematically determining the presence or absence of firefly toxin Lucibufagins (LBGs) across fireflie species and inferring the time of origin of LBGs. We confirmed the presence of LBGs in subfamily Lampyrinae, but more importantly, we revealed the absence of LBGs in other lineages including subfamilies of Luciolinae, Ototretinae and Psilocladinae, two incertae sedis lineages, and Rhagophthalmidae family. Ancestral state reconstructions for LBGs based on firefly phylogeny constructed using genomic data suggested that the presence of LBGs in the common ancestor of Lampyrinae subfamily was highly supported, but unsupported in more ancient nodes including firefly common ancestors. Our results suggest that firefly LBGs probably evolved much later than the evolution of bioluminescence. We thus conclude that firefly bioluminescence did not originally evolve as direct warning signals to toxic LBGs and advise that future studies should focus on other hypotheses. Moreover, LBG toxins are known to directly target and inhibit the Î± subunit of Na+,K+-ATPase (ATPÎ±). We further examined effects of amino acid substitutions in firefly ATPÎ± on its interactions with LBGs. We found that ATPÎ± in LBGs-containing fireflies are relatively insensitive to LBGs, which suggested that target site insensitivity contributes to LBGs-containing fireflies' ability to deal with their own toxins.

620 - Seven new reference-quality genomes illuminate the genomic evolution of genetic sex determination in turtles

Leon Hilgers

Leon Hilgers, Michail Rovatsos, Thomas Brown, Martin Pippel, Chetan Munegowda, Peter Praschag, Sylke Winkler, Joana Damas, Harris Lewin, Eugene Myers, Uwe Fritz, Michael Hiller
LOEWE Centre for Translational Biodiversity Genomics (LOEWE-TBG), Senckenberg Research Institute, Goethe University, Charles University , Leibniz Institute for Zoo and Wildlife Research, Uppsala University, Max Planck Institute of Molecular Cell Biology and Genetics, LOEWE Centre for Translational Biodiversity Genomics (LOEWE-TBG), Turtle Island, Max Planck Institute of Molecular Cell Biology and Genetics, University of California, University of California, Max Planck Institute of Molecular Cell Biology and Genetics,

Senckenberg Research Institute, LOEWE Centre for Translational Biodiversity Genomics (LOEWE-TBG), Senckenberg Research Institute, Goethe University

Turtles are ideal to study sex chromosome evolution. They repeatedly evolved genetic sex determination (GSD), with male (XY) and female (ZW) heterogamety and micro- as well as macro-sex chromosomes. Additionally, karyotype-based studies indicate that GSD evolution drastically changed genome architecture increasing chromosome numbers in GSD turtles. However, turtle sex chromosome origins, their gene contents and local as well as genome-wide consequences of GSD evolution remain unknown.
Here, we generated seven reference-quality turtle genomes, re-sequenced males and females of GSD turtles, and use a haplotype-resolved assembly including X and Y chromosomes to locate an elusive sex locus. Next, we identified gene losses as well as genes under positive selection and reconstruct chromosome evolution across all turtles. This allows us to explore genome evolution linked to GSD and trace the evolution of micro- and macro-sex chromosomes. In contrast to previous hypotheses, GSD does not coevolve with chromosome numbers in turtles. Instead, correlation of GSD and chromosome numbers is driven by dynamic chromosome evolution with numerous fusions in a sister clade with temperature dependent sex determination. Our data show that chelid turtle sex chromosomes originated once from a micro-chromosome >80 mya, but remained highly homomorphic. Macro-sex chromosomes evolved by subsequent chromosome fusion. In combination, this study illuminates sex chromosome evolution in turtles, adds to mounting evidence that not all sex chromosomes are â€œborn to dieâ€• and paves the way to establish turtles as a model to study sex chromosome evolution.

767 - A new scenario for the macroevolution of the Caulimoviridae based on the analysis of endogenous caulimovirids using CAULIFINDER

Héléna Vassilieff

HÃ©lÃ©na Vassilieff, Nathalie Choisne, Florian Maumus, Pierre-Yves Teycheney, Saad Serfraz, Andrew D.W Geering
University Paris-Saclay, INRAE, URGI, University Paris-Saclay, INRAE, URGI, University Paris-Saclay, INRAE, URGI, CIRAD, UMR PVBMT, CABB, University of Agriculture Faisalabad, Queensland Alliance for Agriculture and Food Innovation, The University of Queensland

Endogenous viral elements (EVEs) result from the integration of viral sequences into the genome of their hosts. EVEs can be considered molecular fossils and provide access to ancient or unknown viral sequences that can be used in paleovirological approaches to reconstruct the evolution of related viruses over extended periods of up to tens of millions of years. In plants, most characterized EVEs belong to the family Caulimoviridae, the only family of retrotranscribed plant viruses, and are termed endogenous caulimovirids (ECVs). Previous studies of ECVs have proposed two alternative scenarios to explain the evolutionary history of the Caulimoviridae: a coevolution scenario and a host-switching scenario. Although different, these two scenarios assume a very early emergence of the Caulimoviridae, which may well have preceded that of the euphyllophytes. However, the bioinformatic analyses on which these studies were based were limited by the lack of automated tools for annotating ECVs in plant genomes and the underrepresentation of basal plant genomes in their datasets. To address these issues we developed CAULIFINDER, a tool for the automated detection and annotation of ECVs in plant genomes, and used it to analyze a dataset of 56 genome assemblies covering all subdivisions of the tracheophytes, including basal ones such as ferns and lycophytes. We discovered 1449 representative ECVs that were organized in 35 new taxonomic units among the Caulimoviridae. By intersecting their phylogenetic relationships with the taxonomy of host plants, we propose a new macroevolutionary scenario for the family Caulimoviridae, reconciling host-virus coevolution and host swaps.

771 - Phylogenetic modeling of gene expression shifts in the mole-rat clade

Maëlle Daunesse

MaÃ«lle Daunesse, Ã‰lise Parey, Diego Villar, Camille Berthelot

EMBL-EBI, Institut Pasteur, University College London, Queen Mary University of London, Institut Pasteur, IBENS

Understanding the genetic basis of phenotypic adaptations poses a significant challenge in evolutionary genomics. Despite the morphological and physiological diversity in mammalian traits, their coding genomes exhibit a high degree of conservation, implying that changes in gene expression and regulation are pivotal in driving phenotype evolution. This study aims to identify and establish a connection between shifts in gene expression and cis-regulatory elements, and their potential impact on phenotypic adaptations. Using African mole-rats as a model, renowned for their unique phenotypic adaptation traits like cancer resistance and hypoxia tolerance, we aimed to elucidate the genome-wide gene expression patterns underlying these traits that have been mainly characterised at the level of candidate genes and in individual species. Profiling gene expression in heart and liver tissues across two mole-rat species and two rodent outgroups, we used a phylogenetic comparative approach to identify genes with expression shifts within the mole-rat clade and in specific genera. These shifted genes are associated with functions pertinent to known adaptations in naked mole-rats, such as myogenesis and glycolysis in the heart. Furthermore, our analysis revealed concordant changes in the regulatory landscape of these genes. By employing a phylogenetic comparative approach, we offer new insights into the interplay between gene expression, regulation, and phenotypic evolution in mammals. Our findings shed light on the molecular mechanisms driving the evolution of unique traits in mole-rats and potentially other mammalian species.

831 - Life at extreme elevations: genomic and physiological insights into mechanisms of hypoxia adaptation in Andean mice

Jay F Storz

Jay F Storz, Schuyler Liphardt, Naim M Bautista, Nathanael D Herrera, Marcial Quiroga-Carmona, Juan C Opazo, Timothy B Wheeler, Zachary A Cheviron, Guillermo D'ElÃ-a, Jeffrey M Good
University of Nebraska, University of Montana, University of Nebraska, University of Montana, University of Nebraska, Universidad Austral de Chile, Universidad de San SebastiÃ¡n, Integrative Biology Group, Millennium Nucleus of Ion Channels-Associated Diseases (MiNICAD), University of Montana, University of Montana, Universidad Austral de Chile, University of Montana

We report results of integrated genomic and physiological analyses to identify mechanisms of hypoxia adaptation in Andean leaf-eared mice (Phyllotis vaccarum) that live at extreme elevations (>6700 m [>22,000â€™]) that were previously considered to be completely uninhabitable by mammals. Our high-altitude surveys in the Central Andes have yielded numerous specimens of P. vaccarum from extraordinarily high elevations, including the summits of multiple >6000 m (>20,000â€™) volcanoes, far surpassing all previous specimen-based records for mammals. Population genomic analysis of specimens collected from sites spanning >6700 m of elevation â€“ from the Pacific Coast of northern Chile to the crest of the Andean Cordillera â€“ revealed relatively low levels of genome-wide differentiation. The analysis also revealed evidence for altitude-related selection on specific genes and pathways that we then targeted for functional experiments on captive mice reared in common-garden conditions. Our results provide insights into the genetic and physiological mechanisms that enable P. vaccarum to survive at the environmental limits of vertebrate life.

865 - Population genetics of the Pacific acorn barnacle and insights into evolution in large populations

Angel Gabriel Rivera-Colón

Angel G Rivera-ColÃ³n, Scott T Small, Jiseon Min, Alexandra Bangs, John P Wares, Peter Ralph, Andrew D Kern
Institute of Ecology and Evolution, University of Oregon, Institute of Ecology and Evolution, University of Oregon, Institute of Ecology and Evolution, University of Oregon, Institute of Ecology and Evolution, University of Oregon, Department of Biology, Odum School of Ecology,

University of Georgia, Institute of Ecology and Evolution, University of Oregon, Department of Mathematics, University of Oregon, Institute of Ecology and Evolution, University of Oregon

The strength and effects of stochastic and deterministic evolutionary forces are mediated by the size of the populations in which they act. To date, the most detailed efforts at describing population genomic variation have focused on species with small to moderately sized historical populations (e.g., human, Drosophila), with fewer characterizations at extremely large population sizes. Here, we use the Pacific acorn barnacle (Balanus glandula) as a model for the study of evolution across both large population sizes and geographic scales. B. glandula is widespread across intertidal ecosystems in the Pacific coast of North America, from Baja California to Alaska. This species lives at densities between thousands and tens of thousands of individuals per square meter, and have pelagic larvae capable of long range dispersal. We provide the first chromosome-level genome assembly for B. glandula, which we combine with large-scale spatial sampling, to describe the landscape of genome-wide genetic diversity across the speciesâ€™ range. Additionally, we leverage spatial population genetic simulations to model the specific life history of barnacles to more accurately assess the combined effects of evolutionary forces in the observed patterns of genetic variation. By pairing empirical population genomics and spatial genetic simulations we hope to disentangle the dynamics of evolutionary processes in large barnacle populations and set the stage for similar studies in other large-population systems.

907 - Exploring the role of tRNA regulation in the transition to multicellularity of the amoeba Dictyostelium discoideum.

Dulce I. Valdivia Martínez

Dulce I. Valdivia MartÃ-nez, Peter F. Stadler

University of Leipzig, University of Leipzig, University of Vienna , The Santa Fe Institute, National University of Colombia

In eukaryotes, transcription and translation occur in distinct cellular compartments. This division presents an intricate adaptation of regulation in which transfer RNAs (tRNAs) have to maintain the link between these two processes, as they must be transcribed within the nucleus and then exported to the cytoplasm to carry the amino acids needed to satisfy the ribosomal demand for proper messenger RNA-to-protein translation. Interestingly, this suggests that tRNA abundance could be an alternative way to regulate protein synthesis and their downstream associated phenotypes that it is independent of the regulation of their coding genes. Here, we explore the mechanisms that control changes in tRNA pools and their translational consequences during the development of unicellular Dictyostelium discoideum amoebae into multicellular fruiting bodies, a response that allows them to survive periods of starvation. Studying tRNAs at a transcript level is challenging because they are encoded in large gene families (e.g. ~400 genes in the 34Mb nuclear genome of D. discoideum) with many identical copies that produce short transcripts (~75 bp) and, additionally, premature tRNA molecules are cleaved into smaller fragments with potential alternative functions. To overcome this, we employed a combination of epigenomic datasets (ATAC-seq and ChIP-seq) to identify transcriptionally active tRNA gene loci and, in addition, we used LOTTE-seq data, a newly developed method to capture mature tRNAs. Our findings shed light on the fine-tuned contribution of tRNA abundance to adaptive phenotypes, either solely to meet amino acid demands or, potentially, as a crucial regulatory factor.

1111 - Size matters: revealing the genes controlling shot apical meristem gigantism in cacti

Ulises Rosas

Brenda A. LÃ³pez RuÃ-z, Michael D. Purugganan, Ulises Rosas

JardÃ-n BotÃ¡nico, Instituto de BiologÃ-a, UNAM, Center for Genomics and Systems Biology, New York University, JardÃ-n BotÃ¡nico, Instituto de BiologÃ-a, UNAM

The Cactaceae family has larger shoot apical meristems (SAMs) than other seed plants, such as Arabidopsis thaliana. Mechanisms involved in regulating SAM size in cacti may be similar to those found in model plants. Yet, larger SAM may be achieved by reconfiguring existing genetic elements, leading to novelties found in cacti. The Mammillaria genus is a good model system for studying SAM development due to its short life cycle, high diversity and ease of reproduction. We identified five developmental stages of M. san-angelensis plantlets and compared SAM size with adult plants using histological and microscopic electronic analysis. We also de novo sequenced the whole genome of M. san-angelensis using the PacBio long-read sequencing platform and identified two putative gene homologues that control SAM size: WUSCHEL (WUS) and SHOOT MERISTEMLESS (STM). The MsWUS gene consists of 2409 nucleotides with three exons, two introns, and a putative protein of 273 amino acids. We found a high level of conservation at the nucleotide and amino acid level of the homeodomain, WUS-box, and EAR-like domains compared to A. thaliana. However, the acidic region showed less conservation as compared to the A. thaliana sequence but was highly conserved among several members of the Caryophyllales. We discuss that the evolution of the SAM regulatory network is responsible for some of the novelties seen in cacti.

1125 - The genetic and structural basis of a novel iridescence phenotype in the platyfish, Xiphophorus variatus

Nadia B Haghani

Nadia B Haghani, Theresa R Gunn, Tristram O Dodge, Zihao Ou, Gabriel A Preising, Sophia Haase Cox, John J Baczenas, Cheyenne Payne, Daniel L Powell, Molly Schumer

Stanford University, Stanford University, Stanford University, Stanford University, Stanford University, Stanford University, Stanford University, Stanford University, UC Santa Cruz, Stanford University, Stanford University

A basic goal of biology is to understand the molecular changes that underlie diverse phenotypes. The variable platyfish, Xiphophorus variatus, demonstrates a diverse array of heritable skin pigmentation patterns and has become a tractable system to study the genetic changes leading to adaptive traits. We serendipitously discovered a population of X. variatus that displays a unique iridescence trait in their scales. We confirmed this iridescence is attributed to an enrichment of reflective cells containing subcellular nanostructures. To identify loci associated with iridescence, we performed a genome wide association study (GWAS) and detected a single peak associated with the iridescence trait that spans a surprisingly wide, ~3Mb region. It contains a promising gene candidate, anaplastic lymphoma kinase ligand 2 (alkal2), required for iridophore differentiation in zebrafish. We measured transcriptional profiles of scale tissue and found higher expression levels of alkal2 in iridescent individuals. To visualize structural changes that might explain this transcriptional difference, we used HiFi PacBio reads to generate whole-genome assemblies for iridescent and non-iridescent X. variatus and aligned them at the GWAS peak region. We discovered an 18kb insertion near alkal2 in the iridescent individual that blasts to an LTR/ERV-Foamy element. Given LTR retrotransposons can act as enhancers for nearby genes, we hypothesize that this insertion may be a recently active TE that enhances alkal2 expression, suggesting a possible mechanism by which structural variation produces this novel iridescence trait.

348 - Steppe Ancestry in western Eurasia and the spread of the Germanic Languages

Hugh McColl

Hugh McColl, Guus Kroonen, J. VÃ-ctor Moreno-Mayar, Frederik Valeur Seersholm, Gabriele Scorrano, Thomaz Pinotti, Tharsika Vimala, SÃ¸ren M SindbÃ¦k, Per Ethelberg, Ralph Fyfe, Marie-JosÃ© Gaillard, Hanne M. EllegÃ¥rd Larsen, Morten Fischer Mortensen, Fabrice Demeter, Marie Louise S. JÃ¸rkov, Sophie Bergerbrant, Line Olsen, Rasmus Nielsen, Lotte Hedeager, Kristian Kristiansen, Martin Sikora, Eske Willerslev

Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, Leiden University Centre for Linguistics, Leiden

University, Department of Nordic Studies and Linguistics, University of Copenhagen, Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, LaboratÃ³rio de Biodiversidade e EvoluÃ§Ã£o Molecular (LBEM), Universidade Federal de Minas Gerais, Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, Centre for Urban Network Evolutions (UrbNet), School of Culture and Society, Aarhus University, Museum SÃ¸nderjylland, ArkÃ¦ologi, School of Geography, Earth and Environmental Sciences, University of Plymouth , Department of Biology and Environmental Sciences, Linnaeus University, National Museum of Denmark. Environmental Archaeology and Materials Science, National Museum of Denmark. Environmental Archaeology and Materials Science, Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, Eco-anthropologie (EA), Dpt ABBA, MusÃ©um national dâ€™Histoire naturelle, CNRS, UniversitÃ© Paris CitÃ©, MusÃ©e de lâ€™Homme 17 place du TrocadÃ©ro, 75016, Laboratory of Biological Anthropology, Globe Institute, University of Copenhagen, Department of Archaeology and Ancient History, Campus Gotland, Uppsala Univeristy, Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, Departments of Integrative Biology and Statistics, UC Berkeley, Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, Department of Historical Studies, University of Gothenburg, Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen, GeoGenetics Group, Department of Zoology, University of Cambridge

According to linguistic consensus, the common ancestor of the Germanic languages, including German, English, Frisian, Dutch and the Nordic languages, arrived with the migrations of Bronze Age Steppe Pastoralists into Europe around 5,000 years ago. In much of Europe, there has also been an assumption of genetic continuity since this last major migration, with limited genetic impacts from the surrounding regions. However, the ability to detect later migrations within Europe is hindered by the close relation between populations and limited ancient samples. We sequenced 710 ancient human genomes and analysed them together with 3,940 published genomes suitable for imputing diploid genotypes. By analysing recent coancestry through shared IBD segments we find a series of new clines associated with migrations within Europe. From the Bronze Age onwards, we find two broad clusters within northern and western Europe corresponding with Corded Ware and Bell Beaker cultures. Within the former, we find evidence of a cross-Baltic migration into Eastern Scandinavia that later spread west and south, arriving 800 years later but consistent with the assumed paleo-Germanic range. Later, during the migration period, we identify a region around Southern Jutland / Northern Germany as the source for migrations west to the British Isles and north into the Danish Isles and Southern Sweden, corresponding with the spread of West and North Germanic languages respectively. This work demonstrates the feasibility of detecting migrations within closely-related populations, and reveals a series of archaeologically invisible migrations of archaeological, historic and linguistic significance.

469 - Reconstructing mutation patterns over the course of human evolution

Yulin Zhang

Yulin Zhang, Laurits Skov, Jules Perez, Priya Moorjani
UC Berkeley, UC Berkeley, University of Copenhagen, UC Berkeley, UC Berkeley

One of the most fundamental discoveries in evolutionary biology is the molecular clockâ€“ the notion that mutations occur steadily over time and thus could serve as an "evolutionary clock" for dating past events. Recent whole genome sequencing studies have challenged the validity of the molecular clock by revealing an almost two-fold difference in mutation rate over the course of human evolution. This discordance has been hypothesized to stem from a decrease in mutation rates towards the present in humans. To investigate this

hypothesis, we leverage the history of archaic gene flow into modern humans to learn about mutation patterns at deep evolutionary timescales. Using whole genome sequences of ~900 individuals from the Human Genome Diversity Panel, we identify Neanderthal and Denisovan archaic segments in present-day individuals from diverse populations. We compare the number and types of derived mutations on the archaic and non-archaic syntenic fragments. We find similar overall mutation rates across lineages over the past 500,000 years. Comparison of mutation spectrumâ€"â€“ i.e., composition of different types of mutationsâ€"â€“reveals significant differences between archaic and modern human segments. We discuss the implications and mechanisms related to these changes. Future studies that use older calibration points can help to recover the mutation rate at deeper timescales and facilitate the recalibration of the molecular clock.

472 - Tracing admixture in European early farmers using local ancestry inference

Sandra Oliveira

Sandra Oliveira, Nina Marchi, Adamandia Kapopoulou, Laurent Excoffier
Institute for Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics, Institute for Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics, UMR7206 Eco-anthropologie, CNRS-MNHN-UniversitÃ© Paris CitÃ©, Institute for Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics, Institute for Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics

In Europe, the transition from a hunting and gathering lifestyle to food-production and sedentism unfolded with the expansion of people and culture from Anatolia. The ancestry of the first European farmers, who spread along the Mediterranean or Danubian routes, can be traced back to the Aegean basin, but recent demographic reconstructions show that these early farmers already carried substantial amounts of Western hunter-gatherer (WHG) ancestry. Genomic evidence thus suggests connectivity between Western Europe and Southwest Asia, and potentially early admixture in Anatolia or the Near East. In addition, Neolithic farmers further admixed with local WHG-related groups when expanding to Europe. However, the frequency, timing, and outcomes of these admixture events are still unclear. Here, we use local ancestry inference to investigate these complex admixture processes and their genomic signatures in early Neolithic farmers spreading from Anatolia up to the British Isles, whose genomes were sequenced to high depth. Using different ancient individuals as references for early farmer and WHG admixture sources, we identified and compared the distribution, overlap, and similarity of WHG ancestry tracts across early farmer genomes. Our work provides insights into ancient individual relationships, their admixture history, as well as the role of admixture in shaping our genomes during range expansions.

599 - A high-quality genome from a 200,000-year-old Denisovan

Stéphane Peyrégne

StÃ©phane PeyrÃ©gne, Diyendo Massilani, Yaniv Swiel, Alba Bossoms Mesa, Arev Pelin SÃ¼mer, Michael James Boyle, Cesare de Filippo, Bence Viola, Elena Essel, Sarah Nagel, Julia Richter, Antje Weihmann, Barbara Schellbach, Hugo Zeberg, Johann Visagie, Maxim B Kozlikin, Michael V Shunkov, Anatoly P Derevianko, Benjamin M Peter, Matthias Meyer, Svante PÃ¤Ã¤bo, Janet Kelso
Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Department of Anthropology, University of Toronto, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Department of Physiology and Pharmacology, Karolinska Institutet, Max Planck Institute for Evolutionary Anthropology, Institute of Archaeology and

Ethnography, Siberian Branch, Russian Academy of Sciences, Institute of Archaeology and Ethnography, Siberian Branch, Russian Academy of Sciences, Institute of Archaeology and Ethnography, Siberian Branch, Russian Academy of Sciences, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology

It is a little over a decade since the discovery of Denisovans, a human group that split from a common ancestor with Neandertals around 400 thousand years ago (ka). Nearly all we know about Denisovans is inferred from the single high-quality genome sequence of the Denisova 3 individual, who lived approximately 60ka. However, the presence of multiple Denisovan ancestry components in modern human genomes suggests that there were multiple genetically diverse Denisovan populations. Reconstructing more Denisovan genomes is therefore essential for understanding the population history underlying this genetic diversity.
We describe a new high-coverage Denisovan genome reconstructed from DNA extracted from Denisova 25, a molar found at Denisova Cave. We estimated that this male individual lived around 200ka, in agreement with dating of the corresponding sediment layer. By comparing the genomes of Denisova 25 and Denisova 3, we determined that he belonged to an early Denisovan population that split about 220ka from later Denisovans at the cave, highlighting a population turnover. Approximately 12% of his genome is in long runs of homozygosity, suggesting that he lived in a small population and that there was recent inbreeding among his close ancestors. We also found evidence of recurrent mixing with Neandertals among the Denisovans at Denisova Cave; Denisova 25 inherited Neandertal ancestry from contacts with yet unknown divergent Neandertal populations. Furthermore, using both Denisovan genomes, we dissected Denisovan ancestry components in present-day human genomes. Denisova 25 provides a new view on Denisovan history and their interactions with other hominins.

628 - Overcoming bias and postmortem damage to improve the accuracy of ancient genome analysis

Dilek Koptekin

Dilek Koptekin, Etka Yapar, KÄ±vÄ±lcÄ±m BaÅŸak Vural, Ekin SaÄŸlÄ±can, Anna-Sapfo Malaspinas, Can Alkan, Mehmet Somel
Department of Computational Biology, University of Lausanne, Department of Biological Sciences, Middle East Technical University, Department of Biological Sciences, Middle East Technical University, Department of Biology, Lund University, Department of Biological Sciences, Middle East Technical University, Department of Health Informatics, Graduate School of Informatics, Middle East Technical University, Department of Computational Biology, University of Lausanne, Swiss Institute of Bioinformatics, University of Lausanne, Department of Computer Engineering, Bilkent University, Department of Biological Sciences, Middle East Technical University

Ancient DNA (aDNA) contains invaluable information about population history and evolution, but its analysis is plagued by challenges such as reference bias and postmortem damage (PMD). Here, we propose new strategies to address these challenges. First, we show that masking variable genomic positions before alignment or using a graph genome both effectively reduce reference bias in ancient human genome data. However, since quality-filtered BAM files and 1240K capture data retain bias, this solution depends on data availability in FASTQ or unfiltered BAM formats. Second, we evaluate various strategies for removing PMD, including trimming, rescaling base quality, and a new algorithm, bamRefine, that masks polymorphic regions vulnerable to PMD. Compared to trimming and rescaling, bamRefine significantly improves genotyped loci by up to 20% while increasing accuracy. Overall, we advocate for the adoption of graph alignment in conjunction with bamRefine to minimize data loss and bias in aDNA research. Furthermore, our study highlights the importance of publishing FASTQ files for comprehensive analysis. By addressing these technical challenges, our work increases the reliability and depth of information obtained from ancient DNA, which in turn will improve our inference of population history and evolution.

839 - Archaic ancestry inference in imputed ancient human genomes

Marco Rosario Capodiferro

Marco Rosario Capodiferro , LÃ©o Planche, Linda Ongaro, Emily Breslin, Flora Jay, Lara Cassidy, Emilia Huerta-Sanchez
Smurfit Institute of Genetics, Trinity College Dublin, Interdisciplinary Laboratory of Digital Sciences, UniversitÃ© Paris-Saclay, Smurfit Institute of Genetics, Trinity College Dublin,, Smurfit Institute of Genetics, Trinity College Dublin, Laboratory of Digital Sciences, UniversitÃ© Paris-Saclay, Smurfit Institute of Genetics, Trinity College Dublin, Smurfit Institute of Genetics, Trinity College Dublin, Ecology and Evolutionary Biology and Center for Computational and Molecular Biology, Brown University

A little over a decade ago, studies showed that when Anatomically Modern Humans expanded into Eurasia, they interbred with Archaic Humans (Neanderthal and Denisovan). This discovery redefined the origin of humans and opened new research directions to study the impact of archaic introgression in human evolution. Up to now, the majority of archaic introgression studies have focused on contemporary individual genomic data, revealing that non-African populations harbour varying levels of Neanderthal and Denisovan ancestry. Despite recent advances in identifying and quantifying archaic introgression in humans, little is known about the evolution of archaic variants within modern human populations after the introgression event (40,000 years ago). Archaic variation continued to evolve within humans and was likely shaped by a populationâ€™s unique demographic history as well as natural selection. While ancient DNA offers the potential to address this scientific gap, its poor quality has hindered the exploration of ancient genomes.
Here, we investigate the feasibility of using imputation to enhance global and local archaic ancestry inference in ancient genomes. We downsampled 20 high-coverage genomes, representing individuals from diverse temporal and geographical contexts, to 0.0625X, 0.5X, 1X and 2X. Following the imputation of downsampled genomes, we compared them with high-coverage genomes to assess global and local archaic ancestry inference, using also archaic reference genomes to ascertain the inferred introgressed segments origin. We unveil the significant capacity to infer archaic ancestry in imputed ancient genomes. Notably, local ancestry inference in imputed genomes outperforms that of the original high-coverage genomes.

270 - The UV sex chromosomes of brown algae act as genomic cradles for new genes that evolve de novo

Josué Barrera-Redondo

JosuÃ© Barrera-Redondo, Jia Xuan Leong, Christos Papadopoulos, CÃ¡tia Igreja, RÃ©my Luthringer, Fabian B. Haas, M. Mar AlbÃ , Susana M. Coelho
Max Planck Institute for Biology, Max Planck Institute for Biology, Hospital del Mar Medical Research Institute, Max Planck Institute for Biology, Max Planck Institute for Biology, Max Planck Institute for Biology, Hospital del Mar Medical Research Institute, Max Planck Institute for Biology

Despite the vast diversity of sex-determination systems in nature, there are only three known types of sex-chromosome systems: the XY, the ZW, and the largely unexplored UV systems. Although these three systems share several core features, there are important differences between them that have broad evolutionary and genomic implications. Brown algae (Phaeophyceae) are complex multicellular eukaryotes that harbor an ancestral UV sex-determination system, but have also evolved a derived XY system, as well as several transitions to hermaphroditism. Thus, brown algae represent exceptional models for investigating the evolutionary dynamics of sex chromosomes. We show that the UV sex chromosomes in brown algae are at least 450 million years old. When we calculate the relative age of each gene in the brown algal genomes, we observe that the UV sex chromosomes are consistently enriched in younger genes compared to the autosomes, particularly in the pseudoautosomal regions. This pattern is accompanied by higher substitution rates, more repressive chromatin marks, and higher TE content. However, the brown algae that evolved an XY system or transitioned to a

hermaphroditic lifestyle lost this pattern, suggesting that young gene enrichment is a feature of the UV system. Through a combination of synteny analyses, RNAseq, and riboseq data, we tested whether new genes are arising in the UV sex chromosomes through de novo gene birth processes (i.e., the birth of new protein-coding genes from noncoding DNA). We propose that the UV sex chromosomes act as genomic 'cradles', fostering the birth of new genes.

424 - Cephalopod sex determination and its ancient evolutionary origin revealed by chromosome-level assembly of the California two-spot octopus

Gabrielle Coffing

Gabrielle Coffing, Silas Tittes, Scott Small, Jeremea Songco-Casey, Judit Pungor, Denise Piscopo, Adam Miller, Cris Niell, Andrew Kern

University of Oregon, University of Oregon, University of Oregon, University of Oregon, University of Oregon, University of Oregon, University of Oregon, University of Oregon, University of Oregon

The coleoid cephalopods comprise a remarkable branch of the tree of life, possessing a distinct set of novel traits. With the recent proliferation of genome sequencing technologies, we have the capability to explore the genomic foundations of these distinctive characteristics. In 2015, the California two-spot octopus (Octopus bimaculoides) became the first cephalopod to be sequenced, and since then, numerous additional reference genomes for various members of this group have been generated. Here, using PacBio long-read sequencing of genomic DNA and Iso-seq sequencing of full-length mRNA transcripts, we provide a novel chromosome-scale reference genome and annotation for O. bimaculoides. We use our assembly in combination with existing cephalopod genomes to create the first whole genome alignments from this group and characterize the landscape of sequence conservation in octopus. Notably, our assembly reveals evidence of a heterogametic chromosome in females, consistent with a ZZ/ZO male/female systemâ€”the first evidence of genetic sex determination in cephalopods. We compared this heterogametic chromosome to other cephalopods and our data support a single origin of this sex chromosome on the lineage leading to coleoid cephalopods. The Z chromosome is an evolutionary outlier with respect to divergence and repetitive element patterns. Our results suggest that the coleoid cephalopod Z chromosome originated between 455 and 248 million years ago, placing it among the oldest conserved animal sex chromosomes known.

528 - TBD - Sent to session organizers

Melissa A Wilson

Melissa A Wilson

Arizona State University

TBD - Sent to session organizers

600 - Do Z chromosomes in Butterflies and Moths Show Evidence of Increased Rates of Adaptation?

Claudia C Weber

Claudia C Weber

Tree of Life, Wellcome Sanger Institute

Chromosomes present in a single copy in one sex often have unusual patterns of sequence evolution, since the effects of recessive mutations are not masked. Therefore, substitution rates may be elevated compared to the autosomes. This observation has prompted speculation that tortricid moths, which are agricultural pests, are evolutionarily successful due to enhanced adaptation of newly Z-linked

detoxifying enzymes (without directly examining the implicated proteins). However, both â€œfastâ€• and â€œslowâ€• Z evolution have been reported in lepidopterans, depending on the species and genes examined. Previous analyses were limited to small sets of species, and it remains unclear if positive selection on the Z chromosome is common across the taxon.

Using hundreds of lepidopteran genomes from the Darwin Tree of Life project, we can revisit the evolution of the Z chromosome and ask: Is there evidence for accelerated evolution of Z-linked sequences across lepidopterans? Does the pattern differ for recently translocated sequences? Drawing on a catalogue of genomic rearrangements, we can also consider if newly Z-linked genes evolve faster than comparable sequences that remain autosomal.

The dataset also provides an opportunity to examine if any observed acceleration is consistent with adaptation rather than drift. In addition to identifying fast-evolving proteins, pinpointing rapidly diverging sites and examining the nature of the changes will be informative. Past work on mammals reports that positively selected sites in xenobiotic-metabolising enzymes cluster in functionally important regions. Considering the structural contexts of substitutions might, therefore, provide clues about their functional consequences.

633 - Using phased genome assemblies to examine the evolution of heteromorphic sex chromosomes in Cannabaceae

Sarah Carey

Sarah Carey
HudsonAlpha Institute for Biotechnology

The Cannabaceae family has a deep history regarding their dioecious flowers and sex chromosomes. One century ago, the XY pair that controls the development of the sexes was identified in the hop Humulus lupulus var. lupulus, owing to the cytologically smaller Y chromosome relative to the X. Curiously, examinations across other species in Humulus and in the sister genera Cannabis uncovered varying cytological differences. In Cannabis, the Y chromosome is bigger than the X, while other varieties of Humulus remain homomorphic. Despite these early discoveries, we know little about the Cannabaceae sex chromosomes at the molecular level. This is largely due to the complexities of assembling XY pairs in genome references. Here we use a combination of Illumina, PacBio HiFi, and Dovetail Omni-C data to assemble fully-phased genomes for XY males of Cannabis and Humulus. We show that our assemblies match the known cytological differences in these XY pairs. We next use these assemblies to explore the timing of evolution, gene gain and loss, and structural complexity of these XY chromosomes to ultimately uncover genes that control the development of the economically valuable females of hemp and hop.

637 - Shining a light on early stages of neo-sex chromosome evolution in Australian honeyeaters (Aves: Meliphagidae)

Sophia Catherine MacRae Orzechowski

Sophia Catherine MacRae Orzechowski, Doaa Doudin, Alex Drew, C. Jonathan Schmitt, Foyez Shams, Chris Wilson, Leo Joseph, Tariq Ezaz, Scott V Edwards
Harvard University, University of Canberra, Australian National Wildlife Collection, Harvard University, University of Canberra, Australian National Wildlife Collection, Australian National Wildlife Collection, University of Canberra, Harvard University

Mounting discoveries of avian neo-sex chromosomes are yielding rich opportunities to understand early stages of sex chromosome

evolution in birds. Neo-sex chromosomes form when autosomes fuse to ancestral sex chromosomes, representing a type of sex chromosome turnover. We discovered and characterized the structure and phylogenetic distribution of neo-sex chromosomes in Australian honeyeaters by integrating long-read genomic, cytogenetic, and chromatin conformation data. We leveraged a phylogenetic framework to chart the progression of sex-linked recombination suppression across the newly added region. Finally, we generated transcriptomic data to investigate patterns of dosage compensation in the Blue-faced Honeyeater. Despite overall evidence for a single bout of recombination suppression 10-20 MYA involving 68% of the 55 Mb added region, we find little evidence of major gene loss or disruption of gene expression along the added W. However we find extensive proliferation of transposable elements (TEs) on the added W, consistent with the refugium hypothesis for active TEs on sex-limited chromosomes. Our chromosome-level assembly reveals major rearrangements on the neo-W, including a 20 Mb inversion and a translocation of an ancestral W chunk that we hypothesize led to recombination arrest. Overall, our findings offer insight into the progression and consequences of recombination suppression on neo-sex chromosomes. Our â€˜chromosomicsâ€™ approach provides us with fresh ammunition to get at the long-standing question of what happens when autosomes become sex-linked, and the extent to which each sex chromosome system has its own unique molecular journey.

878 - The struggle of being a sex chromosome: Why X has a lower diversity than expected in primates

Erik Fogh Sorensen

Erik F Sorensen, Kasper Munch
Aarhus University, Aarhus University

Primates exhibit a wide range of population sizes, mating systems, dispersal patterns, and demographic histories, each expected to change the relative genetic diversity of X chromosomes and autosomes (X/A ratio). Interrogating this relative diversity across 120 species from lemurs to great apes reveals that the typical X/A ratio is closer to 1/2 than the Â¾ predicted in neutral randomly mating populations of constant size, with a mean of 0.62 across all species, with some exhibiting a ratio lower than 0.5.

We remove the species-specific effects of demography and sex-biased admixture by comparing observed X/A ratios to the values expected from their effective population size trajectories (SMC++). Still, the X/A ratios remain depressed suggesting that linked selection on the X chromosome, in the form of selective sweeps, may play a stronger role than on autosomes. Such sweeps would affect species individually, producing larger differences in heterozygosity between species across genomic regions. Consistent with this prediction, we find that the coefficient of variation on pi is not only larger on X but that its relative size correlates with the X/A ratio. Lastly, observed mating patterns and testis size have a significant effect on the X/A ratio, showing that differences in reproductive success across genders strongly affect the relative diversity on chromosome X.

110 - The evolutionary genomics of population's responses to climate change

Jonas Andres Aguirre Liguori

JonÃ¡s AndrÃ©s Aguirre-Liguori, Santiago RamÃ-rez-Barahona, Abraham Morales Cruz, Brandon S Gaut
Universidad AutÃ³noma de YucatÃ¡n, Instituto de BiologÃ-a, UNAM, Lawrence Berkeley National Laboratory, University of California Irvine

Climate change is becoming a serious threat to biodiversity. Populations can respond to this threat by migrating to suitable areas or adapting locally to persist in the future. Usually, the impact of climate change is evaluated using species distribution models. However, these models omit evolutionary processes. Recently, researchers have used genomic tools and landscape genomics to incorporate demographic history and adaptive genetic diversity in the predictions of how populations will adapt to climate change. However, other

evolutionary processes are still missing in the future projections of climate change. In this presentation, we first show how demographic history, local adaptation, gene flow, population dispersal and genomic load can be combined to predict the potential response of specific populations to climate change. Second, we analyze 5 wild Vitis species to exemplify how these layers of information can be used to identify crop wild relatives (CWRs) of Vitis species (grapes) that will be resistant in the future. We find five V. mustangensis accessions that are expected to respond adequately to climate change. Finally, we project which of these five V. mustangensis accessions could be adapted in the future if they were actively moved to locations where V. vinifera might be at risk if climate continues to change. In theory, these accessions, after experimental validations, could potentially be used to introduce climate adaptive traits in V. vinifera.

125 - Low Rates of Gene Duplication in an Endangered Freshwater Mussel

Rebekah L. Rogers

Rebekah L. Rogers, Gabe O'Reilly, Jeffrey T Garner, John Wares
UNC Charlotte, UNC Charlotte, Alabama Department of Conservation, University of Georgia

Nearly neutral theory predicts that evolutionary processes will differ in small populations compared to large populations, a key point of concern for endangered species. The nearly-neutral threshold, the span of neutral variation, and the adaptive potential from new mutations all differ depending on Ne. To determine how genomes respond in small populations, we have created a reference genomes for a US federally endangered IUCN Red List freshwater mussel, Elliptio spinosa and compare it to genetic variation for a common and successful relative Elliptio crassidens. We find higher rates of background duplication rates in E. spinosa consistent with proposed theories of duplicate gene accumulation according to nearly-neutral processes. Along with these changes we observe fewer cases of adaptive gene family amplification in this endangered species. However, TE content is not consistent with nearly-neutral theory. We observe substantially less recent TE proliferation in the endangered species with over 500 Mb of newly copied TEs in Elliptio crassidens. These results suggest a more complex interplay between TEs and duplicate genes than previously proposed for small populations. They further suggest that TEs and duplications require greater attention in surveys of genomic health for endangered species.

495 - Epigenome-wide scans identify DNA methylation markers for monitoring sublethal thermal stress in endangered sea turtles

Eugenie Charley Yen

Eugenie C Yen, James D Gilbert, Alice Balard, InÃªs O Afonso, Kirsten Fairweather, DÃ©bora Newlands, Artur Lopes, Sandra M Correia, Albert Taxonera, Stephen J Rossiter, JosÃ© M MartÃ-n-DurÃ¡n, Christophe Eizaguirre
Queen Mary University of London, Queen Mary University of London, Queen Mary University of London, Queen Mary University of London, Project Biodiversity, Project Biodiversity, Project Biodiversity, Instituto do Mar, Project Biodiversity, Queen Mary University of London, Queen Mary University of London, Queen Mary University of London

As global warming progresses, there is a pressing need to quantify temperature-mediated sublethal effects on individuals, since they reduce average population fitness without increasing mortality. Detecting these cryptic effects in the wild is however challenging. Conservation epigenomics emerges as a promising discipline because epigenetic variation represents a direct, molecular link between the genome and the environment. DNA methylation in particular can be thermally induced and stably maintained. Here, we focused on endangered loggerhead sea turtles (Caretta caretta) that nest in Cabo Verde (West Africa). We set up a split-clutch field experiment in an in-situ hatchery, where nests of wild loggerheads were buried at different depths to expose them to different temperatures in biologically relevant conditions. Upon emergence, we collected blood samples from hatchlings for whole genome bisulfite sequencing. We identified over 700 differentially methylated sites between incubation treatments, including on genes with neurodevelopmental and cytoskeletal

functions. These methylome differences were linked to their performance in multiple fitness tests of locomotion ability. Hatchlings that developed in the warmer treatment were weaker than their siblings from the cooler treatment, with direct correlations found between methylation levels and fitness measurements. Overall, we identified a set of biomarkers for monitoring exposure to sublethal thermal stress and showcase the applicability of this epigenomic approach for conservation. After successfully trialling nanopore sequencing in the field, our work also contributes to the movement towards precision conservation for protecting the health and resilience of endangered species.

686 - Towards genome-wide data of Arctic species obtained from snow footprints

Theo Phanu Serivichyaswat

Theo Phanu Serivichyaswat, Samantha LopÃ©z-Clinton, Chenyu Jin, Anja Mezger, Marc FriedlÃ¤nder, Tom van der Valk
Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Centre for Palaeogenetics, -, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Centre for Palaeogenetics, Department of Zoology, Stockholm University, -, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Centre for Palaeogenetics, Department of Zoology, Stockholm University, National Genomics Infrastructure, Science for Life Laboratory, Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Centre for Palaeogenetics

Environmental DNA (eDNA) analysis has emerged as a powerful and non-invasive tool for monitoring species without direct observation, but the genomic data obtained has extremely low coverage and is thus insufficient for population genetic studies. This study focuses on utilizing single-cell sorting , whole genome amplification techniques and Nanopore sequencing to assess the possibility of obtaining genome-wide data from environmental samples. Through sampling snow footprints, we aim to non-invasively obtain whole genomes of multiple arctic species. Our methodology involves sample preservation, single-cell sorting, genome amplification, next-generation sequencing and bioinformatic analysis of genome-wide data. By coupling the current techniques with population genomic analysis, we aim to gain invaluable insights into the spatial distribution, population dynamics, and habitat preferences of arctic species, aiding in their monitoring and conservation efforts. Our findings will underscore the utility of this novel approach as a tool for biodiversity and population monitoring, particularly for species in the Arctic that are challenging to study using traditional methods.

690 - The genetic architecture of an adaptive phenotype conditions vulnerability and populations evolutionary response to climate change

Genís Garcia-Erill

GenÃ-s Garcia-Erill, Francisco Campuzano-JimÃ©nez, Perrine Gauthier, Palle Villesen, Andreas H. Faber, Marc Palmada-Flores, Sylvain Santoni, John D. Thompson, Bodil K. Ehlers, Thomas Bataillon
Bioinformatics Research Center, Aarhus University, Bioinformatics Research Center, Aarhus University, CEFE, Univ Montpellier, CNRS, EPHE, IRD, Univ Paul ValÃ©ry Montpellier, Bioinformatics Research Center, Aarhus University, Department of Ecoscience, Aarhus University, Laboratori Ictiologia Genetica, Departament de Biologia, Universitat de Girona, UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, CEFE, Univ Montpellier, CNRS, EPHE, IRD, Univ Paul ValÃ©ry Montpellier, Department of Ecoscience, Aarhus University, Bioinformatics Research Center, Aarhus University

Common thyme (Thymus vulgaris) is a dominant species in the Mediterranean ecosystem that influences the composition of its associated plant communities. Thyme populations exhibit genetically determined individual chemotype variability at a very local scale. Chemotypes can be grouped by their chemical composition into phenolic and non-phenolic ecotypes, which have been shown to be

adaptive for different local habitats. Ongoing climate change is already resulting in shifts in the frequency of locally adaptive ecotypes in the Mediterranean region. Understanding the evolutionary history of this polymorphism and its maintenance is important for predicting the vulnerability to climate change of thyme populations and thyme associated species in Mediterranean ecosystems and ultimately guiding conservation action. We sequenced 900 thyme individuals from over 50 sites spanning 9 gradients of altitude and environmental conditions across Southern France. Combining genetic, environmental and phenotypic data allowed us to catalogue and link genetic and phenotypic diversity to environmental variation across both local and regional spatial scales. Using a newly assembled draft reference genome, we mapped the loci that determine chemotype in thyme populations. These include structural variants at loci encoding for terpene synthases involved in the biosynthesis of the monoterpenes that define the chemotypes. We use a combination of population and ecological genetics to understand how past demography, genetic connectivity and parallel adaptations shape thymeâ€™s functional and neutral genetic diversity. Finally, we explore with simulations how SNPs and environmental data can be used to predict the vulnerability of thyme populations and their potential to respond to climate change.

1042 - Predicting Pathogenic Variation for conservation: a deep learning approach

Paulina G. Nuñez-Valencia

Paulina G. NuÃ±ez-Valencia, Mafalda Dias, Jonathan Frazer

Center for Genomic Regulation, Center for Genomic Regulation, Center for Genomic Regulation

Biodiversity loss is reaching critical levels, with over 42,100 species currently facing extinction. Understanding the genetic composition of inbred populations is crucial for guiding conservation strategies, especially in the face of rapid environmental changes induced by climate and human activities. Yet, our understanding of genetic variation's effects remains limited. Particularly, identifying the distribution, prevalence, and impacts of pathogenic variants in non-model organisms is largely unexplored. The rapid increase in genomic data, coupled with new developments in statistical modeling, has opened the door to new modeling opportunities to identify pathogenic variants. Capitalizing on the vast repository of genomic data available, we leverage insights from sequence variation distribution across species and employ deep probabilistic models to identify and categorize pathogenic variants in endangered species. This approach yields predictions enriched by a comprehensive understanding of protein context and a notion of fitness across the vast spectrum of proteins and their evolutionary histories. Using the black-footed ferret (Mustela nigripes) as a case studyâ€”a highly inbred and endangered speciesâ€”we uncover distinct trends in pathogenicity scores among pre- and post-founder event samples. We also observe sex differential mutation burdens and an enrichment of pathways related to reproduction in predicted pathogenic variants. By doing so, we not only illuminate the genetic health of endangered species but also test the broader utility of unsupervised deep learning models in cross-species variant effect prediction. Ultimately, this research not only enhances our understanding of genetic variation's impacts but also informs conservation strategies to safeguard biodiversity.

297 - Challenges in joining molecular ecology and on-site sequencing to engage and improve research skills in young students in Amazonian Ecuador

Hugo Mauricio Ortega Andrade

H. Mauricio Ortega-Andrade, Andrea Carrera, Katherine Apunte, Jorge Celi, Miguel Quishpe, Miguel Herrera-Robledo, Christine Van der heyden, Walter Quilumbaquin

Biogeography and Spatial Ecology Research Group, Life Sciences Faculty, Universidad Regional AmazÃ³nica Ikiam, Herpetology Division, Instituto Nacional de Biodiversidad, Universidad Regional AmazÃ³nica IKIAM, Museo de ZoologÃ-a. Escuela de Ciencias BiolÃ³gicas. Facultad de Ciencias Exactas y Naturales. Pontificia Universidad CatÃ³lica del Ecuador, Water and Aquatic Resources Research Group,

School of Earth and Water Sciences, Universidad Regional AmazÃ³nica Ikiam, Population and Environment Research Group, Life Sciences Faculty, Universidad Regional AmazÃ³nica Ikiam, School of Chemical Sciences and Engineering, Yachay Tech University, HOGENT - University of Applied Sciences and Arts, Department of Biosciences and Industrial Technology, Health and Water Technology Research Centre, Molecular Biology and Biochemistry Lab. Universidad Regional AmazÃ³nica IKIAM

The development of fast biodiversity detection methods, particularly DNA-based techniques, is critical for effective environmental management strategies. Molecular techniques commonly used for diversity estimation in ecology, such as DNA/PCR assays and sequencing, often require laboratories with sophisticated and costly equipment. Field laboratories with portable devices have emerged as promising tools for rapid and accurate in-situ biomonitoring. We present our experience in the challenges of joining molecular ecology and direct Nanopore sequencing to engage and improve research skills in young students from the Amazon in Ecuador. Our method includes three key components: 1) Financial support through research project collaboration, 2) a well-planned training program on molecular ecology in the context of research projects (environmental DNA, DNA barcoding), and 3) to engage motivated students in molecular ecology. Projects were led by researchers who worked out in a training program to engage students from Ecosystem, Hydrology, and Biotechnology Engineering at Universidad Regional AmazÃ³nica Ikiam. After five years of the program, at least 22 students and 12 research theses were developed. The advantages of the use of portable devices for in-situ sequencing include their ease of transportation to the sampling site, convenient handling, and immediate results for students. However, in Ecuador, as in other tropical countries with developing economies, equipment and supplies from local suppliers can be up to two to five times higher than in the US or Europe. Pitfalls and challenges in accessing cutting-edge technologies to engage young Ecuadorian students in molecular ecology are discussed.

488 - Evaluating community perceptions and ethical considerations in genetics research in small scale northern Kenyan populations

Rebecca L Siford

Rebecca L Siford, Carla Handley, Sarah Mathew, Melissa Wilson, Jason Robert, Mercy Y Akinyi, Joseph Kamau, Anne C Stone
School of Human Evolution and Social Change, Arizona State University, Institute of Human Origins, Arizona State University, School of Human Evolution and Social Change, Arizona State University, School of Human Evolution and Social Change, Arizona State University, Institute of Human Origins, Arizona State University, School of Life Sciences, Arizona State University, School of Life Sciences, Arizona State University, One Health Centre, Kenya Institute of Primate Research, One Health Centre, Kenya Institute of Primate Research, School of Human Evolution and Social Change, Arizona State University, Institute of Human Origins, Arizona State University

As genomic research expands to be inclusive of global diversity, indigenous perspectives about genetic research and data management should be considered, and the success of dissemination efforts to participants should be evaluated. Kenya, where participation in genetic research is amongst the highest in Africa, is also one of the most understudied for participant perspectives from small scale societies. Here, we address how Western ethical practices used in genetic studies accurately align (or fail to align) with the understandings, attitudes, and perceptions of the Turkana, Borana, Rendille, and Samburu pastoral populations of northern Kenya. Genetic research by our group has been conducted in collaboration with these populations which facilitates the opportunity to explore ethical practices and concerns. We collected semi-structured interviews from previous genetic study participants and from those who had not previously participated, to explore the effectiveness of dissemination efforts and to characterize views on privacy, consent, data use, and governance of data. From over 550 interviews collected in 2023, we generally find that dissemination of genetics results is expected and should be presented with cultural and religious considerations. Also, broad consent is an inappropriate model and permission for new studies should be sought. Finally, communities favor a participatory model of research. We provide a general set of guidelines and

population specific governance frameworks that may aid Kenyan research ethics committees and other researchers. Our study highlights that participant and community perspectives offer culturally relevant motivations and are necessary to ensure ethical governance of genetic data.

748 - TE Hub: A Community for Developing Tools, Trainings, and Resources to Analyze DNA Repeats

Sarah Schaack

Sarah Schaack, TE-Hub Consortium TE-Hub Consortium
Reed College, TE-Hub

The analysis of repetitive regions (RRs) throughout the genome is a formidable challenge in bioinformatics and computational biology. Not only do RRs (including transposable elements (TEs), endogenous viruses, and repeats of as-of-yet-unknown origin) constitute a significant portion or the majority of some genomes, they can also be the most dynamic components of genome architecture due to their ability to replicate, mobilize, horizontally transfer, or facilitate non-homologous recombination. Many biologists hope to analyze RRs of the genomes they study, either to understand their impact, history, or function. Because of their ubiquity and importance, a dizzying array of tools have been developed to find, characterize, and quantify repetitive DNA. Unfortunately, most programs depend on species-specific libraries or focus on a particular type of repeat. TE-Hub is a community of scientists working on repeats formed to address the ongoing challenges associated with analyzing and understanding genomes, and with the goal of sharing resources and knowledge with those new to this sector of DNA. Our aims are to provide connections among TE researchers, make new tools and resources, facilitate the development of community norms and bench-marks for software development, and offer online and in-person workshops to help introduce and train the next generation of biologists on extant tools. We hope these efforts will advance our understanding of DNA repeats, and further support the growing community of scientists dedicated to the challenge of understanding this dynamic part of the genome.

956 - Disease Detectives: Using Minecraft to explore virus evolution and epidemiology

Kate Duggan

Kate Duggan, Maria Fanourgiaki, Richard Fitzpatrick, Perrin Grant
University of Edinburgh, University of Edinburgh, University of Edinburgh, University of Edinburgh

The COVID-19 pandemic illustrated the importance of clear, accessible public engagement, amidst mistrust and conspiracy theories surrounding SARS-CoV-2. However, the pandemic also served as a high-profile example of evolution in action. Our public engagement project uses now-familiar concepts such as variants of concern and infection waves to engage audiences with evolution and its applications to public health. Based around the popular game Minecraft, we created a workshop that introduces young people to data collection, sequence alignment and public health decisions. Our main aim was to increase familiarity with the use of genomics to respond to viral outbreaks and monitor pathogen evolution. Participants investigate a series of simulated outbreaks, which spread in Minecraft recreations of Scottish cities. The workshop is highly accessible, requiring no coding experience. The popularity of Minecraft and the exploration-based nature of the workshop ensured that even students with minimal interest or experience in evolutionary biology were engaged in gathering and analysing data. Our pilot phase saw the workshop reach over 1000 people across various settings, from drop-in events at music festivals to workshops with schools and community groups. In addition to helping us develop our outreach materials, the pilot allowed us to build new connections with young people and partner organisations across the Scottish Highlands. By using participantsâ€™ lived experience of COVID-19 and familiarity with Minecraft, our workshop serves as an accessible, adaptable method of communicating key concepts in evolution and public health.

204 - Spatial ancestry reconstruction using parsimony on tree sequences

Michael Grundler

Michael Grundler, Jonathan Terhorst, Gideon Bradburd
University of Michigan, University of Michigan, University of Michigan

Many statistical methods for quantifying ancestry average over the ages of shared ancestors in a sample, effectively "flattening" the temporal component of the genealogy that connects all individuals within a species. This flattening has the effect of painting a static notion of ancestry, rather than one that changes as it proceeds backwards in time. Here, we describe a method that capitalizes on the rich genealogical information encoded in genomic tree sequences to infer the geographic locations of the shared ancestors of a sample of sequenced individuals. After validating our method with simulations, we illustrate its application to an empirical dataset and show how a spatiotemporal summary of ancestors offers an intuitive and informative path toward understanding ancestry, demography, and major population movements through time.

210 - A space-time model for jointly inferring gene flow and natural selection

Alessandro Lopez-Hernandez

Alessandro Lopez-Hernandez, Vicente Diego Ortega-Del Vecchyo
Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico, N/A, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico

The computational methods developed to infer past demographic events and the impact of natural selection usually employ statistical models that do not use ancient DNA data and do not take spatial information into account. The use of ancient DNA data is particularly important to infer past demographic events that could not be observed from the analysis of present-day genomic information. On the other hand, the use of spatial information is important to perform inferences of local adaptation events and to do a geographical characterization of migration events. Here we present a new likelihood-based methodology employing a statistical model that uses temporal genomic information and spatial data to infer past demographic events and the impact of natural selection acting on new alleles. Our approach uses partial differential equations to model changes in allele frequencies considering parameters that model gene flow, genetic drift and the impact of natural selection. First, we perform simulations to show that our approach can be used to infer the amount of gene flow across a geographical landscape and to infer the selective coefficient of one advantageous allele or a set of deleterious alleles. Then, we apply our method to genomic data from the â€œAllen Ancient DNA Resourceâ€• of 3672 European individuals from different temporalities that span the present up to 10,000 years ago. We show that our method can infer the amount of gene flow and the impact of natural selection acting on putative deleterious and advantageous alleles from European individuals.

385 - Using biological invasions to understand rapid adaptation to new environments: a genomic reconstruction of the house sparrow global spread

Jack Harper

Jack Harper, Mark Ravinet, Britt Heidinger, Simon Griffith, Lynn Martin
University of Nottingham, University of Oslo, University of Oslo, North Dakota State University, Macquarie University Sydney, University of South Florida

The well-known house sparrow (Passer domesticus) is a small bird renowned for its charismatic nature and close association to humans. While the native range of the house sparrow is Eurasia, due to multiple intentional and accidental introductions from the mid-19th century

onwards, the species is also invasive in Australasia, Southern Africa, and the Americas. In modern North America, the house sparrow now exists from Southern Panama all the way to the Northwest territories of Canada, attesting to the speciesâ€™ ability to survive in a diverse range of environments. This research aims to understand evolutionary success in biological invasion by reconstructing the parameters of the house sparrow invasive spread. Using whole genome resequencing of 500 birds, we demonstrate substantial population structure present within the introduced range, suggesting high levels of divergence between different introduction sources. We also studied parallel latitudinal clines within the United States and Australia to quantify the extent of rapid adaptation to novel environments from bottlenecked founder populations, suggesting that certain genome regions and traits may be more beneficial for evolutionary success. Invasive species are an excellent model for understanding complex evolutionary processes, and this research will inform future work on understanding the genomics of successful adaptation to new environments.

408 - Inference of long-range gene-flow events in a background of spatially heterogeneous isolation-by-distance

Vivaswat Shastry

Vivaswat Shastry, John Novembre
University of Chicago, University of Chicago

Population structure in genetics arises from an evolutionary history of local genetic drift that is offset by gene flow. On one extreme, gene flow can be continuous through time, as might be expected for gene flow between neighboring populations, or it might involve instantaneous bursts of rare long-range events across the habitat. Disentangling the effects of both of these modes of gene flow in a single modeling framework is a largely open challenge.
Here, we present a fast and accurate method called FEEMSmix (an extension to FEEMS from Marcus, Ha et al 2021) that infers long-range gene flow events using single nucleotide variant data in a background of spatially heterogeneous diffusive migration. The method uses observed deviations from a baseline spatially heterogeneous isolation-by-distance model to detect plausible pairs of populations that were involved in long-range gene flow events. Then in a coalescent-based framework, the method infers the magnitude, timing, and direction of these events assuming an instantaneous pulse in the last few generations. The method produces geographic maps annotated with long-range edges corresponding to gene flow events and maps of uncertainty in the geographic location of a particular source. We validate the method using simulations. When we apply this framework to existing data sets from North American Canis lupus (gray wolves), Homo sapiens (humans) from Afro-Eurasia, and Phylloscopus trochiloides (greenish warblers) in Central Asia, we find the method paints a fuller picture of the genetic structure and evolutionary history of each sample.

440 - In space no one can hear you sweep: a novel signal of selective sweeps in continuous geographic space

Clara T Rehmann

Clara T Rehmann, Peter L Ralph, Andrew D Kern
University of Oregon, University of Oregon, University of Oregon

As organisms adapt to environmental changes, natural selection modifies the frequency of non-neutral alleles. For beneficial mutations, the outcome of this process may be a selective sweep, in which the allele rapidly increases in frequency and perhaps reaches fixation within a population. Selective sweeps have well-studied effects on patterns of local genetic variation in panmictic populations, but much less is known about the dynamics of sweeps in continuous space. In particular, because limited movement across a landscape leads to unique patterns of population structure, spatial dynamics may influence the trajectory of mutations under positive selection.

Here, we use forward-in-time, individual-based simulations in continuous space to track a beneficial mutation as it sweeps through a population. We describe the interconnected effects of selection, allele frequency, allele age, and continuous spatial structure on the dynamics of a selective sweep. With realistic simulation, we show that selection changes the joint distribution of allele frequency and geographic range and demonstrate that this signal can be used to identify selective sweeps.
We then leverage this signal to identify in-progress selective sweeps within the malaria vector Anopheles gambiae, a population under strong selection pressure from vector control measures. By considering space, we identify multiple previously-undescribed variants with potential phenotypic consequences, including mutations impacting known IR-associated genes and altering protein structure and properties. Our results demonstrate a novel signal for detecting selection in spatial population genetic data that may have implications for genomic surveillance and understanding geographic patterns of genetic variation.

906 - A spatially explicit close kin mark recapture method for estimating census size from genetic data

Gilia Patterson

Gilia Patterson, Chris Smith, Andrew Kern, Peter Ralph
University of Oregon, University of Oregon, University of Oregon, University of Oregon

The number of individuals in a population, or census size, is one of the most important pieces of information for species that we want to conserve, control, or harvest, but reliably estimating it can be very difficult. Widely used genetic capture-recapture methods require capturing the same individual multiple times, which can be impossible for elusive species or species that are sampled lethally. Close kin mark-recapture (CKMR) is a promising new method that does not require recaptures of the same individual and can make use of a single temporal sample. It uses genetic data to identify related pairs of individuals, and then estimates census size from these pairs. However, a major shortcoming of current CKMR methods is that they do not account for spatial population structure or spatial heterogeneity in sampling, patterns that strongly affect estimates.

We developed a spatially explicit CKMR method using a convolutional neural network trained on simulated populations. The method uses as input maps of parent-offspring pairs on the landscape and maps of sampling intensity, and is able to accurately estimate census size, even when dispersal is limited and some areas of the landscape are more intensely sampled than others. Parent-offspring pairs can be identified from as little as hundreds of SNPs. We are extending our method to use more distantly related individuals, and to estimate a map of population density rather than a single census size. These novel methods have the potential to inform and evaluate conservation efforts for many species.

118 - Structural Enzymology, Phylogenetics, Differentiation and Symbolic Reflexivity at the Dawn of Biology

Charles W. Carter, Jr

Charles W. Carter, Jr, Guo Qing Tang, Sourav Kumar Patra, Laurie Betts, Mark Ditzler, Milena Popovic, Peter R. Wills, Remco Bouckaert, Jordan Douglas
University of North Carolina at Chapel Hill, University of North Carolina at Chapel Hill, University of North Carolina at Chapel Hill, University of North Carolina at Chapel Hill, NASA Ames Research Center, Blue Marble Space Institute of Science, University of AucklanD, University of Auckland, University of Auckland

How did Nature discover the coding rules needed to read blueprints templated in the sequences of their own genes? The reflexive

translation of symbols in one chemical language to another defined genetics. Yet, the co-linearity of codons and amino acids is so commonplace an idea that few even ask this question. That readout is done by two sets of proteins, called aminoacyl-tRNA synthetases (AARS). AARS must enforce the rules first used to assemble them. We will find the root of translation only when we can describe and experimentally validate the earliest AARS*tRNA â€œcognate pairsâ€• and the â€œstructural codesâ€• they used to recognize both amino acid and RNA substrates. Phylogenetics and AI now help identify new AARS â€œurzymesâ€• from novel sources. Weâ€™ve shown that urzymes actually prefer to acylate TÎ¨Cminihelices, rather than full-length tRNAs. Catalysis by urzymes does not need active-site amino acids that were not present at early times. Studies of RNA specificity confirm details of the code used by urzymes to recognize cognate RNA minihelices. We have adduced substantial evidence that both strands of a single ancestral gene coded for Class I and II AARS on opposite strands. Translated bidirectional gene sequences are inside out, and so fold into different 3D structures. That leads to contrasting amino acid and RNA substrate binding modes. These experimental data therefore argue that these initial substrate specificities are rooted in the base pairing between their coding sequences. This work provides phylogenetics a platform for discovering the origin of genetics.

255 - Foldtree. Empirical benchmarking of structural phylogenetics methods

David Moi

David Moi, Charles Bernard, Christophe Dessimoz
UNIL DBC, Institut Pasteur , UNIL DBC

With the emergence of accurate AI generated structural models it has become possible to generate phylogenetic trees of protein families incorporating this structural information. It is still unclear which is the best evolutionary model to represent the phylogenetic signal of the final folded protein as we transition away from methods that represent this object as a series of independently evolving positions towards a more holistic view incorporating tertiary structure and interactions between residues.

In this work we present an empirical benchmarking of thousands of protein families using a battery of structural phylogenetics approaches currently available. Currently our experiments show that the use of structural alphabets to align residues improves alignment quality and subsequent tree topology inference. However, constraining this structural character alignment to statistical models typically used for maximum likelihood phylogenetics does not appear to improve tree topology quality in our benchmarking experiments over the use of simpler models.

The methods we find to be most apt at recapitulating species tree signals from our benchmarking experiments are showcased using difficult phylogenies sometimes involving horizontal transfers and long evolutionary histories stretching back billions of years. We show structural phylogenetics applied to the structural superfamilies of Fusexins (containing Hap2, Fsx1 and viral class II fusogens), the RRNPPA peptidic quorum sensing family and other illustrative examples.

We will also present preliminary work regarding the expansion of structural alphabets tuned specifically for phylogenetic analysis.

537 - Structural phylogenetics of the jelly roll fold sheds light on viral evolution

Desiree Beate Langer

Desiree B. Langer, Ashar J. Malik, Peter F. Stadler, Peter W. Hildebrand, Caroline Puente-Lelievre , Nicholas Matzke, Jane R. Allison, Anthony M. Poole

School of Biological Sciences, University of Auckland, -, School of Chemistry and Molecular Biosciences, University of Queensland, Computational Biology and Clinical Informatics, Baker Heart and Diabetes Institute, Bioinformatics Group, Institute of Computer Science, University of Leipzig, Institute of Medical Physics and Biophysics, University of Leipzig, School of Biological Sciences, University of Auckland, School of Biological Sciences, University of Auckland, School of Biological Sciences, University of Auckland, School of Biological Sciences, University of Auckland

Comparing virus coat architectures has revealed the presence of a common jelly-roll fold in viruses infecting all three domains of life, spanning both RNA and DNA viruses. While structural similarities are evident, protein sequence conservation is low, making traditional sequence-based tools impractical for building viral evolutionary trees. Our lab developed a novel method for constructing evolutionary trees based on protein structure, which can aid in reconstructing deep relationships even when sequence similarity is limited. Our method makes use of molecular dynamics to generate statistical support values for phylogenies built from protein structure. These results will be compared with a method that overcomes the dependence on distance-based methods and uses a structural alphabet (3Di), which is used for the recently developed program Foldseek for deep homology search.
Our phylogenies reconstruct expected viral relationships, indicating that evolutionary signal can be extracted from the comparison of protein structures. Our results also shed light on the evolution of jelly roll architectures themselves, revealing that double jelly roll (DJR) folds have emerged multiple times from the single jelly roll (SJR) fold via ancient gene duplication.

596 - Polyphyletic insertions in the bacterial DNA-directed RNA polymerase

Claudia Alvarez Carreno

Claudia Alvarez Carreno, Angela Huynh, Christine Orengo, Loren Dean Williams
University College London, Georgia Institute of Technology, University College London, Georgia Institute of Technology, NASA Center for the Origin of Life

Transcription is the Central Dogma process in which the RNA polymerase (RNAP) transcribes DNA into RNA. Bacterial RNAP contains five subunits called alpha1, alpha2, beta, betaâ€™, and omega. All these subunits have archaeal orthologs.
Though RNAP-beta and RNAP-betaâ€™ are universally distributed, their multi-domain architectures vary significantly between archaea and bacteria, and among bacterial species. Here, we use sequences and structures to retrace the deep evolution of RNAP-beta and RNAP-betaâ€™ in bacteria.
We present the results of a comparative structural analysis of experimentally determined structures and AlphaFold predictions to investigate the deep divergence of bacterial RNAP-beta and RNAP-betaâ€™.
Our combined analysis suggests that phylogenetic distributions of bacterial lineage-specific insertional domains follow the Tree of Life. We observe that bacterial lineage-specific domains in RNAP-beta belong to a group of domains that we call BEAN (Broadly Embedded ANnex) and that in RNAP-betaâ€™, bacterial lineage-specific domains are HAmmerhead/BArrel-Sandwich hybrid (HABAS) domains. We discuss mechanisms that might account for the presence of homologous insertional domains in non-equivalent locations of the same protein.

658 - Exploring protein evolution through natural and simulated sequences using experimental and modeled 3D structures

Hector Romero

Hector Romero, Mauricio Langleib, Martin GraÃ±a
Universidad de la RepÃºblica, Institut Pasteur, Institut Pasteur

Evolving new molecular functions and fine tuning existing ones is the core of evolutionary processes at the molecular level. In turn, in many cases this is the foundation of changes in higher order structures such as complexes, networks (metabolic, signalization, etc.), both at cellular and organismic levels. The path followed by molecular machines through this process has fascinated biologists for decades. Coupling present sequencing capabilities with the power of deep learning methods for protein structure prediction, we explore the fate of theoretical mutations and substitutions within predicted structures and compare it with natural observations using both experimental (when available) and modeled data. First, we took a forward in time approach, in which we simulate evolutionary processes at the sequence level using well known markov models of molecular evolution both at the DNA level (e.g. HKY, GTR, etc.) and directly at the protein level (e.g. Dayhoff, JTT, LG, etc.). Sampling through different evolutionary times and modeling 3D structure in each moment performing 100 different trajectories. And then compare the foldability and difference with the original structure. Then, we used actual sequence data of homologs to model (and when available used experimental 3D structures) to obtain â€˜naturalâ€™ variants of the potential trajectories.
Measures like TMscore, RMSD, stability of the protein, etc. are taken into account in order to evaluate the viability of the variants, and study their distribution in the different models.
Finally we compare the actual sequences and their variability with the simulated ones.

905 - AlphaHOG, a protein structure-based reference classification to improve orthology inference

stefano pascarelli

stefano pascarelli, Christophe Dessimoz
University of Lausanne, ETHZ, Swiss Institute of Bioinformatics, University of Lausanne, Swiss Institute of Bioinformatics

The increasing availability of genomic sequences is driving forward our understanding of the diverse life forms on Earth. However, the ability to generalize organism-specific knowledge is limited by how well we relate genomes to each other. Genomes can be compared in terms of orthologous genes â€” genes of different species that derive from a single gene in the last common ancestor. Currently, the majority of orthology prediction software is based on the amino acid sequence, the most abundant information about proteins. However, the sequence signal of distantly related genes is weak, distributed in saturated positions, and confounded by evolutionary forces. In this work, I show how the more conserved protein 3D-structure can improve orthology prediction. I devised a method that combines sequence k-mers with k-mers generated from a local structural alphabet. The enriched sets of k-mers can be used to generate a reference classification of proteins into Hierarchical Orthologous Groups (HOGs) â€” a coarse-grained representation of protein families. The structure-informed reference HOGs, here named AlphaHOGs, can be exploited to infer orthology in a set of proteomes, by using the recently developed software FastOMA. As a test case, we reconstruct the ancestral genome of the first multicellular animal with an unprecedented resolution, paving the way to higher-level analyses such as the ancestral gene content and protein interaction network, potentially shedding light on the current uncertainties about the origin of the animal lineage.

44 - Contribution of variable TE content on phenotype and plasticity in Drosophila melanogaster

Anaïs Larue

AnaÃ¯s Larue, Miriam Merenciano, Sonia Janillon, AgnÃ¨s Vallier, Marie Fablet, Patricia Gibert, Abdelaziz Heddi, SÃ©verine Chambeyron, Rita Rebollo, Cristina Vieira
Laboratoire de Biologie Fonctionnelle Insectes et Interactions, Laboratoire de BiomÃ©trie et Biologie Evolutive, Laboratoire de BiomÃ©trie et Biologie Evolutive, Laboratoire de BiomÃ©trie et Biologie Evolutive, Laboratoire de Biologie Fonctionnelle Insectes et Interactions,

Laboratoire de BiomÃ©trie et Biologie Evolutive, Laboratoire de BiomÃ©trie et Biologie Evolutive, Laboratoire de Biologie Fonctionnelle Insectes et Interactions, Institut de GÃ©nÃ©tique Humaine, Laboratoire de Biologie Fonctionnelle Insectes et Interactions, Laboratoire de BiomÃ©trie et Biologie Evolutive

Phenotypic plasticity is the ability for one genotype to yield different phenotypes in response to environmental changes. This process may provide organisms with the capacity to adapt as the environment varies. Some mechanisms at play in generating phenotypic plasticity are epigenetic mechanisms, such as histone marks, DNA methylation or small RNAs, allowing differential expression of genes, and eventually result in phenotypic changes. Epigenetic mechanisms have often been documented to fluctuate with the environment. Not only these mechanisms regulate gene expression, but they are also involved in the repression of transposable elements (TEs). TEs are repeated DNA sequences potentially capable of moving and replicating within genomes. Thanks to their transposition mechanisms, but also to their inherent regulatory sequences, TEs are known to impact gene expression, genome architecture and integrity. We hypothesize that TEsâ€™ control mechanisms and regulatory components, might also vary with the environment and potentially lead to changes in TE expression and transposition rate, and in turn alter gene expression. To test this hypothesis, we investigated the influence of TE content on phenotypes in Drosophila melanogaster, and as a potential driver of phenotypic plasticity. We took advantage of genetically engineered fruit flies that bear different TE content but identical genetic background, to tackle this question in a controlled manner. We investigated the plastic response of diverse life history traits in a variety of environments and stressors. Results depict genotype by environment interactions for multiple traits, suggesting that indeed TEs contribute to plasticity.

71 - DNA Transposons favour de novo transcript emergence through enrichment of transcription factor binding motifs

Marie Lebherz

Marie Lebherz, Bertrand Fouks, Julian Schmidt, Erich Bornberg-Bauer, Anna Grandchamp
Institute for Evolution and Biodiversity, Institute for Evolution and Biodiversity, Institute for Evolution and Biodiversity, Institute for Evolution and Biodiversity, Department of Protein Evolution, Max Planck Institute for Biology TÃ¼bingen, Institute for Evolution and Biodiversity

Transposable elements (TE) are known to play an essential role in driving evolvability and thus facilitating fast acting molecular adaptive processes at the DNA level. We report here on how TEs can affect the recently detected and meanwhile well confirmed process of de novo gene emergence. Specifically, we focused on the gain of new transcription events. In many species a continuum between absent and very prominent transcription has been reported for essentially all regions of the genome. In this study we searched for de novo transcripts by using newly assembled genomes and transcriptomes of seven inbred lines of Drosophila melanogaster, originating from seven geographically diverse populations. This setup allowed us to detect line specific de novo transcripts, and compare them to their homologous non-transcribed regions in other lines, as well as genic and intergenic control sequences. We studied the association with TE and the enrichment of transcription factor motifs upstream of de novo emerged transcripts. We found that de novo transcripts overlap with TEs more often than expected by chance. The emergence of new transcripts correlates with epigenetic marks islands and regions of TEs activity. Moreover, upstream regions of de novo transcripts are highly enriched with regulatory motifs. Such motifs abound in new transcripts overlapping with TEs, particularly DNA TEs, and are more conserved upstream de novo transcripts than upstream their non-transcribed homologs. Overall, our study demonstrates that TE insertions play an important role in new transcript emergence, partly by introducing new regulatory motifs from DNA TE families.

337 - Double trouble: two retrotransposons triggered a cascade of horizontal transfers in Drosophila

species within the last 50 years

Riccardo Pianezza

Riccardo Pianezza, AlmorÃ² Scarpa, Sarah Signor, Robert Kofler

Veterinary Medicine University, Vienna Graduate School of Population Genetics, Veterinary Medicine University, Vienna Graduate School of Population Genetics, North Dakota State University, Veterinary Medicine University

Horizontal transfer of genetic material in eukaryotes has rarely been documented in short evolutionary timescales. We show that two retrotransposons, Shellder and Spoink, invaded the genomes of multiple species of the D. melanogaster subgroup in the past 50 years. Through horizontal transfer, Spoink spread in D. melanogaster during the 1980s, while both Shellder and Spoink invaded D. simulans in the 1990s.

Likely following hybridization, D. simulans infected the island endemic species D. mauritiana (Mauritius) and D. sechellia (Seychelles) with both TEs after 1995. Shellder additionally invaded D. teissieri, a species confined to sub-Saharan Africa, in approximately the same time-frame.

The donors of Shellder and Spoink are likely American Drosophila species from the willistoni, cardini, and repleta groups. Thus, the described cascade of TE invasions only became feasible after D. melanogaster and D. simulans extended their habitats into the Americas 200 years ago, likely aided by human activity.

Our work reveals that cascades of TE invasions, likely initiated by human-mediated habitat expansions, could have a profound impact on the genomic and phenotypic evolution of many geographically dispersed species.

Within a few decades, TEs can invade many species, including island endemics, with habitats very distant from the initial donor of the TE.

366 - Evolution of transposable elements in arbuscular mycorrhizal fungi: insights into genome variability and gene regulation

Jordana I.N. Oliveira

Jordana I.N. Oliveira, Catrina Lane, Ariane M. Nicol, Nicolas Corradi

University of Ottawa, University of Ottawa, Queen's University, University of Ottawa

Arbuscular mycorrhizal fungi (AMF) are plant symbionts that provide nutrients from the soil to their hosts in exchange for carbohydrates and lipids. AMF exibit low morphologic variability and no sign of plant specificity, raising questions about how these organisms have adapted to a wide spectrum of host species. Genomic analysis has revealed that AMF species harbor a significant portion of transposable elements (TEs), which play a crucial role in promoting genome variability by generating new copies and insertions. However, on average 60% of TEs in AMF genomes remain unknown. We hypothesize that investigating the TE content can unveil AMF diversity and detect the impact of these sequences in promoting genetic variability. To achieve this, we constructed a TE library using 26 AMF species, identifying the presence of transposition domains in consensus sequences. The annotation of TEs revealed that the major groups of AMF share patterns of transposition, with ancestral Paraglomerales exhibiting lower TE content and smaller genome sizes compared to more derived groups Archeosporales, Glomerales, and Diversisporales, which show clade-specific expansions. Furthermore, a significant proportion of TE insertions are located in exons, suggesting an impact of these repetitive sequences on protein diversity through exonization processes. Expression analysis in the model species Rhizophagus irregularis showed a fourfold increase in upregulated TEs during symbiosis compared to extraradical mycelium, indicating a role of TEs in communicating with the host. We propose that TE variability in AMF may contribute to the evolution and diversification of genes, influencing their successful symbiotic relationships with plants.

688 - Independent transitions to sociality in Stegodyphus spiders are associated with transposable element expansions

Jilong Ma

Jilong Ma, Anna Igolkina, Jesper Bechsgaard, Palle Villesen, Trine Bilde, Magnus Nordborg, Mikkel Heide Schierup
Aarhus University, Gregor Mendel Institute of Molecular Plant Biology, Aarhus University, Aarhus University, Aarhus, Gregor Mendel Institute of Molecular Plant Biology, Aarhus University

Transposable elements (TEs) can insert into functional genomic regions and induce ectopic recombinations in the genome. Due to these larger effect mutations, transposable elements are a possible threat to the genome integrity in evolution. The threat is expected to be larger when nature selection is less efficient in inhibiting TE proliferation. Social evolution in the spider genus Stegodyphus is a case where the development of a life history trait reduces selection efficiency. The evolution of sociality is repeatable in Stegodyphus across three independent transitions within the past hundred-thousands years. All social species appear to be evolutionary short-lived in the phylogeny. We wish to test whether TE could be a reason for the hypothesized evolutionary dead end, exploiting de novo haplotype-phased long-read genomes of 7 species in Stegodyphus. We used polymorphisms of structural variants from alignments of haploid genomes to identify all mobile TE families across Stegodyphus. We then constructed and contrasted time-resolved TE proliferation rate per TE family in closely related pairs of social and subsocial species. We find that more TE families expand, and each to a larger extent, in social species. The more TE proliferation in social lineages correlates with the establishment of sociality over time. The recently acquired TE loci of these TE families also show a higher expression level in social species. We conclude that the social evolution in spiders, representing the development of a high-inbreeding mating system, leads to the expansion of TEs.

694 - TEforest: Enhancing short-read transposable element detection with machine learning

Austin T Daigle

Austin T Daigle, Logan Whitehouse, Dan Schrider
UNC Chapel Hill, UNC Chapel Hill, UNC Chapel Hill

Transposable elements (TEs) are genomic elements that are ubiquitous across the tree of life and play a crucial role in genome evolution and function. Advances in long-read sequencing have allowed highly accurate TE detection, though at a higher cost than short-read sequencing. Recent long-read studies have shown that existing short-read TE detection methods perform inadequately when applied to real data. In this study, we use a machine learning approach (called TEforest) to discover and genotype TE insertions and deletions with short-read data by leveraging true-positive annotations from long-read genome assemblies. Our method first uses a highly sensitive algorithm to discover potential TE insertion or deletion sites in the genome, extracting relevant features from short-read alignments. To discriminate between true and false TE insertions, we train a random forest model with a labeled ground-truth dataset for which we have measured the same set of short-read features. We conduct a comprehensive benchmark of TEforest and traditional TE detection methods using real data, finding that TEforest identifies more true positives and fewer false positives across datasets with different read lengths and coverages, while also accurately inferring genotypes and the precise breakpoints of insertions. By learning short-read signatures of TEs previously only discoverable using long reads, our approach bridges the gap between large-scale population genetic studies and the accuracy of long-read assemblies. This work provides a user-friendly tool to study the prevalence and phenotypic effects of TEs across species.

97 - Novel topological units characterise a distinct regulatory architecture underlying the evolution of complex traits in coleiod cephalopods

Thea F. Rogers

Thea F. Rogers, Oleg Simakov
University of Vienna, University of Vienna

Our understanding of how genomic changes translate into organismal novelties is often confounded by complex and multi-layered genome architecture. Coleoid cephalopods (squid, octopus and cuttlefish) have large, uniquely structured brains, as well as many species-specific innovations, such as camouflaging ability, complex behaviours and novel organs. These have recently been identified to be associated with an interplay of various of genomic mechanisms, including large-scale genome reorganisation in the ancestor of all coleoid cephalopods, repetitive element accumulation and tandem gene duplications resulting in significant genome expansion. However, regulatory aspects of this genomic architecture remain understudied. We use Micro-C, ATAC-seq and transcriptomic data from several tissues and developmental stages in coleiod cephalopods to assess how a distinct evolutionary history has impacted gene regulation and led to the emergence of cephalopod innovations. We classify locally conserved interacting gene pairs as coleiod-specific â€˜spatiosyntheniesâ€™ and uncover distinct patterns of regulatory features associated with their emergence, including chromatin remodelling, alternative splicing, RNA editing, noncoding RNAs and repetitive elements. We find that these interactions often occupy pronounced â€˜bi-TADâ€™ arrangements in coleiod genomes, and suggest that relaxed regulatory constraints on cephalopod genomic regions led them to undergo rearrangement and expansion, resulting in their topological linkage and such novel regulation. Taken together, our results describe a distinct â€˜unitâ€™ of regulatory evolution for coleiod cephalopods.

287 - Gene expression variability implies stable cis-regulation of brain genes across non-nervous organs

Christabel-Floi Bucao

Christabel-Floi Bucao, ConsolÃ©e Aletti, SÃ©bastien Moretti, Julien Bobe, Ingo Braasch, Yann Guigen, John H. Postlethwait, Marc Robinson-Rechavi
Department of Ecology and Evolution, University of Lausanne, SIB Swiss Institute of Bioinformatics, Department of Ecology and Evolution, University of Lausanne, Department of Ecology and Evolution, University of Lausanne, SIB Swiss Institute of Bioinformatics, INRAE, LPGP, Department of Integrative Biology, Michigan State University, INRAE, LPGP, Institute of Neuroscience, University of Oregon, Department of Ecology and Evolution, University of Lausanne, SIB Swiss Institute of Bioinformatics

Gene expression variability broadly refers to a geneâ€™s tendency to vary in expression due to stochastic fluctuations or differences in genetic, epigenetic, or environmental factors. Variability due to transcriptional noise and cell-to-cell heterogeneity has been well-studied in isogenic populations of bacteria and yeasts. However, for complex organisms with multiple cells, tissues, and organs sharing a genetic background, the interplay between variability, gene and organ function, and gene regulation remains an open question. We used 3`-end bulk RNA barcoding and sequencing (BRB-seq) to generate transcriptome profiles spanning at least nine organs in outbred individuals of three ray-finned fishes: zebrafish, northern pike, and spotted gar. Per organ, we measured inter-individual expression variation per gene independent of mean expression. We observed that lowly variable genes are enriched in cellular housekeeping functions whereas highly variable genes are enriched in stimulus-response functions. Furthermore, highly variable genes evolve under weaker purifying selection at the coding sequence, indicating that intra-species gene expression variability predicts inter-species protein sequence divergence. For genes with organ-biased expression, we inferred differences in selective pressure on gene regulation depending on their top organ. Genes that are most expressed in the brain show low expression variability across non-nervous organs, suggesting stabilizing selection on cis-regulatory evolution of brain-biased genes. By contrast, liver-biased genes have highly variable expression across other organs, implying weaker cis-regulatory constraints. Thus, investigating expression variability across organs can provide new insights into gene

regulation in animals.

376 - Population genomics of wild avian embryos uncovers hidden genetic diversity

Justin Wilcox

Nikolas Vellnow, Quentin Foucault, Justin Wilcox, Yu-Chi Chen, Sahar Javaheri, Toni Gossmann
TU Dortmund University, Computational Systems Biology, Faculty of Biochemical and Chemical Engineering, TU Dortmund University, Computational Systems Biology, Faculty of Biochemical and Chemical Engineering, TU Dortmund University, Computational Systems Biology, Faculty of Biochemical and Chemical Engineering, TU Dortmund University, Computational Systems Biology, Faculty of Biochemical and Chemical Engineering, TU Dortmund University, Computational Systems Biology, Faculty of Biochemical and Chemical Engineering, TU Dortmund University, Computational Systems Biology, Faculty of Biochemical and Chemical Engineering

Population genomics analyses are usually based on samples obtained from genetic variation in adult individuals. However, conducting genetic analyses using adult samples limits our ability to understand the full range of genomic variation such as during early development. Songbirds are ideal to study such variation because they possess an additional chromosome in the germline that is eliminated from somatic tissue. However, the precise mechanism of such chromosome elimination remains obscure. To address the intriguing question of developmental genomic variation, we obtained embryonic DNA from 14 great tit individuals isolated from eggs harvested in a local forest population in Bielefeld (Germany). We conducted a series of whole genome population genetic analyses by contrasting embryo samples to adult great tit individuals sampled across whole Europe. Interestingly, our embryo samples had a higher genetic diversity than the adult blood samples. We also found remarkable differences in genetic diversity between female and male embryos. Our observed patterns are hard to explain with standard Mendelian genetics or processes of selection like heterozygote advantage. Instead, they may suggest other forms of genomic variation, e.g. sex- and age-specific structural variation or a cryptic karyotype polymorphism. This illustrates that relying on standard pipelines for population genomic analysis may lead to false conclusions. Here, our analysis of wild embryos has made us able to detect life-stage dependent genetic variation, which would otherwise have been unnoticeable.

477 - Beyond averages: transcriptional variability in outbred Drosophila melanogaster and its environmental and genetic dependence

James Phipps-Tan Sheng Yi

James Phipps-Tan, Julien F. Ayroles, Luisa F. Pallares
Friedrich Miescher Laboratory of the Max Planck Society, Lewis-Sigler Institute for Integrative Genomics, Princeton University, Department of Ecology and Evolutionary Biology, Princeton University, Friedrich Miescher Laboratory of the Max Planck Society

Why do different individuals of a population have different phenotypes? Most approaches to this age-old question compare average trait values of individuals with different genotypes or exposed to different environments. However, this ignores individual variation around averages (variability), which is increasingly recognized as a potentially adaptive property arising from developmental noise, bet-hedging and environmental robustness. To investigate the adaptive potential of gene expression variability, we collected the largest-known set (~2000) of matched head transcriptomes and genomes of outbred Drosophila melanogaster individuals raised for a generation on either a control or high sugar diet. This gave us unrivalled statistical power to test whether transcriptional variability is (1) gene-specific, (2) modified by environment and, for the first time in any animal system, (3) has an environmentally-dependent genetic basis. We show that a geneâ€™s transcriptional variability, measured by the outlier-robust median absolute deviation statistic (MAD), correlates positively with its nucleotide diversity and heritability, negatively with its essentiality, and non-linearly with its connectivity. Strikingly, high sugar

increases transcriptional variability in 77% of genes and changes the identities of almost half of the 5% least variable genes. In stark contrast, the 5% most variable genes remain largely the same across diets. Finally, we mapped the genetic basis of transcriptional variability using a new MAD-based mixed modelling approach that explicitly tests for genotype-by-environment interactions. Taken together, our observations that transcriptional variability changes with environmental and genetic variation hint towards it being an adaptive and evolvable trait.

742 - On mutational neighbourhoods and their influence on evolutionary processes

Paco Majic Bergara

Paco Majic
EMBL

The inheritance of genetic information and the adaptive consequences of variation in that information are concepts that are at the heart of Evolutionary Biology. However, another concept that is typically overlooked is the inheritance of the mutational potential that is associated to each genetic sequence. During reproduction, what is passed on to the descent is not only the genotype that informs phenotypic development, but also that genotypeâ€™s mutational neighborhood - that is, the likelihood that each genotype has of mutating into each of every other genotype. I will discuss two ways in which such hidden mutational potential can impact the evolutionary trajectories of populations and entire lineages. Firstly, at a microevolutionary scale, I will focus on how the inheritance of mutational neighborhoods can bias selection and facilitate adaptation by affecting phenotypic development. Secondly, at a macroevolutionary scale, I will describe how mutational neighbourhoods might explain patterns of phylogenetic diversification that have traditionally been ascribed to other evolutionary forces. Summoning cases from developmental biology and palaeontology, my discussion will orbit the framework of genotype spaces, genotype-phenotype maps and adaptive landscapes. Overall, the view of evolution I will present is a view that does not consider species as standalone products of genomic information, but which considers the evolutionary fate of species to be influenced by the material and historical context of that information.

1112 - The evolution of transcriptome complexity using Oxford Nanopore long-read sequencing technology.

Gabriela Santos Rodriguez

Gabriela Santos-Rodriguez, Favour Oyelami, Agin Ravindran, Akanksha Srivastava, Chi Kin Ip, Pallavi Gupta, Jeannette Villanueva, Helen Elizabeth King, Abigail Grootveld, Ishaan Gupta, Graziele Santos Vieira Helaine, Nikolay Shirokikh, Eduardo Eyras, Robert James Weatheritt
School of Biotechnology and Biomolecular Sciences, RNA Institute, University of New South Wales, John Curtin School of Medical Research, Australian National University, John Curtin School of Medical Research, Australian National University, John Curtin School of Medical Research, Australian National University, Garvan Institute of Medical Research, Department of Biochemical Engineering & Biotechnology, Indian Institute of Technology Delhi, Victor Chang Cardiac Research Institute, Garvan Institute of Medical Research, Garvan Institute of Medical Research, Department of Biochemical Engineering & Biotechnology, Indian Institute of Technology Delhi, Garvan Institute of Medical Research, John Curtin School of Medical Research, Australian National University, John Curtin School of Medical Research, Australian National University, EMBL Australia, Garvan Institute of Medical Research, School of Biotechnology and Biomolecular Sciences, University of New South Wales

The increase of biological complexity is partly due to the dynamic generation of unique cell-specific transcriptomes. Co- and post-transcriptional mechanisms including alternative splicing (AS), alternative promotor usage (APU); alternative polyadenylation (APA)

have been identified as major drivers of transcriptome diversity and are known to drive lineage-specific phenotypic effects. However, this research is limited to the assessment of individual events (exons / 3â€™UTRs) due to the limitations of short-read sequencing

To gain insight into the conservation of full-length transcript expression and the coordinated regulation of post-transcriptional mechanisms, we used Oxford Nanopore long-read sequencing technology to analyse isoform conservation across six tissues from five mammalian species and an outgroup. We found that multiple transcript isoforms have conserved tissue-specific expression, where >40% of the analysed genes in mammals had conserved transcripts with major isoform switching in a tissue-specific manner. Additionally, we found that conserved transcripts are enriched in alternative first exon splicing events. Furthermore, we studied the evolutionary conservation of coordinate-splicing events (the co-association of splicing events in the same transcript). We found that mammalian conserved events are depleted in mutually exclusive coordinate-splicing events and enriched in mutually associated events (present in the same transcript) and with a strong enrichment for tissue-specific expression. Together, our work uncovers extensive conservation of post-transcriptional regulation and provides a resource for understanding the transcriptome conservation across mammalian evolution.

330 - RNA virus genomes from historical specimens of great apes

Sojung Han

Sojung Han, Christian Urban, Olivia Cheronet, Ron Pinhasi, Verena Schuenemann, SÃ©bastien Calvignac-Spencer, Martin Kuhlwilm
Department of Evolutionary Anthropology, University of Vienna, Human Evolution and Archaeological Sciences (HEAS), University of Vienna, Institute of Evolutionary Medicine, University of Zurich, Department of Evolutionary Anthropology, University of Vienna, Human Evolution and Archaeological Sciences (HEAS), University of Vienna, Department of Evolutionary Anthropology, University of Vienna, Human Evolution and Archaeological Sciences (HEAS), University of Vienna, Institute of Evolutionary Medicine, University of Zurich, Department of Environmental Sciences, University of Basel, Helmholtz Centre for Infection Research, Department of Evolutionary Anthropology, University of Vienna, Human Evolution and Archaeological Sciences (HEAS), University of Vienna

Infectious diseases are one of the main drivers of evolutionary change, with RNA viruses such as measles, influenza and coronaviruses being a major threat. With their potential for spillover between species, understanding pathogen genomes across species is important. Pathogens also appear to have quite varying mutation rates, which highlight the necessity of sampling their genomes at different historical time points. Compared to ancient DNA virus genomes, ancient RNA virus genomes are much less studied, due to technical difficulties: RNA is less stable than DNA, and RNases are abundant in our environment, easily degrading RNA molecules. However, it has been shown that RNA virus genomes could be sequenced from lung tissue kept in formalin, and that RNA fragments could be retrieved from museum muscle and skin specimens older than a century. It is necessary to obtain knowledge on the types of historical specimens that may preserve sufficient amounts of RNA for sequencing, in particular from viruses. Here, we extract RNA from different types of historical great ape museum specimens, such as formalin-preserved tissue, teeth, dental calculus, skin and fixative. Using protocols for ancient nucleotides adapted for RNA, we determine differences in host RNA content and complexity across these sample types for a total of 100 specimens. Since RNA viruses are expected to be present, if at all, at low abundance, we perform enrichment capture for a range of lineages. This allows us to obtain insights into the historical RNA virome of great apes, and delineate limitations of specimen types.

597 - Deciphering the emergence and evolutionary history of a bacterial crop pathogen: insights from historical herbarium specimens

Paola Elvira Campos

Paola E. Campos, Olivier Pruvost, Karine Boyer, Myriam Gaudeul, ClÃ¡udia Baider, Timothy M.A. Utteridge, Nathalie Becker, Adrien

Rieux, Lionel Gagnevin
UCD, CIRAD, CIRAD, MNHN, Mauritius Herbarium, Royal Botanic Gardens, MNHN, CIRAD, CIRAD

Crop pathogens have been a threat to mankind since the beginnings of agriculture. In order to better understand current crop diseases and prevent future epidemics, it is essential to appreciate the factors underlying the emergence, adaptation and spread of pathogens. Recent methodological developments in molecular epidemiology now allow reconstructing disease dynamics in space and time. While the majority of studies previously carried out are entirely based on the sampling of contemporary specimens dating from the last decades, the advent of paleogenomics today makes possible the reconstruction of historical genomes and the study of pathogens evolutionary history with greater precision.
Here, we reconstruct 13 historical genomes of the bacterial crop pathogen Xanthomonas citri pv. citri (Xci) from infected Citrus herbarium specimens. Following ancient DNA authentication, we compare them to modern genomes representative of the worldwide genetic diversity to estimate their phylogenetic relationships, pathogenicity-associated gene content and several evolutionary parameters on a global scale. Our results indicate that Xci originated in Southern Asia ~11,500 years ago (perhaps in relation to Neolithic climate change and the development of agriculture) and diversified during the beginning of the 13th century, after Citrus diversification and before spreading to the rest of the world (probably via human-driven expansion of citriculture through early East-West trade and colonisation). Our work emphasises the importance of historical data in the reconstruction of crop pathogens evolutionary history, valourising naturalist collections and generating knowledge bearing the potential of improving disease monitoring and sustainable control of current and future epidemics.

1005 - Characteristics of a 5,500-year-old Treponema pathogen from Colombia

Nasreen Z Broomandkhoshbacht

Nasreen Z Broomandkhoshbacht, Davide Bozzi, Jane Buikstra, Carlos Eduardo G Amorim, Gilbert Greub, Lars Fehren-Schmitz, Miguel Delgado, Anna-Sapfo Malaspinas, Elizabeth A Nelson
UCSC Human Paleogenomics Lab, Department of Anthropology, University of California Santa Cruz, UC Santa Cruz Genomics Institute, University of California Santa Cruz, Department of Computational Biology, University of Lausanne, Center for Bioarchaeology, Arizona State University, Department of Biology, California State University Northridge, Institute of Microbiology, Lausanne University Hospital and University of Lausanne, UCSC Human Paleogenomics Lab, Department of Anthropology, University of California Santa Cruz, UC Santa Cruz Genomics Institute, University of California Santa Cruz, CONICET-Anthropology Division, Faculty of Natural Sciences and Museum, National University of La Plata, Department of Anthropology and Human Genetics, Fudan University, Department of Computational Biology, University of Lausanne, Microbial Paleogenomics Unit, Department of Genomes and Genetics, Institut Pasteur

The geographic origin, evolution and spread of treponemal diseases remains one of the most debated topics of infectious disease history. Treponema pallidum subspecies, the pathogens known to cause modern treponemal diseases that have been genomically characterized, are closely related yet cause various disease syndromes. Each subspecies was traditionally thought to be linked to its own distinct clinical manifestation and environmental context. However, recent genomics research on both modern and ancient Treponema pallidum subspecies has cast doubt on this assumption.
Genomic reconstructions of Treponema pallidum subsp. endemicum from pre-European contexts in Brazil and Treponema pallidum subspecies from 15th-century Europe have provided new perspectives on the evolutionary history and phylogeographic patterns of these subspecies. In this study, we unveil a 5,500-year-old Treponema pallidum-like pathogen discovered in Sabana de BogotÃ¡, Colombia. This work provides the oldest reconstructed Treponema genome to date and offers the first pre-Hispanic representation of Treponema in this

region. Our contribution expands the existing genomic dataset of treponemal pathogens by millennia, significantly advancing our understanding of treponematosis in the past. Phylogenomic analysis of the 1.7-fold coverage genome indicates this pathogen is basal to all Treponema pallidum subspecies for which genomic data are available, including all published ancient strains. In conjunction with archaeological and paleopathological evidence, our findings contribute to discussions on the disease landscape of mid-Holocene populations of the Americas, treponemal evolution, and the geographic presence of treponemal disease in the past.

1020 - Exploring Pathogens and Demographic Dynamics in Colonial Mexico City through Paleogenomics

Laura Carrillo-Olivas

Laura Carrillo-Olivas, Miriam B. Bravo-Lopez, Viridiana Villa-Islas, Ernesto Garfias-Morales, Alejandra Castillo, Julie Wesp, Marcela Sandoval-Velasco, Lourdes Marquez-Morfin, Jorge A. Gomez-Valdes, Maria de la Luz Moreno-Cabrera, Juan Gabriel Cervantes-Rosado, Alejandro Meraz-Moreno, Emilia Huerta-Sanchez, Flora Jay, Maria Avila-Arcos

Postgraduate in Biological Sciences, Institute of Ecology, National Autonomous University of Mexico (UNAM), International Laboratory for Human Genome Research (LIIGH), National Autonomous University of Mexico (UNAM), International Laboratory for Human Genome Research (LIIGH), National Autonomous University of Mexico (UNAM), International Laboratory for Human Genome Research (LIIGH), National Autonomous University of Mexico (UNAM), International Laboratory for Human Genome Research (LIIGH), National Autonomous University of Mexico (UNAM), International Laboratory for Human Genome Research (LIIGH), National Autonomous University of Mexico (UNAM), Department of Sociology & Anthropology, North Carolina State University (NCSU), Center for Genomic Sciences, Autonomous National University of Mexico (UNAM), Escuela Nacional de Antropologia e Historia (ENAH), Escuela Nacional de Antropologia e Historia (ENAH), Direccion de Salvamento Arqueologico, Instituto Nacional de Antropologia e Historia, Direccion de Salvamento Arqueologico, Instituto Nacional de Antropologia e Historia, Direccion de Salvamento Arqueologico, Instituto Nacional de Antropologia e Historia, Department of Ecology and Evolutionary Biology, Brown University, Laboratoire Interdisciplinaire des Sciences du Numerique, CNRS, INRIA, Universite Paris-Saclay, Postgraduate in Biological Sciences, Institute of Ecology, National Autonomous University of Mexico (UNAM), International Laboratory for Human Genome Research (LIIGH), National Autonomous University of Mexico (UNAM)

After European colonization, new pathogens caused epidemics that decimated the Indigenous population in the Americas. To understand post-colonization pathogen circulation, we analyzed the paleogenomics of individuals from El Templo de la Inmaculada ConcepciÃ³n 'La Conchita' (n=15) and the Hospital Real San JosÃ© de los Naturales â€˜HSJNâ€™ (n=82), dated to 19th 16th-17th centuries respectively. We generated shotgun sequencing for both human and nonhuman genomes. For the human fraction, we determined the mitochondrial (mtDNA) haplogroups and the molecular sex. Both sites reveal a majority of Indigenous mtDNA lineages (n=75) with high confidence, and one HSJN individual has an African mtDNA haplogroup. This aligns with La ConchitaÂ´s reported Indigenous demographics and documentary evidence that HSJN was designated for mainly Natives (â€œNaturalesâ€•), but had a more diverse patient population. Sex distribution analysis indicates a high proportion of males in La Conchita, contrasting with a more balanced sex ratio in HSJN. The non-human fraction was taxonomically classified using Kraken2, and pathogenic bacterial species were identified through a thorough comparison with the virulence factors database VFDB2022. In HSJN, we identified the oral pathogen Tannerella forsythia (n=12), the agent of paratyphoid fever Salmonella enterica Paratyphi C (n=3), and tuberculosis Mycobacterium tuberculosis (n=1). Streptococcus pneumoniae, associated with respiratory infections, was identified in both sites (n=2). These results emphasize the multiplicity of pathogens circulating in colonial Mexico. Further sequencing of the ancient pathogen genomes of interest will allow phylogenetic analyses to identify potential epidemiological relationships between these ancient strains and other ancient and present-day pathogens.

1121 - Tuberculosis and Sociocultural Dynamics:Critical Insights from Paleogenomics in South

America and Beyond

Elizabeth A. Nelson

Elizabeth A. Nelson, Jane E. Buikstra, Nicolas Rascovan
Institute Pasteur, Arizona State Unverisity, Institute Pasteur

For millennia, tuberculosis (TB) has exacted a profound toll on human populations, marked by significant mortality rates and societal upheaval. Despite its ancient history, TB remains a leading cause of death worldwide, second only to COVID-19 in recent years. The emergence of antimicrobial-resistant strains underscores the urgent need for an improved comprehensive understanding of TB persistence on a molecular and social level. Paleogenomics is perfectly positioned to contribute contextualized genomic insights into TB evolution, shedding light on large-scale social forces influencing pathogen dynamics and persistence. Paleogenomics has already revolutionized our understanding of TB in the ancient Americas through molecular confirmation and genomic characterization of TB-causing pathogens in pre-Hispanic contexts, providing pivotal insights. Drawing from published literature and ongoing research, we explore the transformative impact of paleogenomic investigations on our understanding of TB histories in the Americas. We examine the TB presence and strain diversity across varied geocultural and ecological landscapes of pre-Hispanic South America. Through an interdisciplinary lens encompassing archaeology, paleopathology, and paleogenomics, we view these results within biocultural and environmental contexts, elucidating the intricate interplay between pathogens, people, and places over time. Considering TB epidemics' association with socio-political turbulence, climatic disruptions, and subsistence strategies, we argue that contextualizing TB evolution yields critical insights into its prevalence and transmission patterns. Looking forward, we highlight key considerations and future directions to enhance the utility of pathogen paleogenomics in advancing our understanding of infectious disease dynamics in the Americas and beyond.

Poster presentations

100 - Protein language models uncover differential tolerance to substitutions and deletions in the human proteome

Grant Goldman

Grant Goldman, Vasilis Ntranos

Biological and Medical Informatics Graduate Program, University of California, San Francisco, Bakar Computational Health Sciences Institute, University of California, San Francisco, Department of Epidemiology & Biostatistics, University of California, San Francisco, Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, Diabetes Center, University of California, San Francisco

Deep mutational scanning (DMS) experiments have been successfully leveraged to understand the genotype to phenotype mapping, with broad implications for protein engineering, human genetics, and beyond. To date, however, the overwhelming majority of DMS have focused on substitutions, excluding other classes of variation such as deletions. In order to further our understanding of proteinsâ€™ (in)tolerance to deletions, we leveraged the ESM-1b model, a transformer-based protein language model, to analyze every possible single amino acid deletion in the human proteome. We find that ESM-1b is a highly accurate predictor of clinical annotations and deletion scanning experiments, making it a suitable model for proteome-wide deletion scanning. Moreover, we discovered approximately 300,000 sites in the human proteome that display opposing behavior for deletions versus substitutions, i.e. a site that can tolerate being deleted but not substituted, or vice versa. Sites that are both deletion intolerant and substitution tolerant are disproportionately found within an alpha helix. We reason this is a consequence of the periodic nature of the structure, which completes a full turn every 3.6 residues; substitutions tend to preserve this periodicity, while deletions tend to disrupt it. Alternatively, sites that are simultaneously deletion intolerant and substitution tolerant show a strong tendency to be found within homogenous sequence contexts. We argue these regions can withstand deletions due to the proximity of similar residues, but substitutions that disrupt the homogeneity will be damaging. Overall, our results underscore the importance of considering deletions alongside substitutions in DMS studies.

129 - Discovering genotype-phenotype relationships with machine learning and the Visual Physiology Opsin Database (VPOD)

Seth A Frazer

Seth A Frazer, Todd H Oakley, Mahdi Baghbanzadeh, Ali Rahnavard, Keith A Crandall

UC Santa Barbara, -, -, UC Santa Barbara, Computational Biology Institute, Department of Biostatistics and Bioinformatics, Milken Institute School of Public Health, The George Washington University, Computational Biology Institute, Department of Biostatistics and Bioinformatics, Milken Institute School of Public Health, The George Washington University, Computational Biology Institute, Department of Biostatistics and Bioinformatics, Milken Institute School of Public Health, The George Washington University, Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution

Predicting phenotypes from genetic variation is foundational for fields as diverse as bioengineering and global change biology, highlighting the importance of efficient methods to predict gene functions. Linking genetic and phenotypic changes has been a goal of decades of experimental work, especially for some model gene families including light-sensitive opsin proteins. Opsins can be expressed in vitro to measure light absorption parameters, including Î»max - the wavelength of maximum absorbance - which strongly affects

organismal phenotypes like color vision. Despite extensive research on opsins, the data remain dispersed, uncompiled, and often challenging to access, thereby precluding systematic and comprehensive analyses of the intricate relationships between genotype and phenotype. Here, we report a newly compiled database of all heterologously expressed opsin genes with Î»max phenotypes called the Visual Physiology Opsin Database (VPOD). VPOD_1.0 contains 864 unique opsin genotypes and corresponding Î»max phenotypes collected across all animals from 73 separate publications. We use VPOD data and deepBreaks to show regression-based machine learning (ML) models reliably predict Î»max, account for non-additive effects of mutations on function, and identify functionally critical amino acid sites. The ability to reliably predict functions from sequences alone using ML will allow robust exploration of molecular-evolutionary patterns governing phenotype, inform functional and evolutionary connections to an organismâ€™s ecological niche, and may be used broadly for de-novo protein design. Together, our database, phenotype predictions, and model comparisons lay the groundwork for future research applicable to families of genes with quantifiable and comparable phenotypes.

158 - Novel approaches for detecting ghost admixture

Martin Kuhlwilm

Xin Huang, Josef Hackl, Konstantina Cheshmedzhieva, Aigerim Rymbekova, Martin Kuhlwilm
University of Vienna, University of Vienna, University of Vienna, University of Vienna, University of Vienna

Many populations experienced a complex history of admixture. In present-day humans, the availability of thousands of genomes allows to detect introgressed fragments with high confidence. However, gene flow has been suggested from lineages other than Neanderthals and Denisovans into humans, and ghost admixture has been found in our closest living relatives, the great apes. A comparative analysis has not been possible yet due to limits of the detection methods, which have often been tailored to the admixture history of modern humans, or even specific datasets. However, demographic histories and available datasets differ vastly between populations and species, introducing the need for more comprehensive tools.

Here, we present a framework to detect introgressed fragments from unknown source populations based on previously applied strategies, using summary statistics or densities of private variants, as well as supervised machine learning algorithms. Furthermore, we implement different machine learning architectures, improving the performance specifically in datasets which substantially differ from present-day humans from Eurasia. In particular, this includes smaller sample sizes (on the order of tens of individuals), unphased data, and approximate instead of fine-tuned demographic models. A modularized framework allows versatile treatment of input data, and flexible handling of different datasets.

Under this framework, it becomes possible to choose the most appropriate strategy for a given dataset, evaluate the reliability of the results based on built-in simulation functions, and generate fine-grained maps of introgressed fragments. This will open new avenues into the exploration of ghost introgression genomic landscapes.

206 - Illuminating the functional landscape of the dark proteome across the Animal Tree of Life through natural language processing models

Gemma Martínez-Redondo

Gemma I MartÃ-nez-Redondo, Israel Barrios-NÃºÃ±ez, MarÃ§al VÃ¡zquez-Valls, Patricia Medina-Burgos, Idelfonso Cases, Ana M Rojas, Rosa FernÃ¡ndez
Institute of Evolutionary Biology (CSIC-UPF), Andalusian Center for Developmental Biology (CABD-CSIC), InsInstitute of Evolutionary Biology (CSIC-UPF), Andalusian Center for Developmental Biology (CABD-CSIC), Andalusian Center for Developmental Biology (CABD-CSIC), Andalusian Center for Developmental Biology (CABD-CSIC), InsInstitute of Evolutionary Biology (CSIC-UPF)

Understanding how coding genes and their functions evolve is a key aspect of evolutionary biology. Protein-coding genes poorly characterized at the functional level may be related to important evolutionary innovations, potentially leading to incomplete or inaccurate evolutionary change models, and limiting the identification of conserved or lineage-specific features. Traditionally, functions are assigned based on homology against model organisms, often failing in a large fraction of the coding gene repertoire in non-model species. This is particularly relevant in animals, where many coding genes yield no functional annotation. Here, we leverage machine learning and natural language processing-based protein models to investigate 1) their performance in model organisms (where there are proteins with unknown functions), and 2) the functional annotation in the â€˜dark proteomeâ€™ (â€˜unknown functional landscapeâ€™) of ca. 1,000 gene repertoires of virtually all animal phyla, totaling ca. 23.2 million coding genes. We found that protein language models are the most informative, being able to transfer gene ontology annotations to virtually all genes. Moreover, our exploration of the animal â€˜dark proteomeâ€™ revealed enriched functions related to immunity, viral infection, response to stimuli, development, or signaling. Furthermore, we present an open-access pipeline -FANTASIA- to implement and benchmark these methodologies in any dataset. Our results highlight the use of protein language models for large-scale functional annotation that uncovers the putative functions of poorly understood protein-coding genes across the Animal Tree of Life, contributing to a more comprehensive understanding of the molecular basis of animal evolution.

234 - QuartetNet: Novel Phylogenetic Quartet Tree Reconstruction Using Neural Networks

Trong Nhan Ly

Zixin Zhuang, Nhan Ly-Trong, Minh Bui
Australian National University, Australian National University, Australian National University

Phylogenetic inference is a fundamental biological problem that explores the evolutionary relationships among living organisms on Earth. Phylogenetic methods typically take a sequence alignment of species (taxa) as the input to reconstruct an evolutionary tree. While machine learning has been applied to phylogenetics since around 2018, these studies are still in their early stages, typically reconstructing small trees with a few taxa (e.g., four) and facing limitations - only work with gap-free alignments with a fixed number of sites, can predict either tree topology or branch lengths but not both. This study proposes QuartetNet, which overcomes these limitations. QuartetNet can reconstruct a quartet (4-taxon) tree with branch lengths from an alignment (with/without gaps) with an arbitrary number of sites using neural networks (NNs) . Notably, QuartetNet outperforms the widely used maximum likelihood-based IQ-TREE software in both speed and accuracy, being 20 times faster and slightly more accurate in simulations and empirical data.

382 - CNNs for selection inference: what cues do they take from us, and what can we learn from them?

Lauren Alpert Sugden

Lauren A Sugden
Duquesne University

As machine learning tools prove to be more and more powerful for inference in population genetics, we need greater transparency and tools for interpretation that can help us understand why these models work the way they do. In part, this is important as a defense against potential artifacts and other robustness concerns. But perhaps more importantly, by trying to understand these sophisticated models, we are presented with a valuable opportunity to learn from them in ways that could potentially enrich our own theoretical understanding of evolutionary processes. To this end, we examine the use of convolutional neural networks (CNNs) for identifying selective sweeps as a case study. Using tools from the field of machine learning explainability, together with the deep well of theory surrounding selective

sweeps, we examine what patterns the CNN is learning from raw haplotype images, how these compare with known signatures, how much model complexity is really necessary for this and related tasks, and how the model's learning strategy can be swayed by changes in upstream data preprocessing decisions.

401 - Detecting Introgressions from Hybrids using Evolutionary Sparse Learning

Alessandra Pavan Lamarca

Alessandra P. Lamarca, Sudhir Kumar

Universidade Federal do Rio de Janeiro, Temple University

Recently evolved species, recognized as distinct evolutionary units, may experience genetic introgression from hybrid individuals. This introgression complicates population management and conservation efforts, particularly when hybrids present a continuum of genetic mixtures between parent species. Traditional genotype analyses, commonly used to detect such individuals, rely on unrealistic assumptions of Hardy-Weinberg and linkage equilibrium. They also become computationally challenging for large datasets. We have developed an Evolutionary Sparse Learning (ESL-hybrids) approach to classify specimens and identify species hybrids automatically. ESL-hybrids is a machine learning method that does not require pre-training or any assumptions and models needed by traditional genotyping analysis. To demonstrate its effectiveness, we applied ESL-hybrids to identify hybrids within the Ibis genus Plegadis, comparing its performance to a classical genotype-based method for hybrid identification. We found ESL-hybrids to produce similar results in a fraction of the time used my traditional methods. We also introduce two novel metrics, Î£SPS and Î”SPS, to detect the hybrid status of any new sequence. These results demonstrate ESL's ability to construct a molecular classification model with significantly reduced computational demands compared to traditional methods, offering a practical and efficient approach to differentiate species and their hybrids.

423 - Ecolocator: A supervised machine learning model for location and climate-of-origin prediction

Jordan Rodriguez

Jordan Rodriguez, Richard Cronn, Andrew Kern

University of Oregon, University of Oregon, United States Forest Service, University of Oregon

As the Earthâ€™s climate changes, nearly every organism on the planet will shift its geographical range in response. For ecologically and economically important species, it is vital that scientists and managers be able to make informed decisions about future suitable environments for populations. In previous work, we have shown that supervised machine learning methods can predict origin from genomic variation using georeferenced training data. However, the spatial genetic variation signal that we are capturing may be the result of both neutral processes and local adaptation. Reasoning that the connection between genotype and environment mediated by local adaptation might be discernable from neutral variation, we introduce Ecolocator, a deep learning method that jointly predicts location of origin as well as environmental variables of that location. Ecolocator trains on a dataset of genotypes with known geographic location and associated climate variables such that it can predict the climate-of-origin and geographic location based on the genotype of an unlabeled sample. To validate our approach, we applied our method to one of the most important timber species in the world, Pseudotsuga menziesii var. menziesii, which has been shown to have strong local adaptation. Using a dataset of georeferenced genotypes and associated climate variables for those locations, we are able to perform accurate joint prediction of location and climate. Our open-source tool clarifies the biological implications of the connection between environmental variation and genetic variation so that informed conclusions regarding the conservation, adaptation, and resiliency of this species can be made.

494 - Identifying heterogeneous genome-wide signatures of speciation using deep learning

Raya Esplin

Raya Esplin, Arun Sethuraman

San Diego State University, San Diego State University

A standing question in evolutionary biology is what happens during speciation across heterogenous parts of the genome? Genomic drivers of speciation have been identified in numerous species complexes - from BDM incompatibilities, genomic "islands", meiotic drives, gene conversion, linked natural selection and genetic hitchhiking, to stochasticity due to drift. However, computational methods to infer genomic drivers of speciation fail to account for the heterogeneity in demography and selection landscapes across the genome, while methods that utilize summary statistics (e.g. Fst - differentiation, Dxy - divergence outliers) are prone to different histories resulting in similar summary statistic distributions. To address these issues, we use forward time simulations using SLiM to generate whole genomes with heterogeneous signatures across the genome (i.e. different processes across different loci) under various speciation scenarios - allopatry, sympatry, speciation with gene flow, delayed cessation of gene flow post divergence, and secondary contact. We then use these data to train a deep-learning framework to predict evolutionary history at each locus. We then apply the training framework to predict evolutionary processes in the i5K, AADR, and 1KGP biobanks.

895 - "A Machine Learning Approach to Predicting Tumor Status through Telomere Length Variation (TLV) Analysis"

Priyanshi P Shah

Priyanshi Shah, Arun Sethuraman

San Diego State University, San Diego State University

Telomeres, the protective caps at the ends of chromosomes, consist of a repetitive nucleotide sequence (TTAGGG) in humans and play a crucial role in chromosomal replication. With each cell division, telomeres shorten due to the end-replication problem. Telomere shortening is therefore associated with age-related
diseases, early mortality, and has implications for cancer biology, particularly in the context of cellular senescence and tumor progression. Our preliminary analyses of Telomere Length Variation (TLV) in the 1000 Genomes Project data indicate significant variation in Telomere Length (TL) among human populations, between sexes, and age classes. A recent study (Burren et al., 2023) also performed a large-scale estimation of telomere length variation across the UK BioBank data and identified several associated rare variants that correlate with myeloid cancer precursors. Here we use data from The Cancer Genome Atlas (TCGA) in congruence with variants identified by Burren et al., 2023
to achieve three primary objectives: 1. Quantify TLV across different cancer types and tissues. 2. Identify genetic variants associated with telomere length variation, spanning both cancer and non-
cancer genomes. 3. Utilize estimated telomere lengths from the whole-genome sequencing (WGS) data, rare variant data, and phenotypic data available on TCGA to train a supervised machine learning model that predicts tumor status (cancer versus non-cancer).

937 - Graph embeddings for population genetics data visualization

Micaela Long Grosso

Micaela Long Grosso, MatÃ-as Reyes

Centro Interdisciplinario en Ciencia de Datos y Aprendizaje AutomÃ¡tico - Universidad de la RepÃºblica, Instituto de MatemÃ¡tica y

EstadÃ-stica - Facultad de IngenierÃ-a - Universidad de la RepÃºblica, Facultad de Ciencias - Universidad de la RepÃºblica

One of the challenges in population genetics data analysis is high dimensionality. Dimensionality reduction via principal components analysis (PCA) is commonly used for visualization of genetic variation and is usually used to make inferences on population structure in combination with other methods such as ADMIXTURE. However, PCA focuses on capturing global linear relationships in the data, and as local variation is often found on higher principal components, a two-dimensional plot does not always capture all the relevant genetic variation.

Non-linear methods such as t-distributed Stochastic Neighbor Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP) have become very popular as visualization and dimensionality reduction tools. Although these algorithms emphasize local patterns and complex relationships, they usually fail at preserving global structure, because distances between clusters are not always meaningful and visualizations are usually difficult to interpret.

In this work, we explore the performance of Node2Vec, an embedding algorithm that generates vector representations of nodes in a graph. Node2Vec learns low-dimensional representations for nodes, based on random walks in the graph, following the idea that nodes on a graph can be treated like sentences in a corpus.

Relative to the other algorithms, Node2Vec appears to better preserve global geometry, showing results that are similar to those obtained with PCA, with an improvement in clustering groups with small sample sizes.

1067 - A Deep Learning approach for HLA typing on ancient DNA data.

Alan Vladimir Godinez Plascencia

Alan Vladimir Godinez Plascencia, Andres Rivera Ramirez, Adriana Michel Avila Garcia, Carlos-Francisco Mendez-Cruz, Federico Sanchez Quinto

International Laboratory for Human Genome Research, UNAM, Center for Genomic Sciences, UNAM, Faculty of Sciences, UNAM, Center for Genomic Sciences, UNAM, International Laboratory for Human Genome Research, UNAM

The allelic characterization of the Human Leukocyte Antigen system (HLA typing) has proven crucial to deepen our understanding of immunological function, population diversity and susceptibility to disease. Recently, computational methods have surged as a cost-effective strategy to carry out HLA typing for clinical and research purposes. Furthermore, novel advancements in ancient DNA (aDNA) recovery and sequencing techniques have uncovered important questions regarding the evolution of the HLA system, allowing researchers to potentially take a glimpse into the immunogenic makeup of individuals from the past. Nevertheless, the application of computational HLA typing on ancient samples still poses a significant challenge due to intrinsic properties of aDNA; such as being highly fragmented, the presence of cytosine deamination at the ends of sequencing reads, and low levels of endogenous DNA present.

To address these difficulties, we present a Deep Learning architecture tailored for conducting HLA typing on aDNA data, which aims to predict database-derived labels corresponding to HLA haplotypes directly from query sequences in a mapping-agnostic manner. This strategy takes advantage of the vast amount of HLA sequence information that is broadly accessible, as well as the ability to incorporate damage patterns akin to those present in aDNA for the generation of a comprehensive training dataset. Our approximation fills the need of a fully-automated way to perform HLA typing that takes into consideration aDNA and its caveats. Overall, this work represents a significant step towards enhancing the suitability of computational methods to paleogenomics, facilitating greater insights into human evolutionary history.

1128 - Can AI Outperform Well-Established Antimicrobial Resistance Gene Databases such as ResFinder and CARD?

Kristiene Bianca Recto

Kristiene Bianca Recto, Pleuni Pennings
San Francisco State University, San Francisco State University

The rise in antimicrobial resistance genes diminishes antibiotics' ability to effectively fight harmful microorganisms. To combat infections and customize effective treatments, itâ€™s essential to know whether an infection is resistant to a certain drug. Machine learning (ML) combined with genomic data is a promising tool in this field, yet achieving the desired higher accuracy values remains a challenge. Escherichia coli (E coli) is a diverse species that can be categorized into phylogroups. Itâ€™s unclear how these phylogroups influence prediction accuracies in drug resistance. We aim to bridge this gap by investigating the relationship between phylogroups and ML prediction accuracies in hopes of improving the ability to combat drug-resistant infections.

Using an openly accessible dataset from the UK, we evaluate the prediction accuracies of Gradient Boosted Tree (XGBoost) ML model against ResFinder and Comprehensive Antibiotic Resistance Database (CARD). XGBoost combines predictions of multiple decision trees, with each tree correcting the errors of the previous tree to improve overall predictive accuracy. ResFinder is a database that identifies acquired genes and chromosomal mutations associated with antimicrobial resistance. CARD offers analysis tools such as BLAST and the Resistance Gene Identifier software to predict resistomes. We compare these models across four E. coli Phylogroups (A, B1, B2, D/F) and six antibiotics to determine whether ML or the traditional CARD/Resfinder approach yield better results for each drug-phylogroup combination. Our project is dedicated to enhancing ML methods for quickly and accurately predicting antibiotic resistance. This advancement holds the potential to improve patient health outcomes significantly.

92 - Long-liveness in Aves: Convergent Evolution Supports a More Conserved Cell Cycle and Division Compared to Short-lived Birds

Mirko Martini

Mirko Martini, Mariangela Iannello, Giovanni Piccinini, Liliana Milani
University of Bologna, University of Bologna, University of Bologna, University of Bologna

Ageing is the unavoidable process of progressive deterioration of physiological functions of any biological object, which leads to a higher incidence of diseases and death. Birds have proven to be valuable new models for longevity studies thanks to their remarkable lifespans and low rates of ageing. Employing comparative genomic analyses, this work aims to increase the knowledge about the shared genetic mechanisms at the base of longevity in vertebrates, focusing on the class Aves. Through the implementation of three longevity indexesâ€”LQ, p-PCA, and MLSâ€, we first tried to most sensibly group long- and short-lived bird species. Signals of convergent evolution were searched by applying analyses based on relative evolution rates. Our findings revealed a strong convergent evolution signal in long-lived bird genomes, evident only when using MLS as the longevity index. These convergently evolved genes were enriched in functions involved in chromosome segregation during cellular division. Additionally, comparing interesting genes found in long-lived species to short-lived ones provided insights into the genetic pathways underlying opposite longevity phenotypes. Our results suggest that an efficient cell cycle regulation may underpin the long-liveness phenotype in birds: long-lived birds could have evolved a more accurate cell division machinery, which decreases chromosome instability and aneuploidy onset, offering a potential explanation for their delayed ageing. While this study highlights the utility of comparative genomic approaches, further research is warranted to unveil the precise roles of these genes and their evolutionary trajectories across avian phylogeny.

99 - Clonal Hematopoiesis of Indeterminate Potential and HIV infection synergistically affect

all-cause mortality

Valeriia Timonina

Valeriia Timonina, Konstantin Popadin, Jacques Fellay
Ã‰cole polytechnique fÃ©dÃ©rale de Lausanne, Swiss Institute of Bioinformatics, attendancesmbe2024@gmail.com, 4Immanuel Kant Baltic Federal University, Ã‰cole polytechnique fÃ©dÃ©rale de Lausanne, Swiss Institute of Bioinformatics, Lausanne University Hospital and University of Lausanne

Understanding the mechanisms of normal, healthy aging can be enhanced by studying the cases of accelerated aging. It is well established that both Clonal Hematopoiesis of Indeterminate Potential (CHIP) and human immunodeficiency virus (HIV) are linked to accelerated aging, as evidenced by increased all-cause mortality risks. Previous research has identified a higher prevalence of CHIP in People Living with HIV (PLWH). In this study, we investigated whether CHIP and HIV interact synergistically to exacerbate aging outcomes. Utilizing data from the Swiss HIV Cohort Study of people living with HIV (N = 1889), we found that CHIP carriers exhibit a hazard ratio for all-cause mortality of 1.35, which is in the range of the ratios of 1.18-1.4 in the general population reported in the literature. To corroborate these findings, we analyzed the effect of both HIV and CHIP on all-cause mortality in age- and sex-matched PLWH (N=481) and control individuals (N=2405) from the UK Biobank. Our results revealed a significant interaction between HIV and CHIP (Hazard ratio for HIV = 1.98, p<0.001; Hazard ratio for CHIP = 1.16, p=0.5; Hazard ration for HIV and CHIP interaction = 2.99, p=0.05), further elevating the risk of all-cause mortality. Investigating the underlying mechanisms of an interaction (potentially through increased inflammation) and exploring similar interactions of CHIP with other infections represents a promising avenue for future research on aging.
This research was conducted using the UK Biobank Resource under Application #84415.

507 - Metabolic pathway stimulation alleviates aging pathologies in C. elegans

Serah Maia Osborn

Serah M Osborn, Joseph Dubie, Justin Havird
University of Texas at Austin, University of Texas at Austin, , University of Texas at Austin

The accumulation of mitochondrial damage is one of the primary theories of aging. As an essential metabolic regulator and the principal organelle responsible for oxidative metabolism, it is crucial to understand the role of mitochondria in disease and aging phenotypes.
We tested the effects of chemicals known to increase the lifespan of wild-type Caenorhabditis elegans in strains with different mitochondrial phenotypes and lifespans. The chemicals included rapamycin (mTor signaling pathway inhibitor), carnosol (increases ROS scavenging), and curcumin (increases insulin/IGF-1 signaling).
We reared five strains of C. elegans exposing them to each of the chemicals from birth to death: LB138 (Î”uaDf5) is a wild-type strain with a 3kbp mitochondrial genome deletion shown to exhibit lower mean lifespan and metabolic dysfunction. DA631 (eat-3) has impaired outer mitochondrial membrane fusion and increased lifespan, but decreased health. CB4858 and MQ130 are wild isolates with respectively short and long lifespans relative to the wild-type, reference N2 strain.
We measured lifespan in each strain exposed to each chemical and without any chemical exposure (i.e., 20 treatments). We also approximated worm health throughout the experimental regime by measuring the frequency of longitudinal bends in liquid media.
Results suggest health benefits from exposure to these anti-aging chemicals are unique to the pathology of the strain. This experiment provides insight into the impact of metabolic signaling pathways on mitochondrial function, and potential avenues for treatments of mitochondrial disease.

560 - Longitudinal multi-omics data analysis of cynomolgus macaque genome throughout their lifespan reveals age-related immune patterns

Hyeon-Mu Cho

Hyeon-Mu Cho, Se-Hee Choe, Ja-Rang Lee, Hye-Ri Park, Min-Gyeong Ko, Yun-Jung Lee, Hwal-Yong Lee, Sung Hyun Park, Sang-Je Park, Young-Hyun Kim, Jae-Won Huh

Korea Research Institute of Bioscience and Biotechnology (KRIBB), University of Science & Technology (UST), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea Research Institute of Bioscience and Biotechnology (KRIBB), University of Science & Technology (UST), Korea Research Institute of Bioscience and Biotechnology (KRIBB), University of Science & Technology (UST), Korea Research Institute of Bioscience and Biotechnology (KRIBB), University of Science & Technology (UST), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea Research Institute of Bioscience and Biotechnology (KRIBB), University of Science & Technology (UST),

Despite aging research obtains huge benefits from recent technological advances, its biological mechanism remains unknown. Therefore, we investigated multi-omics data of cynomologus macaques, a laboratory monkey, well protected from environmental confounding factors other than aging. With transcriptomic data, both immune cell composition analysis and gene ontology analysis showed that the genes of innate immune cell such as myeloid cells and NK cells were upregulated while the genes of adaptive immune cell such as Th cells and B cells were downregulated as they get older. We found that the cis-regulatory regions of the upregulated genes were mostly down-methylated in aging according to the whole genome bisulfite sequencing data. Additionally, the correlation analysis indicated that the number of genes with increased expression by down-methylation on cis-regulatory region were twice as many as that with decreased expression by up-methylation. In ATAC sequencing analysis, we observed that chromatin keeps opening with down-methylation on a genome scale. Conversely, we found eight peaks that chromatin was gradually closing while aging. For instance, the peak in the intronic region of TP73 5â€™UTR were gradually closed with its downregulated gene expression. As further analysis is still ongoing, we will put more efforts into the correlation analysis of these multi omics data.

591 - Comparative Analysis of primate genomes to unveil the genetics of maximum lifespan determination.

Fabio Barteri

Fabio Barteri, Noelia Rodriguez Perez, Miguel Ramon Alonso, Alejandro Valenzuela Seba, David De Juan, Arcadi Navarro Cuartiellas

BarcelonaBeta Brain Research Center, Pasqual Maragall Foundation, IBE, Institute of Evolutionary Biology (UPF-CSIC), Department of Medicine and Life Sciences, Universitat Pompeu Fabra., IBE, Institute of Evolutionary Biology (UPF-CSIC), Department of Medicine and Life Sciences, Universitat Pompeu Fabra., IBE, Institute of Evolutionary Biology (UPF-CSIC), Department of Medicine and Life Sciences, Universitat Pompeu Fabra., IBE, Institute of Evolutionary Biology (UPF-CSIC), Department of Medicine and Life Sciences, Universitat Pompeu Fabra., IBE, Institute of Evolutionary Biology (UPF-CSIC), Department of Medicine and Life Sciences, Universitat Pompeu Fabra., IBE, Institute of Evolutionary Biology (UPF-CSIC), Department of Medicine and Life Sciences, Universitat Pompeu Fabra ., BarcelonaBeta Brain Research Center, Pasqual Maragall Foundation., InstituciÃ³ Catalana de Recerca i Estudis AvanÃ§ats (ICREA) and Universitat Pompeu Fabra., Center for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology., Department of Medicine and Life Sciences, IBE, Institute of Evolutionary Biology (UPF-CSIC), Universitat Pompeu Fabra.

Maximum life span (MLS) is a trait that varies widely across primates â€“ from 2 to 60 years when excluding humans â€“ but remains

consistent within each individual group. The genetic architecture of MLS involves numerous genes and regulatory elements that belong to different networks, many of which related to senescence and/or the onset of age-related diseases. Understanding the genetic background of MLS can provide insights into various aspects of the evolution of aging and may suggest interventions upon lifespan and healthspan in humans.

Here, we will discuss our comparative study of MLS within the primate phylogeny leveraging an unprecedented dataset formed by the genomes of primate species that was recently released by Kuderna et al. (2023). We analyzed 16,133 protein-coding regions across the primate phylogeny, identifying genes and even individual Convergent Amino Acid Substitutions (CAAS) associated with the variation of Longevity Quotient (LQ, MLS normalized for body mass). Additionally, we investigated genes whose rates of protein evolution correlate with LQ.

Our preliminary results have returned a set of 2,191 genes, the 11.6% of which (255) had been already retrieved amongst the 996 genes associated with longevity in mammals identified by FarrÃ© et al. in 2021. These genes are involved in pathways that have been associated with male infertility, chronic respiratory conditions and degenerative disorders, such as Alzheimerâ€™s Disease and Huntington's Disease.

Our work represents an opportunity to exploit the potential of comparative genomics approaches to identify the genetic background of aging.

715 - Don't grow old alone: Horizontal transfer of transposable elements from the microbiome during vertebrate host senescence

Sergio A González-Mollinedo

Sergio A GonzÃ¡lez-Mollinedo, Chuen Lee, Terry Burke, Jan Komdeur, David Richardson, Hannah Dugdale

University of Groningen, University of East Anglia, University of Sheffield, University of Groningen, University of East Anglia, University of Groningen

Senescence - the accelerating loss of fitness during later life - continues to spark debate ever since it was proposed as being an evolved trait. The search for explanations for the evolution of senescence have revealed interesting mechanistic explanations of why it occurs, feeding the development of new theoretical frameworks. Recent findings suggest that gene expression dysregulation associated with loss of heterochromatin control in late life, causing a decline in cellular and organismal function, may be a mechanistic pathway . This loss of control causes a spike of transposable element (TE) expression at the onset of senescence, since heterochromatin regions are rich in TEs. Also, it has been suggested that TEs from the microbiome can cross kingdom barriers and affect host cells, potentially causing senescence. In our study on the Seychelles warbler, we find increasing levels of bacterial transposases in the gut microbiome with age, coinciding with the onset of reproductive senescence. We identified these transposases in individual genomes and in the reference genome. These transposases belong to endogenous retroviral elements (ERVs), suggesting that horizontal transfer has occurred from the microbiome to the germline of the host through ERVs. The increase of transposases in the microbiome during senescence suggests that the microbiome can contribute ERV particles that can cause inflammation and senescence, and/or serve as repository for the ERVs that are expressed in the host. Our findings suggest a role for the microbiome in senescence and contributes to a pangenomic approach to our understanding of this phenomenon.

959 - Insights into Mitochondrial Dynamics in early Human Embryonic Development: Enhancing Preimplantation Genetic Testing for Aneuploidy with MODICUM

Evgenii O. Tretiakov

Evgenii O Tretiakov, Ilya Mazunin, Konstantin Popadin
Medical University of Vienna, Center for Molecular and Cellular Biology, Skolkovo Institute of Science and Technology, ICARM (Interdisciplinary Clinical Association for Reproductive Medicine), ICARM (Interdisciplinary Clinical Association for Reproductive Medicine), Center for Mitochondrial Functional Genomics, Institute of Living Systems, Immanuel Kant Baltic Federal University, School of Life Sciences, Eì▪cole Polytechnique Feì▪deì▪rale de Lausanne

Mitochondrial integrity is a crucial determinant of embryonic development and its potential to progress to a successful pregnancy. Traditional methods for evaluating embryo viability before in vitro fertilization (IVF) are limited by a lack of precision. Addressing this gap, our study integrates concepts of cell competition, an evolutionary quality-control mechanism, and developmental canalization, which ensures robust development under genetic and environmental perturbations. We investigated the role of mitochondrial dynamics in embryonic fitness by analyzing the mito-nuclear ratio in 15,520 human blastocysts. The MitOchonDrIa-based Cell tUrnover tiMe (MODICUM) was introduced as an innovative metric, combining mito-nuclear ratio and embryonic age at biopsy to estimate cell turnover timeâ€”a marker of developmental competence. Our findings demonstrate a strong association between low MODICUM values, superior morphological scores, and increased likelihood of live birth outcomes in retrospective clinical trial across several cohorts of different age. The most prominent results could be achieved for cohort of age above 40 years old, which also incorporate a higher number of non-recommended for transfer embryos. MODICUM's robustness in predicting embryo viability underscores the evolutionary advantage of eliminating cells with suboptimal mitochondrial function through cell competition, contributing to developmental canalization. This study advances the field of reproductive medicine by offering a scientifically grounded, evolutionary-informed method for IVF embryo selection, enhancing the success rates of assisted reproductive technologies.

980 - Do Methylation Clocks Generalize Across Admixed Populations?

Sebastián Cruz-González

SebastiÃ¡n Cruz-GonzÃ¡lez, Esther Gu, Lissette Gomez, Makaela A. Mews, Anthony J. Griswold, William S. Bush, John A. Capra
University of California, San Francisco, Biological and Medical Informatics Graduate Program, UCSF, University of Miami Miller School of Medicine, University of Miami Miller School of Medicine, Case Western Reserve University, System Biology and Bioinformatics, Department of Nutrition, Case Western Reserve University School of Medicine, University of Miami Miller School of Medicine, Case Western Reserve University, University of California, San Francisco

Biological aging is the progressive accumulation of cellular damage leading to degeneration and eventual organismal death. Unlike chronological age, biological age varies among individuals due to factors including genetics, smoking, and diet. Methylation clocks, which link methylation levels at CpG sites across the genome to chronological age, have emerged as potential biomarkers for biological age. Accelerated biological aging, as measured by methylation clocks, is significantly associated with an increased risk of age-related ailments, such as Alzheimerâ€™s disease (AD) and coronary heart disease.
Despite this promise, existing clocks have been trained and evaluated in predominantly European ancestry cohorts and lack validation in genetically diverse populations, hindering their broader applicability.
We evaluate the accuracy of several common methylation clocks and their ability to serve as biomarkers of Alzheimerâ€™s disease risk in admixed populations from the Americas. We leverage the MAGENTA study, a large cohort of 313 diverse admixed AD patients and 308 control individuals with methylation data and genotyping data. We generate biological age predictions from the methylome for all individuals and compare their biological age to their chronological age and test the association of accelerated aging with AD status. Furthermore, we evaluate the accuracy of the methylation clocks on admixed individuals with varying levels of global European, Amerindigenous, and African ancestry.

This study addresses the lack of diversity in the development and application of genomic tools for precision medicine and assesses the utility of methylation clocks as biomarkers for age-related diseases across diverse populations.

1033 - Cell type specific age-dependent changes in prefrontal cortex of tame and aggressive foxes

Anna V. Kukekova

Jennifer L Johnson, Erin E Hecht, Darya V Shepeleva, Anastasiya V Kharlamova, Anastasiya V Vladimirova, Rimma G Gulevich, Lyudmila N Trut, Rando M Rando, Anna V Kukekova

University of Illinois, Harvard University, Harvard University, Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Smith College, University of Illinois

Geriatric changes are often thought not to be under direct selection because in nature mortality typically precedes senescence. Genetic variants that facilitate selection for specific behaviors in young animals might therefore have trade-offs in later life. We used single nucleus RNA-seq to dissect cell-type-specific cortex differences between six young and six old red foxes from populations selected separately for tame and aggressive behavior. Clustering analysis identified 27 clusters, including 12 clusters of excitatory neurons, 9 of inhibitory neurons, and 6 of non-neuronal cell-types. Comparative analysis found stronger differences between ages than between populations. The number of nuclei in a layer 2/3 cluster of intratelencephalic neurons was significantly larger in young than old foxes, and the opposite pattern was observed in a layer 5/6 cluster of cortico-thalamic neurons. The differential expression analysis identified 923 DE genes between populations and 2,648 DE genes between ages, the largest number of DE genes between young and old foxes was found in layer 2/3 neurons. Testing whether genes that are upregulated in young versus old foxes tend to also be upregulated in an expression direction between tame and aggressive foxes revealed significant enrichment for DE genes upregulated in both tame and young foxes. Together with our previous finding of a late onset Alzheimerâ€™s disease gene involved in regulation of tame behavior, these results indicate that genes under selection for specific behaviors may have broader functions in the brain, and selection for behavior may affect these functions.

1058 - Identifying de novo structural variation in the aging male germline using long reads

Stacy Li

Stacy Li, Peter Sudmant

Department of Integrative Biology, UC Berkeley, Center for Computational Biology, UC Berkeley, Department of Integrative Biology, UC Berkeley, Center for Computational Biology, UC Berkeley

Aging is an emergent phenomenon hallmarked by the deterioration of physiological processes over time. Genomic instability occurs throughout normal and pathological aging, typically driven by unresolved or erroneously remediated DNA damage. De novo germline mutations are of critical importance to reproductive success in aging: Previous work using short-read sequencing in pedigree-based studies identified a parental age-associated increase in small de novo variants, with a strong bias from the paternal germline. De novo structural variants (dnSVs) have profound impacts on health, disease, and genome evolution, but remain challenging to characterize due to the technical limitations of short-read sequencing. To address this, we conducted a study using PacBio HiFi long-read sequencing to identify dnSVs in bulk sperm samples from eight donors of varying ages. Under this framework, we anticipate recall of clonal dnSVs (shared amongst sperm descended from a common progenitor), with the potential to recover singleton/doublet dnSVs arising in meiosis. To this end, we developed a method to identify de novo retrotransposition events, which can be characterized within the span of a single

read. We identified multiple high-confidence AluYb8, AluYa5, and L1HS retrotransposition events. We estimate between 2.1 to 10.2 events per 100 individual cells, with the oldest donor sample yielding the greatest number of events. Furthermore, several L1 events originally raised as insertions were revealed to be complex clonal dnSVs following manual examination. These results suggest that long-read sequencing is a promising method to survey dnSVs in the aging germline.

87 - Comparison and optimization of protocols and whole genome capture conditions for ancient DNA samples

Reyhan Yaka

Reyhan Yaka, Vendela Kempe Lagerholm, Maja Krzewinska , Anna Linderholm, FÃ¼sun Ã–zer , Mehmet Somel, Anders Gotherstrom
Centre for Palaeogenetics, Stockholm University, Department of Archaeology and Classical Studies, Stockholm University, Centre for Palaeogenetics, Stockholm University, Centre for Palaeogenetics, Stockholm University, Centre for Palaeogenetics, Stockholm University, Department of Anthropology, Hacettepe University, Department of Biological Sciences, Middle East Technical University, Centre for Palaeogenetics, Stockholm University

Ancient DNA(aDNA) data obtained from archaeological remains is an expanding and well-established field that has already helped resolve various questions in evolutionary biology, archaeology and anthropology. However, aDNA also exhibits significant technical challenges. Over the last decade, new approaches have mitigated these issues. But obtaining sufficient authentic aDNA still remains a challenge for samples from temperate, warm and humid areas. Here, we studied a set of methods for obtaining aDNA data from archaeological samples spanning different sites and periods, and with various levels of preservation. We compared DNA extraction and library preparation protocols, and tested the efficiency of whole-genome enrichment (WGC) on ancient samples by modifying a number of parameter combinations. We find that the Dabney extraction protocol performs significantly better than alternatives, and observe a positive trend with the BEST library protocol indicating lower clonality. Our results further show that WGC is effective at retrieving endogenous DNA, particularly from poorly preserved samples. This is an important finding suggesting that WGC can work exceptionally well on aDNA samples with very low endogenous content and contrasts with previous studies. We also suggest WGC as an alternative technique to the targeted SNP-capture method. This is because WGC has the potential to generate unbiased data, and thereby, could present a viable option in future aDNA research targeting low quality libraries. Overall, our results suggest that WGC is an efficient and cost-effective method for retrieving endogenous DNA in ancient samples, and holds promise to improve new techniques in the field.

312 - An autosomal genome of a Mexican Columbian mammoth informs about its hybridization dynamics with the wooly mammoth in the Americas

Rigoberto Padilla Bustos

Rigoberto Padilla Bustos, Alejandra Castillo Carbajal, Eduardo A Arrieta-Donato, Miriam Bravo LÃ³pez, Ernesto Garfias Morales, Ã•ngeles Tavares GuzmÃ¡n, Viridiana Villa Islas, Alejandro LÃ³pez JÃ-menez, Mashaal Sohail, JoaquÃ-n Arroyo-Cabrales, MarÃ-a C Ã•vila Arcos, Federico SÃ¡nchez Quinto
International Laboratory for Human Genome Research, International Laboratory for Human Genome Research, International Laboratory for Human Genome Research, Centro de Ciencias GenÃ³micas, International Laboratory for Human Genome Research, International Laboratory for Human Genome Research, International Laboratory for Human Genome Research, International Laboratory for Human Genome Research, Laboratorio de ArqueozoologÃ-a, Centro de Ciencias GenÃ³micas, Laboratorio de ArqueozoologÃ-a, International Laboratory for Human Genome Research, International Laboratory for Human Genome Research

The hybridization dynamics between different species have been a long-studied research subject addressed by multiple disciplines, most

recently by paleogenomics. Recent studies suggest that the Columbian mammoth resulted from hybridization between an ancestral lineage of Siberian woolly mammoths and possibly the steppe mammoth (50-50%), followed by subsequent gene flow (~10%) from the woolly mammoth into the Columbian mammoth in North America. However, more genetic data on the Columbian mammoth is needed to better understand the hybridization dynamics between both species in the Americas, particularly since only one Columbian mammoth genome is available to date. Additionally, this genome stems from an individual from Wyoming, at their putative hybridization zone between USA and Canada, further complicating the interpretation. The discovery of several Columbian mammoth remains at the Basin of Mexico provides a unique opportunity to further understand the demographic dynamics between both species in the continent. We screened 72 libraries and generated a >1.5x nuclear genome for one individual. We used this nuclear genome to investigate the hybridization dynamics between both species in the Americas, by comparing it to the Columbian mammoth individual from Wyoming, and other woolly mammoth genomes from North America and Eurasia. Our study provides a deeper understanding of the relationship between both mammoth species in the American continent and paves the way for future projects to continue to tackle mammoth evolutionary history.

313 - Insights into the evolutionary history of Late Pleistocene Mexican camels through mitochondrial DNA phylogenetic analyses

María José Rodríguez Barrera

MarÃ-a JosÃ© RodrÃ-guez-Barrera, Pablo E. Uribe-Herrera, Alejandra Castillo-Carbajal, Viridiana Villa-Islas, Ernesto Garfias-Morales, Eduardo Arrieta-Donato, Miriam Bravo-LÃ³pez, Alejandro LÃ³pez-JimÃ©nez, JoaquÃ-n Arroyo-Cabrales, MarÃ-a Ã•vila-Arcos, Federico SÃ¡nchez-Quinto

Escuela Nacional de Estudios Superiores Juriquilla, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM)., Facultad de Estudios Superiores Iztacala. Universidad Nacional AutÃ³noma de MÃ©xico., Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), Centro de Ciencias GenÃ³micas, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), Laboratorio de ArqueozoologÃ-a, SubdirecciÃ³n de Laboratorios y Apoyo AcadÃ©mico, Instituto Nacional de AntropologÃ-a e Historia, Laboratorio de ArqueozoologÃ-a, SubdirecciÃ³n de Laboratorios y Apoyo AcadÃ©mico, Instituto Nacional de AntropologÃ-a e Historia, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM)

Camelops genus first appears in the fossil record between 4 to 3 million years ago and lived across North America until its extinction ~12,000 years ago at the end of the Pleistocene. Recent paleogenomic studies have shown that mitochondrial genomes from Late Pleistocene camels (Camelops hesternus) from Yukon, Canada, related as a sister clade to the genetic variation of all present-day camel species. This result confirms previous reports of an American origin for the Camelidae family. However, prehistoric camels once lived across all North America, therefore it is not clear to what extent Yukon camel genetic data represents the genetic variation of camels across the continent. The recent discovery of camel paleontological remains in the Basin of Mexico during the construction of the Mexico City International Airport, provides an opportunity to further study evolutionary history of Camelops genus in America with a broader geographical sampling. In this study we extracted DNA, sequenced and analyzed paleogenomic data from seven camel samples from the Basin of Mexico. We identified the molecular sex of these individuals and performed mitochondrial capture enrichment experiments. We recovered four whole mitochondrial genomes >1x coverage, and carried out phylogenetic analysis to test the relationship of Mexican prehistoric camels to both present and ancient Camelops sequences. This study is pivotal in unraveling the evolutionary history of camels

in America by characterizing the genetic diversity and demographic trajectory of prehistoric camels in Mexico.

321 - Taxonomic identification of the prehistoric horses from the Basin of Mexico using ancient mitochondrial DNA

Pablo Esteban Uribe-Herrera

Pablo Esteban Uribe-Herrera, Viridiana Villa-Islas, Alejandra Castillo-Carbajal, Ernesto Garfias-Morales, Eduardo Arrieta-Donato, Miriam Bravo-Lopez, Alejandro` Lopez-Jimenez, Joaquin Arroyo-Cabrales, Maria Avila-Arcos, Federico Sanches-Quinto
Facultad de Estudios Superiores Iztacala, UNAM, Laboratorio de Investigacion Internacional del Genoma Humano, UNAM, Laboratorio de Investigacion Internacional del Genoma Humano, UNAM, Laboratorio de Investigacion Internacional del Genoma Humano, UNAM, Laboratorio de Investigacion Internacional del Genoma Humano, UNAM, Centro de Ciencias Genomicas, UNAM, Laboratorio de Investigacion Internacional del Genoma Humano, UNAM, Laboratorio de Arqueozoologia, INAH, Laboratorio de Arqueozoologia, INAH, Laboratorio de Investigacion Internacional del Genoma Humano, UNAM, Laboratorio de Investigacion Internacional del Genoma Humano, UNAM

North America was the birthplace of the Equidae family and witnessed most of its evolution. Mexico has an extensive fossil record of prehistoric horses up to the Late Pleistocene. Despite this, the exact number of pleistocene equine species in North America, and their phylogenetic position is problematic due to the limitations of morphological-morphometric characterization methods and the complexity of the Equus genus. Cheek-teeth morphology characterization has proven useful for species identification. However, it has shown limited resolution when individuals display traits with similar morphological features. An example of this are the prehistoric horse species (characterized to date) that inhabited Mexico: E. mexicanus, E. conversidens and Harintonhippus francisci. Ancient DNA provides an opportunity to perform high resolution species identification. Recently, 15 individualized remains of prehistoric horses were found at the Santa Lucia site dated to 12-30kya in the Basin of Mexico. In this study, we extracted and sequenced ancient DNA from seven prehistoric horses, and captured-enriched three mitochondrial genomes to >1x coverage. We performed mitochondrial phylogenetic analyses including our samples with other ancient and present-day North American-Beringian horses as references and identified their taxon affiliation either on the Equus or Harintonhippus lineages. This study emphasizes the importance of species identification through ancient DNA, as well as provides a dataset to further understand the diversity and evolutionary trajectory of prehistoric horses in Mexico and North America.

350 - Molecular data from South African species dating to 2 million years ago

Ioannis Patramanis

Ioannis Patramanis, Palesa P Madupe, Claire Koenig, Clement Zanolli, Lauren Schroeder, Fernando Racimo, Jesper V Olsen, Rebecca R Ackermann, Enrico Cappellini
Globe Institute, University of Copenhagen, Globe Institute, University of Copenhagen, Globe Institute, University of Copenhagen, Center for Protein Research, University of Copenhagen, University of Bordeaux, University of Toronto Mississauga, Globe Institute, University of Copenhagen, Center for Protein Research, University of Copenhagen, University of Cape Town, Globe Institute, University of Copenhagen

Although the field of ancient DNA has made significant advancements in deep time recovery, enabling scientists to recover genetic material from fossils up to a 1 million years ago, such successes are largely confined to specimens from regions of the world where DNA tends to be well preserved, such as Siberia and northern Canada. In contrast, the tropical and subtropical climates of Africa and southern Asia, where most ancient hominin diversity is found, present substantial challenges for DNA preservation. Here, by employing tandem-mass spectrometry, we sequence the enamel proteomes of multiple 2-million-year-old hominin and mammalian specimens from

South Africa, offering an alternative for extracting valuable genetic information from areas where ancient DNA recovery is not currently viable. Using the recovered protein sequences, we are able to determine the biological sex of the samples as well as identify within-species variation, unlocking the potential to study sexual dimorphism, population diversity and possible substructure in populations of the deep past. We show that the recovered proteins can also be utilized to generate phylogenetic trees, but demonstrate that phylogenetic resolution depends substantially on the amount of protein sequence available to us. Our approach demonstrates the feasibility of recovering informative genetic sequences from the early Pleistocene and from some of the most challenging preservation conditions in the globe, but also highlights how these molecular data can contribute valuable insights into our understanding of the past.

402 - Molecular dating of Mexican Columbian mammoths mitochondrial genomes suggests one hundred-thousand years presence in the Basin of Mexico

Eduardo Arrieta Donato

Eduardo Arrieta Donato, Angeles Tavares Guzman, Alejandra Castillo Carbajal, Miriam Bravo Lopez, Ernesto Garfias Morales, Viridiana Villa Islas, Vanssy Li, Alejandro Lopez Jimenez, Mashaal Sohail, Camilo Chacon Duque, Joaquin Arroyo Cabrales, Maria Avila Arcos, Federico Sanchez Quinto

Center for Genomic Sciences, National Autonomous University of Mexico, International Laboratory for Human Genome Research, National Autonomous University of Mexico, International Laboratory for Human Genome Research, National Autonomous University of Mexico, International Laboratory for Human Genome Research, National Autonomous University of Mexico, International Laboratory for Human Genome Research, National Autonomous University of Mexico, International Laboratory for Human Genome Research, National Autonomous University of Mexico, Centre for Palaeogenetics, Stockholm University, Laboratorio de Arqueozoologia, Subdireccion de Laboratorios y Apoyo Academico, Instituto Nacional de Antropologia e Historia, Center for Genomic Sciences, National Autonomous University of Mexico, Centre for Palaeogenetics, Stockholm University, Department of Zoology, Stockholm University, Laboratorio de Arqueozoologia, Subdireccion de Laboratorios y Apoyo Academico, Instituto Nacional de Antropologia e Historia, International Laboratory for Human Genome Research, National Autonomous University of Mexico, International Laboratory for Human Genome Research, National Autonomous University of Mexico

The Columbian mammoth was the only mammoth species endemic to the Americas with a geographical range which spanned from northern US to Costa Rica. To date its evolutionary history has only been studied using genetic data from samples across the US. The recent discovery of Columbian mammoth remains at the Basin of Mexico provides an opportunity to further investigate its evolutionary trajectory across the American continent. We extracted DNA, performed captured enrichment experiments and recovered 63 Columbian mammoth mitochondrial genomes, of which 28 of them had more than 10x coverage. Phylogenetic analyses suggest that three different lineages were present in the Basin of Mexico. We performed molecular dating on the 28 high-coverage samples using worldwide carbon-14 dated individuals as references for molecular clock calibration, including a Mexican individual C14-dated to 12ky, in order to obtain a tip dating for each individual. Our findings suggest at least three different temporal occupations across a >100ky time transect in Central Mexico, displaying different genetic diversity estimations. This study spans our understanding of the evolutionary history of the Columbian mammoth in the continent, particularly its demographic dynamics in the Basin of Mexico.

644 - What's (read) size got to do with it?: An assessment of reference bias and spurious mapping

Stephanie Dolenz

Stephanie Dolenz, Tom van der Valk, Chenyu Jin, Jonas Oppenheimer, Muhammad Bilal Sharif, Ludovic Orlando, Beth Shapiro, Love

DalÃ©n, Peter D. Heintzman
Centre for Palaeogenetics, Department of Geological Sciences, Stockholm University, Centre for Palaeogenetics, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Science for Life Laboratory, Centre for Palaeogenetics, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Department of Zoology, Stockholm University, Department of Biomolecular Engineering, University of California Santa Cruz, Centre for Palaeogenetics, Department of Zoology, Stockholm University, Centre for Anthropobiology and Genomics of Toulouse, University Paul Sabatier, Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Howard Hughes Medical Institute, University of California Santa Cruz, Centre for Palaeogenetics, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Department of Zoology, Stockholm University, Centre for Palaeogenetics, Department of Geological Sciences, Stockholm University

The alignment of sequencing reads to a reference genome is a critical step in the characterization of ancient genomes. However, reference bias and spurious mappings pose a significant challenge, particularly as cutting-edge wet lab methods generate datasets that push the boundaries of alignment tools designed for less diverged reads. Reference bias occurs when reference alleles are favored over alternative alleles whilst mapping to the same or a closely related species, whereas spurious mappings stem from either contamination or when endogenous reads fail to align to their correct position of origin. Previous work has shown that the extent of these phenomena is correlated with read length but a more thorough investigation of the nature of reference bias and spurious mappings with ancient DNA mapping parameters has been lacking. Here, we use a range of empirical and simulated palaeogenomic datasets to investigate the impacts of mapping tools, mapping quality thresholds, and reference genome choice, on mismatch rates across all read lengths. We find that mapping algorithms and quality threshold choices dictate reference bias and rates of spurious alignment at different read lengths in a predictable and often conflicting manner, suggesting that optimized mapping parameters for each read length will be a key step in alleviating both reference bias and spurious mappings. This will allow for more robust analyses of faunal palaeogenomic datasets and increased nuance of evolutionary inferences than is currently possible for many extinct taxa.

653 - Thinking about the Roman Empire one more time: Roman sheep genomes and the demographic history of domestication in the western Mediterranean

Marianne Dehasque

Marianne Dehasque, Viviana Floridia, Pedro Morell Miranda, Paolo Mereu, Mario Pirastru, Leoni Giovanni Giuseppe, Salvatore Naitana, NicolÃ² Columbano, Marco Zeddi, Mario Barbato, Torsten GÃ¼nther
Department of Organismal Biology, Uppsala University, Dipartimento di Scienze Veterinarie, UniversitÃ degli Studi di Messina, Department of Organismal Biology, Uppsala University, Dipartimento di Scienze Biomediche, UniversitÃ degli Studi di Sassari, Dipartimento di Scienze Biomediche, UniversitÃ degli Studi di Sassari, Dipartimento di Scienze Biomediche, UniversitÃ degli Studi di Sassari, Dipartimento di Medicina Veterinaria, UniversitÃ degli Studi di Sassari, Dipartimento di Medicina Veterinaria, UniversitÃ degli Studi di Sassari, Dipartimento di Medicina Veterinaria, UniversitÃ degli Studi di Sassari, Dipartimento di Scienze Veterinarie, UniversitÃ degli Studi di Messina, Department of Organismal Biology, Uppsala University

Sheep (Ovis aries) were domesticated in the Fertile Crescent over 10,000 years ago and spread to the rest of the world through multiple domestication and expansion waves. The Roman period, a time characterised by high mobilisation, has been hypothesised to have had a big impact on the genomic make-up of present-day sheep populations. However, to date no sheep genomes from this period have been published. In this study, we generate genome-wide data for two Roman sheep samples from Sardinia and one from Spain and compare these to an extensive dataset consisting of 12 ancient Mediterranean samples and modern genomes from over 30 wild and domestic sheep lineages. We investigate demographic history and the signatures of selection in our western Mediterranean dataset. Preliminary

analysis of the oldest Sardinian sample reveals a unique genetic signature when compared to present-day sheep, suggesting that the history of sheep breeds was even more complex than previously thought. Further analysis will provide important insights into the impact of the Roman period on sheep diversity.

751 - Assessing pre-exploitation baseline numbers and population dynamics of the European sardine Sardina pilchardus Walbaum, 1792 using palaeogenomics

Paula F. Campos

GonÃ§alo E. Themudo, Adolfo F. FernÃ¡ndez, Carlos FernÃ¡ndez-rodriguez, Eduardo GG de Aguero, Maria C del Arco Aguilar, SÃ³nia Gabriel, Rute R da Fonseca, Paula F. Campos

CIIMAR/UP, CIIMAR/UP, University of Vigo, University of Leon, Universidad Leon, Universidad La Laguna, FAUP, The Globe, University of Copenhagen, CIIMAR, University of Porto

Overfishing has been a major problem for several marine species, reducing their effective population sizes to remnants of their pristine levels. However due to the lack of accurate fishery catch data we still have a poor understanding of how these impacted their evolutionary path. One of the most impacted species was the European sardine (Sardina pilchardus), an important pelagic fish resource in Atlantic waters, with enormous economic value especially in Southern Europe and Morocco, where it is the main target of purse-seine fleets and represents a major source of income for local populations. The species is distributed from the southern Celtic Sea and the North Sea to Mauritania and Senegal, including the Azores, Madeira and the Canary Archipelagos, being also abundant in the Mediterranean. Extant populations show three genetic clusters: one including individuals from Azores and Madeira, the second encompassing Iberian populations (the centre of the sampling distribution), and the third gathering the Mediterranean and Canary Islands, with individuals from Iberia showing some degree of admixture. Here we use state of the art ancient DNA techniques to look at samples that pre-date periods of intensive fishing and compare those results to the ones obtained from current populations (post exploitation). Doing this will enable us to get a more accurate perspective of the effective population size and genetic diversity levels of the species prior to overfishing, invaluable information for stock delimitation and management and definition of fishing effort in the different Food and Agricultural Organization (FAO) fishing areas.

219 - Principal Component Analysis in Human Genetics: Debunking False Claims and Recommending Best Practices

Mateus Henrique Gouveia

Daniel Shriner, Mateus H. Gouveia, Amy R. R. Bentley, Karlijn A. C. Meeks, Elisabeth F. Heuston, Ayo P. Doumatey, Guanjie Chen, Charles N. Rotimi, Adebowale A. Adeyemo

National Institutes of Health, National Institutes of Health, National Institutes of Health, National Institutes of Health, National Institutes of Health, National Institutes of Health, National Institutes of Health, National Institutes of Health, National Institutes of Health, National Institutes of Health

Population structure is a pervasive phenomenon in genetic datasets, often resulting from barriers to gene flow that cause populations to genetically differentiate due to non-random mating. Principal Component Analysis (PCA) is a statistical technique widely used in population genetics and genetic epidemiology to reduce the dimensionality of multidimensional data by decorrelating its features. PCA is instrumental in assessing population structure, detecting outliers, and mitigating the effects of confounding due to population stratification. Recently, there have been claims regarding the reliability of PCA in human genetics. Additionally, non-linear techniques such as Uniform Manifold Approximation and Projection (UMAP) have been increasingly utilized to assess population structure, sparking

discussion on their interpretability, especially when used to describe ancestry, ethnicity, and race. Here, we evaluate claims suggesting biases, inconsistencies, and irreproducibility in PCA analysis of genetic data in the context of human genetics and genetic epidemiology. Through an examination of a previously described PCA toy example, we demonstrate that none of these claims are true. Furthermore, we apply PCA and UMAP to simulated data and across different scales of empirical genomic data [HapMap (n=924 individuals), 1000 Genomes (n=2,504), and the All of Us Research Program (n=245,388)], and demonstrate that PCA is valid, consistent, and reproducible. We find that nonlinear techniques such as UMAP can produce artefactual visualization and lack the strong interpretability of linear techniques like PCA, particularly in assessing admixture. In conclusion, while acknowledging the caveats to PCA interpretation, we provide updated recommendations for the best practices governing PCA in genetics.

274 - Evaluating accuracies of forensic analyses across genetic ancestries

Elena I Zavala

Elena I Zavala, Rori Rohlfs, Priya Moorjani
University of California, Berkeley, University of Oregon, University of California, Berkeley

DNA analyses have transformed forensics applications from identification of missing individuals to solving cold cases. In these analyses, profile matching between DNA recovered from a biological sample or piece of evidence and an individual is commonplace. The significance of this match, which is of particular importance for partial or mixture profiles, is typically calculated based on reference allele frequencies of a handful of markers. Recently, investigative leads have been provided using varying qualities of sequencing dataâ€“â€“i.e. coverage, number of sites, etc.â€“â€“ to match individuals or genetic relatives. Given the potential impact of these results on people living today, this raises questions as to how new methods should be tested to understand accuracies across diverse populations. In this study, we discuss the utility of using different statistics to describe individual DNA profiles in place of discrete population identifiers (i.e. heterozygosity and f-statistics). We then evaluate the impact of degraded DNA on genotyping accuracy for forensic genetic analysis across diverse populations using five different SNP panels. The outputs of this study demonstrate the potential for performing human identification with degraded DNA at low coverages (<5X), but highlights critical considerations. In particular, it is important to consider genotype refinement, choice of reference panels, and interpretations of resulting DNA profiles. We discuss the critical need to move beyond conventional population definitions in analysis of diverse individuals for forensic applications.

812 - Connections between topological data analysis and identity-by-descent in biobank data

Alex Diaz-Papkovich

Alex Diaz-Papkovich, Cole Williams, Justin Pelletier, Claude BhÃ©rer, Simon Gravel, Sohini Ramachandran
Brown University, Brown University, McGill University, McGill University, McGill University, Brown University

Population structure reflects shared population histories. Identifying structured subgroups in biobanks is key to exploratory analyses, visualization, and downstream study of biomedical, demographic, or geographic relationships. One approachâ€”identity-by-descent (IBD) clusteringâ€”detects chromosomal segments that are shared identical-by-descent between at least two individuals, and finds communities in networks of shared IBD. Another is density clustering on variant data that has been dimensionally-reduced. While numerous dimensionality reduction approaches are applied to biobanks, we focus on uniform manifold approximation and projection (UMAP), which preserves high-dimensional topology of data. IBD clustering and UMAP do not rely on social/overly coarse labels such as continent, race, or ethnicity to generate clusters.

In past applications, we have found that UMAP-identified subgroups correlate strongly with demographic history, environmental, and

phenotypic data and identify fine-scale clusters similar to IBD. However, no explicit connection has been drawn between the approaches. Such connections can inform researchers of relationships between coalescence and clustering, aid in interpretation of non-linear dimensionality reduction, and provide alternative computationally cheap tools to study clusters of structure in large biobanks.

We study the relationships between these approaches and measures such as time to most recent common ancestor (TMRCA). Using 1000 Genomes Project, CARTaGENE, UK biobank, and Allofus data, we also study relationships between population structure clusters and distributions of environmental variables and geographic data. We present different approaches to visualization, offer best practices on using and visualizing clusters via different methods, and highlight the impact of data processing and algorithm parametrization.

90 - Analysis of meiotic recombination in Drosophila simulans shows heterozygous inversions do not cause an interchromosomal effect

Nicole Crown

Bowen Man, Elizabeth Kim, Alekhya Vadlakonda, David Stern, Nicole Crown
Case Western Reserve University, Janelia Research Campus, Case Western Reserve University, Janelia Research Campus, Case Western Reserve University

Chromosome inversions are of unique importance in the evolution of genomes and species because when heterozygous with a standard arrangement chromosome, they suppress meiotic crossovers within the inversion. In Drosophila species, heterozygous inversions also cause the interchromosomal effect, whereby the presence of a heterozygous inversion induces a dramatic increase in CO frequencies in the remainder of the genome within a single meiosis. To date, the interchromosomal effect has been studied exclusively in species that also have high frequencies of inversions in wild populations. We took advantage of a recently developed approach for generating inversions in Drosophila simulans, a species that does not have inversions in wild populations, to ask if there is an interchromosomal effect. We used the existing chromosome 3R balancer and generated a new chromosome 2L balancer to assay for the interchromosomal effect genetically and cytologically. We found no evidence of an interchromosomal effect in D. simulans. To gain insight into the underlying mechanistic reasons, we qualitatively analyzed the relationship between meiotic double-strand break formation and synaptonemal complex assembly. We find that the synaptonemal complex is assembled prior to double-strand break formation as in D. melanogaster; however, we show that the synaptonemal complex is assembled prior to localization of the oocyte determination factor Orb, whereas in D. melanogaster, synaptonemal complex formation does not begin until Orb is localized. Together, our data show heterozygous inversions in D. simulans do not induce an interchromosomal effect and there are differences in the developmental programming of early stages of meiosis.

143 - Role of aneuploidy in the de novo evolution of macroscopic size in Saccharomyces cerevisiae

Sayantan Datta

Sayantan Datta, Kai Tong, Dung T Lac, Gonensin O Bozdag, William C Ratcliff
Georgia Institute of Technology, Georgia Institute of Technology, Boston University, Georgia Institute of Technology, Georgia Institute of Technology, Georgia Institute of Technology

The evolution of genetic architecture has been crucial in the origin of complex multicellularity. However, due to the lack of transitional model systems, our knowledge of how genome structure changes throughout multicellular adaptation is limited. The Multicellularity Long Term Evolution Experiment (MuLTEE) using the snowflake yeast under different metabolic conditions (aerobic, mixotrophic, and anaerobic) provides a unique system to study evolutionary transitions to multicellularity. Populations growing under anaerobic conditions

evolved macroscopic size after 600 days of laboratory selection, paralleled by whole genome duplication and extensive aneuploidies across all five replicate populations. Here, we investigate the role of this aneuploid genomic architecture in the evolution and maintenance of macroscopic size in the MuLTEE. First, we demonstrate that the phenotypic switch from macroscopic to microscopic multicellularity can be detected by distinctive colony morphology on agar plates. Our microscopy analysis shows that the loss of the macroscopic phenotype is associated with changes in cell-level phenotypes, such as reduced cell aspect ratio, cell volume, or both, confirming that cell-level adaptations drive group-level changes. Next, by sequencing the microscopic and macroscopic cluster genomes, we detect karyotype differences between these strains. Further analysis suggests that the phenotypic switch in size during the MuLTEE is associated with the gain or loss of chromosomes featuring specific mutation profiles, stressing the significance of certain de novo variants in this evolutionary process. Overall, our results suggest that a combined role of aneuploidy and novel mutations drives the evolution and maintenance of macroscopic multicellularity in snowflake yeast.

160 - Dissecting the structural and regulatory genomic basis of hypertrophied lip adaptations in Neotropical cichlid fishes

Paolo Franchini

Paolo Franchini, Darrin C Hulsey, Frederico Henning, Gonzalo Machado-Schiaffino, Bai Wei Lo, Ming Li, Axel Meyer
Department of Ecological and Biological Sciences, University of Tuscia, Department of Biology, University of Konstanz, School of Biology and Environmental Science, University College Dublin, Department of Genetics, Institute of Biology, Federal University of Rio de Janeiro, Department of Biology, University of Konstanz, Max Planck Institute for Molecular Genetics, Department of Biology, University of Konstanz, Department of Biology, University of Konstanz

Cichlid fishes offer powerful models to understand the molecular mechanisms that generate biodiversity. In cichlids, the repeated evolution of novel traits provides a framework to study the genomic basis of adaptation. For instance, hypertrophied lips have convergently arisen in several adaptive radiations, including the Neotropical Midas cichlid species complex (Amphilophus cf. citrinellus), where a functional trade-off in feeding behaviour between thick- and normal-lipped ecotypes is likely fueling divergence through disruptive selection. In this study, we combined analyses of QTL, pangenomic, transcriptomic and chromatin interactions to characterize the molecular genetic bases of the hypertrophied lip phenotype in the Midas cichlids. Our findings reveal that more loci than previously thought contribute to variation in this trophic trait, including two loci with large effect size. Additionally, we identified several genes that show differential expression (DE) between fishes with thick and normal lips and link this to associated DE in miRNAs. Moreover, pangenome reconstructions resolved several ecomorph-specific structural variants that co-localize with lip QTL intervals. Lastly, topological associating domain (TAD) analysis revealed altered TADs in those genomic regions harboring both structural variants and QTL. Collectively, our results suggest that variation in lip size â€“ a possible driver of sympatric speciation â€“ is due to a complex interplay of multiple genomic factors.

164 - Patterns of Duplicate Gene Retention After Whole Genome Duplication: Mechanistic and Process-Based Modeling

David A Liberles

Amanda E Wilson, David A Liberles
Temple University, Temple University

Gene duplication is an important process for the evolution of novel gene functions. Previously, we presented the first models that enabled the probabilistic inference of evolutionary mechanisms leading to duplicate gene retention, that differentiated nonfunctionalization,

neofunctionalization, subfunctionalization, and dosage balance. Independently, in analyzing duplicate gene retention following two rounds of whole genome duplication in Atlantic salmon, we observed that genes retained after the second round of whole genome duplication were approximately equally likely to be retained whether they were from retained from the first round or were not. This was surprising in light of the commonly accepted gene duplicability hypothesis, that some genes are inherently more likely to be retained due to either their function or to the number of interacting partners they have (with a dosage balance mechanism). Here, we present the formal expectations of the gene duplicability hypothesis in the framework generated for analysis of the Atlantic salmon genome based upon underlying mechanisms. We generated new data from monocot species with multiple rounds of whole genome duplication that can also generate insight into these evolutionary processes. Parameterizing our model leads to an explanation of the duplicable gene content in genomes and the processes leading to long term retention.

208 - Male reproductive pressures drive the emergence of new genes and sexual conflict in Drosophila

Deanna Arsala

Deanna Arsala, Shengqian Xia, Andrea Gschwend, William Koval, Jianhai Chen, Jared Atlas, Laura Faulere, Grace Yu, Muzi Li, Nicholas VanKuren, Daniel J Sanchez, Dylan Sosa, Joseph Mihaljevic, Manyuan Long, Manyuan Long
The University of Chicago , The University of Chicago , The Ohio State University, The University of Chicago , The University of Chicago , The University of Chicago , The University of Chicago, The University of Chicago, The University of Chicago, The University of Chicago, The University of Chicago, The University of Chicago, Northern Arizona University, The University of Chicago, The University of Chicago

Young gene duplicates often exhibit testis-biased expression, suggesting a role in sexual conflict and evolution. We hypothesized that male reproductive pressures drive the emergence and shape the evolution of new genes. To test this, we conducted an RNAi screen of nearly 20% of young gene duplicates (<25 million years old) and 33 ancient genes in Drosophila. Using the UAS-GAL4 system in D. melanogaster, we measured the reproductive fitness contributions of these genes in the soma and germline in both sexes.

Germline knockdowns revealed that genes across evolutionary ages largely conferred reproductive advantages to both sexes. However, somatic knockdowns showed approximately 55% of genes benefited male fitness at the expense of females, regardless of evolutionary age, indicating widespread intralocus sexual conflict (ISC) that persists through time. Interestingly, the genes with the strongest conflict effects were predominantly X-linked and ancient. These genes have the highest expression in the male and female ganglion, which contains neuronal structures that innervate male reproductive tissues and regulate mating behaviors. These patterns imply an evolutionary trajectory where genes initially become fixed due to their immediate germline benefits but subsequently adapt to cater to male somatic needs, negatively impacting female fitness. Our findings highlight the significant role of male reproductive pressures in driving gene emergence and evolution and reveal the pervasiveness and persistence of ISC across the genome.

211 - Evolution of the peroxiredoxin gene family: an antioxidant enzyme in the context of hypoxia-induced by aquatic mammals dives.

Giovanna Selleghin Veiga

Giovanna Selleghin Veiga, Mariana Nery
UC Berkeley, University of Campinas, University of Campinas

Peroxiredoxins are a group of antioxidant enzymes that play a crucial role in eliminating hydrogen peroxide (H2O2) produced during cellular metabolism. In conditions such as tissue ischemia/reperfusion, H2O2 production is heightened. This scenario is recurrent in

aquatic mammals, such as cetaceans and pinnipeds, when they perform breath-hold dives that lead to depletion of oxygen stores and cell hypoxia due to prolonged submersion. In such circumstances, peroxiredoxins are essential for regulating metabolism and preventing oxidative damage. Thus, our objective was to investigate evolutionary patterns of the peroxiredoxin gene family among aquatic mammalian species. Accordingly, we identified the number of ortholog and paralog genes among mammals, examining events of gene duplication, patterns of chromosomal distribution, and adaptive selection. We found that most subtypes of peroxiredoxins are characterized by single-copy genes with minimal missing data, with PRDX1, the most highly expressed subtype found in both the nucleus and cytoplasm, exhibiting an increase in gene copies in aquatic mammals, particularly in pinnipeds. The distribution of the copies was dispersed along the genome and the majority were retrocopies of the main coding sequence with little divergence among them. These results suggest that the peroxiredoxin gene family has a role in the evolution of aquatic mammals by reducing oxidative damage intrinsic to diving behavior in these species.

276 - A genomic perspective of variation in a highly diverse Dipteran family

Wenjun Zhou

Wenjun Zhou, Chris D Jiggins

Department of Zoology, University of Cambridge, Department of Zoology, University of Cambridge

Structural variation and transposable elements have been shown to be major factors influencing species diversity and adaptation. Here, we explore the structural variation and repetitive landscape of six high-quality chromosome level reference genomes of the soldier fly family, Stratiomyidae, a highly diverse taxa with differences in distribution, behaviours and morphology. Whole genome synteny showed chromosomal rearrangements across the phylogeny. A total of 86,085 genes were assigned to 14,181 orthogroups, with 1,142 of which are species-specific. Result showed a large number of gene duplications throughout their speciation history, with most of them happening at the species nodes in the phylogenetic tree, indicating recent and species-specific duplication events related to their adaptation. All species used in the study showed a high proportion of repetitive elements, varying from 63.73% to 73.18% of the whole genome. Among these species, the black soldier fly (Hermetia illucens) has the highest proportion of long interspersed nuclear elements (LINEs), with 45.14% of its genome consisted of this type of transposable elements. Our result reveals rarely-explored divergence in a wide-spread Dipteran family, and suggests potential link between genomic structural variation and functional adaptation of their genes.

344 - NLRs on the move: Transposition drives the diversification of fungal immune receptors

S. Lorena Ament-Velásquez

S. Lorena Ament-VelÃ¡squez, Hanna Johannesson, Aaron A. Vogan

Stockholm University, University of Bordeaux, Stockholm University, Uppsala University

NOD-like receptors (NLRs) are intracellular proteins that independently evolved as key components of the innate immune system of plants and animals. NLRs also exist in fungi, where they can be highly diverse and fast-evolving. The few fungal NLRs of known function control vegetative incompatibility, a process akin to immunity that prevents the spread of mycoviruses. Using high-quality genome assemblies of the Podospora anserina species complex, we show a rapid turnover of NLR gene copies. Indeed, NLRs are often associated with structural variants within and between species, some of which exhibit signatures of balancing selection and/or introgression. Importantly, these polymorphic NLRs are consistently flanked by specific sequences and 5bp-long short direct repeats typical of DNA transposons. We further show that one Podospora lineage that lost a genome defense mechanism against transposons also carries the expansion of two unrelated NLR families. We speculate that these NLRs act as non-autonomous elements mobilized by an unknown transposon, boosting the diversity of variants with potential immunity functions, while simultaneously promoting divergence in genome architecture between

species.

368 - Exploring MicroRNA Evolutionary Dynamics in Mammals

Sarahjane Power

Sarahjane Power, Zixia Huang
University College Dublin, University College Dublin

The evolutionary history of mammals reflects a continuous process of genetic innovation, giving rise to a diversity of adaptations. To explore the mechanisms underlying this, substantial efforts have been made to investigate the evolution of protein-coding genes. However, genomic analyses have revealed a significant degree of conservation in these genes, inconsistent with the variation observed across mammals. This suggests that protein-coding genes alone cannot fully explain the diversity of adaptations within mammals, and that regulation of non-coding genes likely plays a crucial role. MicroRNAs are short non-coding molecules that play a key role in regulating genes post-transcriptionally. Although there is substantial evidence to implicate microRNAs in organismal complexity, their biological relevance in mammalian evolution remains largely unclear due to their poor characterisation within the clade. The current landscape of curated microRNAs is limited to only a few model species, and their annotation is heavily reliant on RNA-seq availability. To this end, this study aims to unravel the evolutionary dynamics of microRNAs across mammals using a mapping-based approach. Leveraging known microRNA sequences, their orthologues will be identified in various mammalian genomes. By examining their evolutionary trajectories, this research will ascertain the molecular changes of microRNAs across the current mammal phylogeny, such as copy number variation and nucleotide substitution. This research will provide a comprehensive investigation into microRNA evolution in mammals, building our knowledge beyond the coding genome and helping to bridge the gap between genome and phenotype in the complex class Mammalia.

380 - Distinct Ancient Structural Polymorphisms Control Heterodichogamy in Walnuts and Hickories

Jeffrey Groh

Jeffrey Groh, Diane C Vik, Matt Davis, Kristian A Stevens, Patrick J Brown, Charles H Langley, Graham Coop
UC Davis, UC Davis, UC Davis, UC Davis, UC Davis, UC Davis, UC Davis

Some of the most striking polymorphisms in nature are seen in the maintenance of multiple mating types. One example in angiosperms is heterodichogamy - dimorphism for the order of male vs. female flowering in hermaphrodites. This mating system is common throughout Juglandaceae, including globally important nut crops - walnuts (Juglans), as well as pecan and other hickories (Carya). In both genera, heterodichogamy is controlled by a single dominant allele. We fine-map the locus in each genus, and find two ancient (>50 Mya) structural variants involving different genes that both segregate as genus-wide trans-species polymorphisms. The Juglans locus maps to a 20 kb structural variant adjacent to gene involved in a flowering time signaling pathway (TPPD-1). Across species, the dominant haplotype contains a tandem array of 8-12 duplicated sequence motifs, each of which contains an inverted copy of the 3' UTR. In developing male flowers, this acts as a template for the transcription of numerous small RNAs, corresponding with a relative increase in expression of the cis copy of the flowering gene and delayed male flowering of female-first morphs. The Carya locus maps to a cluster of 20 tightly-linked variants across 20 genes, likely acting as a supergene. The dominant haplotype rarely recombines and shows both reduced diversity and increased transposable element accumulation, mirroring patterns seen in heterogametic sex-chromosomes. We did not detect either genetic system in other heterodichogamous genera within Juglandaceae, suggesting that additional genetic systems for heterodichogamy remain undiscovered.

417 - Spontaneous mutation rate for and strength of purifying selection against structural variants in

the C. elegans genome

Charles F Baer

Ayush S Saxena, Charles F Baer

University of Florida, Dept. of Biology, University of Florida, Dept. of Biology

Structural variants (SVs) are the most understudied part of the mutation spectrum, and much less is known about their mutational properties than those of single-nucleotide variants (SNVs) and short indels. We estimated the SV mutation rate from six mutation accumulation (MA) lines of C. elegans using long-read (PacBio) sequencing. We identified 32 SV mutations in four N2 strain lines and 17 in two PB306 lines. We previously identified an average of 56 SNV and 14 short indel mutations per line. The total number of bases affected is ~20kb in N2, with a net gain of ~3kb across 4 lines (15 insertions and 17 deletions). For PB306, ~16kb were affected, with a net gain of 7kb (6 insertions and 11 deletions).

Comparison of the SNV/SV ratio in MA lines to that in three natural isolates provides an estimate of the relative strength of selection against SVs, which is ~7X greater for SVs. Selection against SVs in intergenic regions is ~4X greater than against SNPs in exons. The correlation between the pairwise divergence at SNVs and SVs in the wild isolates was >0.99.

False positives were identified by visual inspection of putative mutants in IGV, resulting in a signal (inferred mutant) to noise (false positive) ratio in the MA lines of ~0.7. False negative rates were determined by simulating variants in the reference genome, and observing 'recall'. Recall rate ranges from >90% for short (~100-bp) indels and declines with increasing SV length, with deletions having higher recall rates than insertions.

418 - Structural variants contribute to Anopheles gambiae and An. coluzzii adaptation to urban environments

Marta Coronado-Zamora

Marta Coronado-Zamora, Giannis Zampetakis, Diego Ayala, Josefa GonzÃ¡lez

Institute of Evolutionary Biology (CSIC, UPF), Institute of Evolutionary Biology (CSIC, UPF), IRD, Institut Pasteur de Madagascar, Institute of Evolutionary Biology (CSIC, UPF)

Anopheles gambiae and An. coluzzii mosquitoes are major human malaria vectors in Africa, accounting for most of the transmission. While urban environments were until recently considered to be unfit for Anopheles larvae development, these mosquito species have rapidly adapted to polluted habitats, posing challenges for malaria control. Therefore, understanding the genetic factors driving this adaptability is crucial. In this work, we have analyzed 375 An. gambiae and An. coluzzii WGS samples from urban and rural areas in six Central African countries. Taking advantage of recent high-quality long-read assemblies for both species, our analysis focused on identifying genetic variants, from SNPs to structural variants including transposable elements (TEs) and inversions. We have created the first manually curated TE library for An. gambiae, containing 295 consensus sequences and including 53 new TE families. By combining three TE annotation tools, PoPoolation2, TEMP2 and TEFLoN, we identified 5,462 and 4,773 euchromatic TE insertions present at high frequencies in An. gambiae and An. coluzzii populations, respectively, that could be potentially involved in adaptation. We also found that the frequencies of the cosmopolitan 2La and 2Rb inversions, previously linked to ecological adaptation in these species, varied from low in urban locations to high in rural locations (6%-86% and 0-100% for 2La, 0%-78% and 0%-28% for 2Rb, in An. gambiae and An. coluzzii, respectively). Overall, our work shows that beyond SNPs, structural variants, including TEs and inversions, contribute to the ecological adaptation of malaria vectors.

438 - Fine-Scale Resolution of the Structure and Synteny of the Complete Major Histocompatibility

Complex (MHC) Region from Human and Six Ape Species

Joanna Malukiewicz

Joanna Malukiewicz, Britta S Meyer, Primate T2T Consortium , Tobias Lenz

University of Hamburg, University of SÃ£o Paulo , University of Hamburg, Penn State University , NIH, University of Washington, University of Hamburg

Adaptively, the major histocompatibility complex (MHC) region, which encompasses about 3 million bases in the human genome, is of keen evolutionary interest. MHC genes play key roles in immune function and exhibit the highest association with diseases across the entire human genome. A suite of unique features likeÂ frequent gene duplications, high gene density, high polymorphism levels, and high repetitiveness define this region. However, due to these evolutionary features, it has historically been difficult to deduce genome-levelÂ structural variation between the MHC of humans and their closest non-human primate relatives. To address this gap, we harness newly available telomere-to-telomere genomic assemblies from human, chimpanzee, gorilla, Bornean and Sumatran orangutans, and siamang. These assemblies from the Primate T2T consortium use cutting edge long-read sequencing technologies, and resolve the diploid, phased structure of â€œproblematic regionsâ€• such as the MHC at unprecedented scale. Here we describe/compare the gene organization, structure, and copy number variation of classical and non-classical MHC genes within the full MHC region of humans and the ape species above. Our characterizations of immunologically important MHC genes include previously annotated and as well as novel loci in apes, and we provide novel characterization of the siamang MHC region. Our comparisons reveal structural variation at both the intra-individual and especially the interspecific level variation which likely resulted from both neutral and adaptive processes, reflecting different evolutionary trajectories of ape species. This study highlights recent sequencing advancesÂ that facilitate understanding inter-specific structural variation in challenging genomic regions.

529 - Transcriptomic and proteomic effects of gene deletion are evolutionarily unconserved

Yang Li

Yang Li, Jianzhi Zhang

Department of Ecology and Evolutionary Biology, University of Michigan, Department of Ecology and Evolutionary Biology, University of Michigan

The function of a gene is the effect for which the gene was selected and/or by which it is maintained. Although gene function is commonly inferred from the phenotypic effects of deleting the gene, not all phenotypic effects reflect the geneâ€™s selected-effect function. To evaluate the degree to which the phenotypic effects of gene deletion inform gene function, we compare the transcriptomic and proteomic effects of systematic gene deletions in budding yeast (Saccharomyces cerevisiae) with those effects in fission yeast (Schizosaccharomyces pombe). Surprisingly, despite evidence for functional conservation of orthologous genes, their deletions result in no more sharing of transcriptomic or proteomic effects than that from deleting non-orthologous genes. Because the wild-type mRNA and protein levels of orthologous genes are significantly correlated between the two yeasts, our observation cannot be explained by rapid evolution or large measurement error of gene expression. Analysis of transcriptomic and proteomic effects of gene deletions in multiple S. cerevisiae strains reveals a high sensitivity of these effects to the genetic background. Together, these observations suggest that most transcriptomic and proteomic effects of gene deletion do not inform selected-effect function. This finding has important implications for assessing and/or understanding gene function, pleiotropy, and biological complexity.

541 - The new complete human genome assembly provides new insights in Copy Number Variations across humans

Lucia Bazan Williamson

Lucia Bazan Williamson, Emily Jane McTavish
University of California, Merced, University of California, Merced

Copy number variations (CNVs) represents a crucial aspect of human genomic diversity, influencing phenotypic variation, disease susceptibility, and population genomics. This study examines CNV patterns across human populations utilizing the latest complete human genome assembly (T2T-CHM13) and contrasts the findings with the previous GRCh37 assembly. Discrepancies were found in CNV frequencies between different assembly, withare observed, with notable differences in the number of duplications and deletions. The T2T-CHM13 assembly unveils substantial differences in copy numbers, including many novel duplications, particularly in telomeric regions, reflecting improvements in genome assembly quality. Population structure analyses highlighted contrasting patterns of population differentiation between assemblies, impacting interpretations of human evolutionary history. The improved T2T-CHM13 assembly enabled discovery of highly differentiated CNVs, known introgressed regions, and gene candidates previously undetected. Overall, these findings underscore the critical importance of leveraging updated, high-quality genome assemblies for comprehensive CNV landscape characterization and accurate population genetic inferences, while highlighting potential biases introduced by previous assemblies.

612 - "Characterization of Inversions in 1000 Individuals across the Human Population using Single-Cell Pooled and Long-Read Data."

Vasiliki Tsapalou

Vasiliki Tsapalou, Wolfram Hoeps, Tobias Rausch, Thomas Weber, Siegfried Schloissnig, Eva Benito-Garagorri, Tobias Marschall, Jan Korbel
European Molecular Biology Laboratory (EMBL), Genome Biology Unit, Department of Human Genetics, Radboud University Medical Center, European Molecular Biology Laboratory (EMBL), Genome Biology Unit, European Molecular Biology Laboratory (EMBL), Genome Biology Unit, Research Institute of Molecular Pathology (IMP), Vienna Biocenter, European Molecular Biology Laboratory (EMBL), Genome Biology Unit, Institute for Medical Biometry and Bioinformatics, Medical Faculty, Heinrich Heine University Dul̀ˆsseldorf, European Molecular Biology Laboratory (EMBL), Genome Biology Unit

Structural variants (SVs) are key contributors to human genomic diversity. Despite technological advances in SV detection, knowledge gaps remain, notably for inversions. These copy-number neutral variations, often flanked by identical repeat sequences, are notoriously challenging to detect. Using data from 1000 individuals in the 1000 Genomes Project, we aim to comprehensively identify inversions, understand their formation mechanisms, and explore their phenotypic impact and evolutionary relevance across populations.

Employing single-cell sequencing (Strand-Seq) and Oxford Nanopore Technologies (ONT) data from the 1000 Genomes IMP/MARVL study, we aim to identify a broad spectrum of previously undetected inversions. Strand-Seq's ability to reveal read directionality is crucial for detecting large inversions, particularly near complex regions. In this regard, we pursued an innovative single-cell data generation method, coupling cell line pooling with strand-specific sequencing. This approach enables efficient sequencing of diverse human samples, optimizing cell numbers for precise inversion identification and genotyping. Concurrently, we leverage ONT data and refine existing computational approaches to capture smaller to mid-sized genomic rearrangements, with a focus on detecting small inversions up to 1kb.

Our research has identified 638 novel polymorphic inversions, establishing the most comprehensive inversion dataset to date. We will

discuss these events, emphasizing their recurrence, functional consequences, and population distribution, including rare haplotypes with diverse phenotypic impacts.

This endeavor is poised to advance large-scale population-genetics studies, providing a unique opportunity to decipher natural and disease-causing structural variations within various population groups, and their role in adaptation and genome evolution.

641 - Exceptional genetic variation at MHC class II genes in Atlantic herring unveiled through long-read sequencing

Fahime Mohamadnejad Sangdehi

Fahime Mohamadnejad Sangdehi, Minal S Jamsandekar, Brian W Davis, Mats E Pettersson, Leif Andersson
Uppsala University, Texas A&M University, Texas A&M University, Uppsala University, Uppsala University, Texas A&M University

Major histocompatibility complex (MHC) genes are renowned for their high polymorphism across vertebrate genomes. They encode cell-surface molecules crucial for presenting antigenic peptide fragments to T-cells, playing a pivotal role in adaptive immune defense. It is believed that the main force maintaining diversity in classical MHC genes is balancing selection driven by arm race between host and pathogens. In this study, we explore class II MHC genes in Atlantic herring, a schooling, long-distance migratory fish species that is distinguished as one of the most abundant vertebrates and occurs in huge shoals providing a favorable target for pathogens. We employed PacBio HiFi long-read approach to sequence 14 Atlantic herring individuals, and constructed 28 de novo haploid assemblies using hifiasm assembler. We then conducted the annotation and manual curation of MHC-II genes in each haploid assembly. Our findings revealed MHC-II genes are distributed across nine loci located on four chromosomes, with copy number variation observed at some loci. Sequence polymorphism in exon 2 of certain MHC-II genes, encoding the peptide-binding region, was strikingly high, with nucleotide diversity exceeding 10%, and dN/dS around 2.0. Clustering MHC-II alleles by sequence similarity primarily demonstrated locus-specific grouping but also highlighted instances of alleles from different loci forming mixed clusters, suggesting genetic exchange between loci. Overall, Atlantic herring display remarkable genetic variation at MHC-II genes, possibly ranking among the highest documented, potentially reflecting an adaptation for defense against the diverse pathogens they encounter.

643 - Evidence that gene copy number plays a role in the evolution of life history traits across mammals

Sophie Matthews

Sophie Matthews, Marc Tollis, Cathal Seoighe
University of Galway, SFI Centre for Research Training in Genomic Data Science, Northern Arizona University, University of Galway, SFI Centre for Research Training in Genomic Data Science

Cancer is a disease of multicellularity, observed across the tree of life. In principle, animals with larger body sizes and longer lifespans should be at increased risk of developing cancer. However, cancer risk is not found to increase with body size or lifespan across mammals. Previous studies have proposed the duplication of cancer-related genes as an explanation of this phenomenon, whereby an increased number of cancer genes enhances the robustness of cancer suppression pathways. However, these studies have not extended beyond genes that are already deemed to be cancer related. In this study, we conducted a phylogenetic generalized least squares (PGLS) analysis for association between copy number of all protein-coding genes and longevity, body size and cancer prevalence, in 94 mammalian species. Applying gene set analysis methods to these results, several processes showed evidence of enrichment for genes that show an association between their copy number and these traits. Notably, many of these pathways were immune-related. We also

found enrichment for copy number associations with longevity in both DNA repair and telomere-related gene sets, suggesting a role for gene copy number changes in these processes in the evolution of extended lifespan. These results provide a more comprehensive picture of processes and pathways that are likely to be important in the evolution of key life history traits, helping to shed light on long-standing questions relating to the mechanisms of adaptation to changes in life history traits that have decoupled these traits from cancer risk.

695 - De novo chromosome-scale assembly of an indica x japonica hybrid rice landrace from Vietnam

Meng-Ting Hsieh

Meng-Ting Hsieh, Rachel Rusholme-Pilcher, Jose De-Vega, Rafal Gutaker, Anthony Hall
Earlham Institute, Earlham Institute, Earlham Institue, Kew, Earlham Institute

After 1960s, the Green Revolution led to a broad introduction of a limited number of elite cultivars. Generally, it led to a move away locally adapted landraces and diminishment of the gene pool. However, the genetic erosion in allelic diversity and fraction of dispensable genes has not been well quantified. In Vietnam, most of the elite varieties and landraces were previously classified into two indica rice subgroups, I1 and I2. We believe that this reflects introduction of elite varieties and narrow genetic diversity. An additional subgroup, I5, was found to be unique to Vietnam and different from all defined sub-populations in 3000 Rice Genetic Project (3K-RGP), a widely used dataset believed to cover global rice genetic variation. We assume that I5 was maintained owing to the local environment or culinary preferences and could be valuable genetic resources. To validate our hypothesizes, we generated a de novo assembly of an I5 representative with Pac-bio. We set out to map the population data onto the genome to measure the genetic diversity and identify structural variation and introgression.

790 - Imperfect plagiarism: the evolutionary fate of laterally acquired structural variants in grasses

Catherine F Collins

Catherine F Collins, Luke T Dunning
University of Sheffield, University of Sheffield

Lateral gene transfer (LGT) is widespread in eukaryotes, including in grasses, wherein gene acquisition is reported in crops and wild relatives alike. These transfers involve the insertion of large multigene DNA fragments, but it is presently unclear whether all the acquired genes are expressed and retained, or if non-beneficial linked genes are purged from the recipientsâ€™ genome. We use RNA-seq to determine the proportion of laterally acquired genes that remain expressed. We then compare the relative expression levels of the laterally acquired genes to their orthologous vertically inherited copy in the recipient. Finally we determined the expression level of the laterally acquired genes prior to the transfer event by sequencing the transcriptome of the donor species. We see a range of patterns in relative expression among the genes, including laterally acquired genes that are silenced, expressed more similarly to the copy in the donor, and those that align with the vertical copy in the recipient. These patterns shed light on the fate of a laterally acquired gene, suggesting that neutral and deleterious genes tend to be silenced or pseudogenised. In contrast, advantageous genes are retained and result in subsequent pseudogenisation of the vertical copy. This research provides an insight into the integration of laterally acquired genes and the ripple effect this acquisition causes to the genome.

797 - Unraveling the Free-Threshing Trait in wheat: Characterization and Evolution

Konstantinos Gidiotis

Konsantinos Gidiotis , Laura BotiguÃ©, Cristobal Uaay, Nikolai Adamski, Assaf Distelfeld

Centre for Research in Agricultural Genomics, Centre for Research in Agricultural Genomics, John Innes Centre, John Innes Centre, Institute of Evolution, University of Haifa

Triticum aestivum, commonly known as bread wheat, plays a vital role in ensuring food security and environmental sustainability. Understanding the evolutionary history and genetic makeup of this plant species can hold immense value for future agricultural endeavors. In this study, we aim to shed light in to the evolution and origin of the free-threshing trait, a trait that distinguishes ancestral wheat from the domestic one, by studying a QTL on chromosome 2B, known to be associated with this phenotype, in different Triticum accessions. We found a 5Mb inversion differentiating hulled and free-threshing wheat accessions that coincides with the QTL. Additionally, we examined the SNP variability across a diverse range of Triticum varieties to investigate the origin of the free-threshing trait and study the patterns of genetic differentiation in a panel of hulled and free-threshing wheat accessions. By using tools such as comparative genomics and population genetics we aim to delve deeper into the evolutionary dynamics of the free-threshing trait. Moreover, our investigation highlights the potential of wild emmer wheat as a genetic resource for the improvement of modern bread wheat varieties, offering valuable traits for adaptation to changing environmental conditions and emerging agricultural challenges. All in all, our study highlights the importance of accounting for structural variation in the study of adaptation, in particular in crops like wheat, where repetitive elements account for more than 80% of the genome.

823 - Evolution of G-Quadruplex Motifs Across Ape Telomere-To-Telomere Genomes

Saswat Kumar Mohanty

Saswat K Mohanty, Kaivan Kamali, Kateryna D Makova, The T2T Primates Consortium
Department of Biology, The Pennsylvania State University, Huck Institutes of the Life Sciences, The Pennsylvania State University, Department of Biochemistry and Molecular Biology, Penn State University, Department of Biology, The Pennsylvania State University, Department of Biology, The Pennsylvania State University, Huck Institutes of the Life Sciences, The Pennsylvania State University, Center for Medical Genomics, The Pennsylvania State University, Department of Biology, The Pennsylvania State University, Genome Informatics Section, National Human Genome Research Institute, 49 Convent Dr, Department of Genome Sciences, University of Washington School of Medicine

G-quadruplexes (G4s) are non-canonical (i.e. different from the right-handed double-helix) DNA structures that can form at approximately 1.5% of the human genome. G4s can fold at guanine-rich motifs due to hydrogen bonds among nonconsecutive guanines, and their formation occurs in living cells. G4 structures facilitate double-strand breaks and genomic instability, and thus induce structural variation and point mutations. Additionally, G4 structures have recently emerged as critical regulators of replication, transcription, and telomere maintenance, and, consequently, have been demonstrated to evolve under purifying selection. Despite their significance for mutagenesis and cellular functions, the evolution of G4s remains understudied. To fill this gap, we conducted a comprehensive analysis of G4 motifs in the recently-released complete, telomere-to-telomere (T2T) genomes of apesâ€”bonobo, chimpanzee, gorilla, Bornean and Sumatran orangutans, and siamang gibbon. Using genome-wide multi-species alignments, we found that in general, G4 motifs accumulate at a rate that is expected based on the phylogeny among the studied species. Whereas many G4 motifs were shared among species, such sharing was not uniform across the genome and was particularly strong at enhancers and promoters, highlighting the role of G4 structures in regulation of transcription. At the same time, we discovered thousands of species- and genus-specific G4 motifs, suggesting their rapid emergence/decay and potential role in evolution of individual ape taxa. Our results open avenues for subsequent studies of individual G4 motifs and structures and of their role in ape evolution.

829 - Investigating inversions in Lake Malawi cichlids for a role in speciation and phenotypic variation

Nikesh M Kumar

Nikesh M Kumar, Jeffrey T Streelman, Patrick T McGrath
Georgia Institute of Technology, Georgia Institute of Technology, Georgia Institute of Technology

Large inversions and other more complicated structural variants have the remarkable property of suppressing recombination between inverted and non-inverted haplotypes, allowing for the accumulation of many functional alleles that are inherited as a single unit. These large variants have been proposed to play an important role in adaptation to local environments and speciation in the presence of gene flow. Here, we use optical genome mapping to explore the presence of inversions in Lake Malawi cichlids, a collection of over 800 species representing one of the largest vertebrate evolutionary radiations known to man. We identify eight new inversions or tandem inversions, between about 0.5 - 22 Mbp in length. Using short-read sequencing of over 141 species, containing members from all seven ecogroups in the lake, we infer the genotypes of these inversions across the lake. Some inversions appear fixed within several of the ecogroups, while others segregate within individual ecogroups in a manner consistent with their spread through hybridization. Their distribution and evolutionary history are supportive for a role in speciation, either in separating the major ecogroups into specific habitats or in subsequent speciation into distinct ecological niches. We will discuss their potential role in controlling sex-determination and prezygotic mating behaviors.

1076 - Intra-genomic coevolution between a DNA satellite and Topoisomerase II triggers reproductive barriers between species

Cara Brand

Cara Brand, Mia Levine
University of Pennsylvania, University of Pennsylvania

Postzygotic reproductive barriers between species are triggered by deleterious interactions among diverged genetic factors. Several incompatibility factors map to either rapidly evolving DNA satellites or to adaptively evolving satellite-interacting proteins. These reports raise the possibility that lineage-specific coevolution between proteins and satellites leads to incompatibilities when brought together in hybrids. However, the functions of these proteins and the satellites with which they interact are largely unknown. Consequently, the cellular pathways perturbed by species incompatibilities are poorly understood. To study the genetic factors and molecular mechanisms underlying species incompatibilities, we probed the 359bp satellite array, which proliferated along the Drosophila melanogaster lineage. This species-specific satellite requires Topoisomerase II (Top2) to resolve DNA entanglements at the array. We show that Top2 evolves adaptively between D. melanogaster and its sister species D. simulans. To determine if 359bp and Top2 coevolve, we introduced a D. simulans version of Top2 (â€œTop2[sim]â€•) into D. melanogaster. Maternally provisioned Top2[sim] in the embryo triggers genome instability and aberrant mitotic divisions, resulting in early embryonic lethality. This mitotic catastrophe is caused by 359bp mis-segregation between dividing embryonic nuclei. Consistent with a history of coevolution, deleting 359bp rescued chromosome segregation and embryogenesis. These data indicate that D. simulans Top2 is incompatible with the D. melanogaster-specific 359bp satellite. We propose that Top2[sim] fails to resolve known DNA entanglements that accumulate at 359bp, triggering chromosome mis-segregation and genome instability. This work provides novel mechanistic insights into how rapid satellite DNA divergence triggers reproductive barriers between species.

1085 - Genomic and Transcriptomic Signatures of Alternative Reproductive Tactics in a Mexican Poeciliid

Gabriel A Preising

Gabriel A Preising, Daniel L Powell, Tristram O Dodge, Alexandra E Donny, John J Baczenas, Theresa R Gunn, Samantha Mason, Terrance T Yang, Kang Du, Oscar Rios-Cardenas, Molly R Morris, Molly Schumer
Stanford University, Stanford University, Stanford University, Stanford University, Stanford University, Stanford University, Ohio University, Stanford University, Xiphophorus Genetic Stock Center, Texas State University San Marcos, Instituto de EcologÃ-a, AC, Ohio University, Stanford University

Intense sexual selection can drive the emergence of alternative reproductive tactics (ARTs), or strategies that increase reproductive success through mechanisms other than courtship. The evolutionary history of ARTs is puzzling, as multiple reproductive morphs can exist together within a population for long periods of time, hinting at fitness tradeoffs between reproductive strategies that maintain multiple phenotypes. The swordtail fish, Xiphophorus multilineatus, exhibits two opposing genetically-determined male size morphs. Large males exhibit sexual ornamentation and are behaviorally fixed for courtship while small males lack sexual ornamentation and are plastic for courtship and coercive mating tactics. Variation in the number of Y-linked non-functional mc4r copies has been shown to account for some of the differences between morphs. However, crosses have pointed to the existence of autosomal factors that contribute to some of the variation between morphs as well. Here, we take a whole-genome approach to dissect the variants contributing to differences between morphs. We perform a genome-wide association study (GWAS) for male morph and demonstrate a polygenic basis for male size while capturing the Y-linked mc4r signal demonstrated by previous studies. We generate PacBio assemblies for each morph to identify differences in structural variation and confirm the variation in Y-linked mc4r copy number. Finally, we explore the transcriptomic profiles of relevant tissues between morphs to refine our GWAS results and further explore physiological differences between them. These results provide a comprehensive characterization of a balanced polymorphism under strong selection in natural populations.

162 - Genomic erosion through time in avian species undergoing population collapse

Hernán E. Morales

HernÃ¡n E Morales, Georgette Femerling, Claudia Fontsere, Xufen Liu, Samuel Speak, Ester Milesi, Barry Hartup, Andy Caven, Ross Crates, Rachel Bristol, Molly Cassatt-Johnstone , Beth Shapiro, Jim Groombridge, Cock van Oosterhout, Tom Gilbert
University of Copenhagen, McGill University, University of Copenhagen, University of Copenhagen, University of East Anglia, DTU, INTERNATIONAL CRANE FOUNDATION, INTERNATIONAL CRANE FOUNDATION, Australian National University, Independent consultor, University of California Santa Cruz, University of California Santa Cruz, University of Kent, University of East Anglia, University of Copenhagen

The escalating biodiversity crisis threatens species with the loss of essential genomic diversity needed to survive and adapt to future environmental change. Of particular concern is that population decline may have a delayed effect on the genomes of threatened species, relative to demographic processes. This delay can result in genomic erosion, even after population size recovery, a phenomenon known as the 'genetic drift debt.' Temporal genomics can offer the necessary resolution to understand these changes. We compared re-sequenced whole genomes from pre-decline (100+ year-old museum samples) and post-decline (modern samples) populations in three closely monitored endangered bird species: the Seychelles paradise flycatcher (recovered from N=28), the whooping crane (recovered from N=16), and the regent honeyeater (declining N<300). These species experienced distinct trends of population decline and conservation management actions over recent decades. Employing a population genomics approach we investigated how different types of genomic variation (deleterious, functional, neutral) have responded to population declines and subsequent conservation efforts. We show how despite their demographic recovery, past declines have left a strong genomic signature, with loss of genetic diversity, inbreeding and accumulation of genetic load. We showcase that the loss of genomic diversity, and the risk of extinction are influenced by

demographic trajectories, the severity of their decline, connectivity, and conservation management strategies. Using forward-in-time simulations, parameterized with empirical temporal data, we make predictions of future trajectories of genomic erosion and identify how we can implement effective conservation actions

354 - Metagenomic reconstruction of the plant community in the stomach content of a steppe bison living over 48,000 C14 years ago

Chenyu Jin

Chenyu Jin, Adrian Forsythe, MiklÃ³s BÃ¡lint, Frank Kienast, Pavel Nikolskiy, Katerina Guschanski, Tom van der Valk
Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Centre for Palaeogenetics, Department of Zoology, Stockholm University, Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Senckenberg Biodiversity and Climate Research Centre, Department of Insect Biotechnology, Justus-Liebig University, LOEWE Centre for Translational Biodiversity Genomics, Senckenberg Research Institute and Natural History Museum, Geological Institute, Russian Academy of Sciences, Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Institute of Ecology and Evolution, School of Biological Sciences, The University of Edinburgh, --, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Centre for Palaeogenetics

By reconstructing past ecosystems, we can gain valuable context for understanding how extant taxa are distributed and how ecosystems may respond to changing climate in the future. During the last cold stage (115 - 11.7 kya), Arctic Eurasia exhibited a distinctive plant composition consisting of steppe and xerophilous arctic plants adapted to a dry climate. This combination formed the Pleistocene mammoth steppe, which provided a habitat for large herbivores like mammoths (Mammuthus primigenius) and steppe bison (Bison priscus). In this study, we take a unique opportunity to analyze a plant paleo-community of the Arctic mammoth steppe through DNA sequencing of a well-preserved stomach content of a steppe bison, which was discovered in the Kolyma Lowlands in northeast Russia in 2009 and dated to >48,000 C14 years before present. Macrofossil analysis of this sample revealed a floral composition resembling a saline meadow. In line with macrofossil evidence, we identified alkali grass (Puccinellia sp.) and Larch (Larix sp.) in the DNA data of the steppe bison sample. We also confirmed the presence of plant genera such as Artemisia, Allium, Silene, Poa, Carex, and Dryas, which are typically associated with arid soil conditions, whereas the detected genera Puccinellia, Festuca, Poa, and Potamogeton are indicative of high salinity. This list includes taxa previously not identified by macrofossil analyses and allows us to obtain a more comprehensive view of the mammoth steppe flora and steppe bison dietary preferences.

549 - Unveiling genetic and ecological factors driving divergence population in Mesoamerican Common Mustached Bat

Tania Garrido-Garduño

Tania GarridoGarduÃ±o, Ricardo LÃ³pezWilchis, Luis GuevaraChumancero, JosÃ© Domingo RuizOrtiz, Karen PonceMuÃ±oz, Carlos Alejandro LunaArangurÃ©, Daniel PiÃ±eroDalmau, Ella VÃ¡zquezDomÃnguez
Iztacala Faculty of Higher Studies, UNAM, Autonomous University of Mexico, Iztapalapa, Autonomous University of Mexico, Iztapalapa, Kansas State university, Iztacala Faculty of Higher Studies, UNAM, Institute of Atmospheric and Environmental Sciences, Institute of Ecology, UNAM, Institute of Ecology, UNAM

The Mexican mormoopids inhabit one of the most biologically complex regions globally, leading to distinct genetic structuring patterns among species distributed partially or entirely within this area. Recent diversification within the genus Pteronotus bats in Mexico has given rise to P. mesoamericanus, which diverged from P. mexicanus around 0.7 million years ago. Despite being recognized as a distinct

species, the evolutionary history of P. mesoamericanus in Mexico remains unclear. Our objective was to employ a comprehensive approach, integrating various methods to evaluate the evolutionary history and genetic patterns influenced by environment. To accomplish this, we utilized data from 67 sequences (21 localities) of the hypervariable domain HVII of the mitochondrial control region, and SNPs from 89 individuals (18 localities) exclusively from Mexico. Our analysis encompassed phylogeographic structure, historical demography, as well as ecological niche modeling and a phyloclimatespace approach. Our findings reveal distinct genetic groups: Gulf of Mexico and the Southeast of the country. This genetic structuring is most pronounced in the Southeast but at the same time with the highest genetic diversity. Notably, in YucatÃ¡n Peninsula, we observed a greater number of lineages with unique haplotypes and alleles. Furthermore, our analysis did not indicate isolation by distance, suggesting that the differentiation of these groups is primarily driven by environmental factors, particularly temperature fluctuations from the past. Understanding the presence of distinct lineages, each geographically and environmentally isolated from one another, provides us with insights into the potential fate of these lineages under climate change.

824 - Ancient environmental DNA from the Pliocene High Terrace deposits in the Canadian High Arctic

Bianca De Sanctis

Bianca De Sanctis, Alexandra Rouillard, Kurt H KjÃ¦r, Mikkel W Pedersen, Alberto Reyes, Natalia Rybczynski, Lasse Vinner, Eske Willerslev

University of Santa Cruz, UmeÃ¥ University, University of Copenhagen, University of Copenhagen, University of Alberta, Canadian Museum of Nature, University of Copenhagen, University of Copenhagen

The Pliocene High Terrace deposits on Ellesmere Island in Canada's High Arctic are known for their exceptional preservation and high fossil yield, especially the Beaver Pond and Fyles Leaf Beds sites. Extensive paleoecological study of these sites has uncovered a larch-dominated boreal forest with an array of wetland flora, grasses, extinct species of beaver, rabbit and wolverine, and even a High Arctic camel. These strata date to approximately 3.9 million years ago, when mean annual temperatures in the Arctic were 14-22C higher than they are today. Here, we report preliminary results of DNA isolated and shotgun sequenced from sediment samples collected from Beaver Pond and Fyles Leaf Beds.

918 - Phylogenomic reconstruction and evolutionary history of the Nothofagus of South America

Gabriela Narváez

Gabriela NarvÃ¡ez, Daly Noll, Josefina Jorquera, Roberto Nespolo, Juliana Vianna

Instituto de Ciencias Ambientales y Evolutivas, Universidad Austral de Chile, Millennium Nucleus Limit of Life (LiLi), Institute Millennium Center for Genome Regulation (CRG), Institute Millennium Center for Genome Regulation (CRG), Millennium Institute Biodiversity of Antarctic and Subantarctic Ecosystems (BASE), Facultad de Ciencias, Universidad de Chile, Millennium Nucleus Limit of Life (LiLi), Millennium Nucleus Limit of Life (LiLi), Institute Millennium Center for Genome Regulation (CRG), Pontificia Universidad CatÃ³lica de Chile, Instituto de Ciencias Ambientales y Evolutivas, Universidad Austral de Chile, Millennium Nucleus Limit of Life (LiLi), Center of Applied Ecology and Sustainability (CAPES), Millennium Institute for Integrative Biology (iBio), Millennium Nucleus Limit of Life (LiLi), Institute Millennium Center for Genome Regulation (CRG), Millennium Institute Biodiversity of Antarctic and Subantarctic Ecosystems (BASE), Pontificia Universidad CatÃ³lica de Chile

The Nothofagus genus has attracted considerable attention due to its complex biogeographic history in the Southern Hemisphere. However, the lack of genome-wide data and the limitation of obtaining samples representative of the geographical distribution of the

different species still represent a challenge to understanding speciation in forest environments and tree plants in general. In South America, the Nothofagus forests are mainly conformed by the subgenus Nothofagus, widely distributed across the Andes Mountains of Chile and Argentina.

We generated high-coverage whole genome resequencing data from 10 representative individuals of the five species of subgenus Nothofagus (N. pumilio, N. antarctica, N. dombeyi, N. nitida and N. betuloides). We have reconstructed the evolutionary history of this group, the phylogenetic tree, and raised the contribution of introgression. We have also compared the gene content and orthologous relationships between these species and other Nothofagus trees. Overall, we highlight interesting features of this subgenus and suggest that the resources presented herein contribute to the complete understanding of the speciation and reconstructions of deep phylogenetic relationships such as those that have been unresolved and in constant debate in all genus Nothofagus.

Therefore, this project will raise high-impact scientific knowledge for Nothofagus, contributing to a better understanding of our local biodiversity, and thus allowing research opportunities to open in this area from the global south.

919 - The impact of whaling in the Northeast Atlantic right whale population (Eubalaena glacialis) using palaeogenomics

Gonçalo E. Themudo

GonÃ§alo E. Themudo, Alba Rey-Iglesia, Paula F. Campos
CIIMAR, University of Porto, The Globe Institute, University of Copenhagen, CIIMAR, University of Porto

The North Atlantic (NA) right whale (Eubalaena glacialis) was a primary target for early industrial whalers, particularly the Basques, who extensively hunted them for their high-quality oil. Once abundant on both sides of the North Atlantic, inhabiting its temperate and subpolar waters, these whales are now functionally extinct in the Northeast Atlantic and critically endangered according to the IUCN Red List. Despite their fragmentary remains making identification challenging, bones are abundant in the archaeological and historical records and reveal a long history and extensive tradition connected to the exploitation of these marine resources.
In this study, using state of the art ancient DNA techniques we aim to genetically characterize the functionally extinct population of NA right whales from the Northeast Atlantic. We sequenced 17 full genomes from historical specimens collected from the seabed near Medieval whaling ports in the Cantabrian Sea, dating from the 13-18th century. Data gathered from these specimens was compared with genomes of modern individuals from the extant population in the Northwest Atlantic. Our results highlight differences in genetic diversity over time, shedding light on the evolutionary impact of whaling on this species.

45 - Genome doubling results in N6-Methyladenosine and physiological response divergence enhanced tetraploid Hordeum bulbosum tolerance to salt stress

Genlou Sun

Genlou Sun, Margaret Scarrow
Saint Mary's University, Saint Mary's University

Polyploidy has played a critical role in increased tolerance to environmental stress compared to diploid progenitors. N6-Methyladenosine (m6A) is suggested to play an important role in stress response. However, whether genome doubling affects m6A to increase autopolyploids stress tolerance is still unclear. This study aims to compare physiological and m6A changes between autotetraploid and diploid wild barley (Hordeum bulbosum) in response to salt (NaCl) stress. Our results showed that: (1) Relative water content was

significantly higher and water loss was significantly lower in autotetraploid H. bulbosum than diploid at 0 days (control), 14 days (salt stress) and 28 days (recovery); (2) proline was significantly higher in autotetraploid barley at day 0 but through treatment and recovery proline was higher in diploids. Both autotetraploid and diploid barley showed a significant increase at treatment and recovery when compared to controls; (3) H2O2 was significantly higher in diploids compared to tetraploids; (4) Chlorophyll A was higher in autotetraploids, however diploids showed a more significant drop under stress than autotetraploids, while chlorophyll B appeared higher in autotetraploid controls except for controls observed at recovery; (5) m6A in total RNA showed no significant difference between ploidies in controls, but during treatment and recovery autotetraploids were significantly higher than diploids. Both had a significant increase in m6A for treatment and recovery when compared to controls. These results suggest that increased m6A might be one of molecular mechanisms that increases salt tolerance in autotetraploid H. bulbosum compared to diploids, which enhances the adaptation of autopolyploids.

78 - Transition from natal downs to juvenile feathers: conserved regulatory switches in birds

CHIH-KUAN CHEN

CHIH-KUAN CHEN, Yao-Ming Chang, Ting-Xin Jiang, ZhiCao Yue, Tzu-Yu Liu, Jiayi Lu, Zhou Yu, Jinn-Jy Lin, Trieu-Duc Vu, Tao-Yu Huang, Hans I-Chen Harn, Chen Siang Ng, Ping Wu, Cheng-Ming Chuong, Wenâ€•Hsiung Li
University of Southern California, Academia Sinica, National Chung Hsing University, Academia Sinica, University of Southern California, Shenzhen University Medical School, University of Southern California, University of Southern California, University of Southern California, National Center for High-performance Computing, University of Michigan School of Medicine, Academia Sinica, University of Southern California, National Tsing Hua University, University of Southern California, University of Southern California, Academia Sinica

The transition from natal downs for heat conservation to juvenile feathers for simple flight is a remarkable environmental adaptation process in avian evolution. However, the underlying epigenetic mechanism for this primary feather transition is mostly unknown. Here we conducted time-ordered gene co-expression network construction, epigenetic analysis, and functional perturbations in developing feather follicles to elucidate four downy-juvenile feather transition events. We discovered that LEF1 works as a key hub of Wnt signaling to build rachis and converts radial downy to bilateral symmetry. Extracellular matrix reorganization leads to peripheral pulp formation, which mediates epithelial - mesenchymal interactions for branching morphogenesis. ACTA2 compartments dermal papilla stem cells for feather cycling. Novel usage of scale keratins strengthens feather sheath with SOX14 as the epigenetic regulator. We found this primary feather transition largely conserved in chicken (precocious) and zebra finch (altricial) and discussed the possibility that this evolutionary adaptation process started in feathered dinosaurs.

224 - Wolbachia induces cellular differentiation in Drosophila cell culture cells following stable infection.

Jodie Jacobs

Jodie Jacobs, Samuel Sacco, William Seligmann, Cade Michanandi, Merly Escalona, Richard Green, Shelbi L Russell
University of California, Santa Cruz, University of California, Santa Cruz, University of California, Santa Cruz, University of California, Santa Cruz, University of California, Santa Cruz, University of California, Santa Cruz, University of California, Santa Cruz

Wolbachia is an endosymbiotic bacterium prevalent in arthropods and nematodes, renowned for its manipulation of host fertility. The phenotypic effects of Wolbachia implicate core developmental mechanisms prompting an investigation into its potential role in mediating cellular differentiation in Drosophila melanogaster. Through sequencing the transcriptome of Drosophila melanogaster grown in vitro we have identified many genes with differential expression between uninfected cultures and those with stable Wolbachia infections. Cluster

analysis of these data suggest alterations in cell state associated with Wolbachia infection. Furthermore, we observed significant changes in the genome-chromatin contact map in infected cells compared to their uninfected counterparts. These alterations in chromatin structure and gene expression propose a potential mechanism through which Wolbachia may influence host cellular differentiation. We employed Gene Ontology (GO) analysis to examine both genes with differential expression and genetic elements within regions of the genome with enriched contacts. The results revealed significant genome-chromatin contact enrichment in immune response pathways, including genes like bonamin, serpin, and tetraspanin in stable Wolbachia infections. GO analysis of the host transcriptome revealed that Wolbachia infections induce significant enrichments in the ecdysteroid kinase-like gene family and folate transporter-like family while suppressing gene expression in the metalloexopeptidases gene group. These findings provide new insights into the molecular mechanisms underlying Wolbachia niche formation and maintenance in Drosophila melanogaster.

263 - Exposure to early-life stress induces developmental shifts in behaviour and DNA methylation in a small prey fish

Janay A. Fox

Janay A. Fox, Simon M. Reader, MÃ©lanie F. Guigueno, Rowan D. H. Barrett
McGill University, McGill University, McGill University, McGill University

Early-life experiences can predict the environments experienced later in life, giving individuals an opportunity to develop behaviour that matches the environment (â€˜adaptive programmingâ€™). Epigenetic mechanisms such as DNA methylation (DNAm) offer a link between the environment and the genome that could underlie mechanism such as developmental plasticity. The goals of this study were to investigate the impact of early-life exposure to predation stress on behaviour and DNAm in the brain of Trinidadian guppies (Poecilia reticulata), a small freshwater fish. We exposed guppies throughout development to either alarm cue (conspecific skin extract), inducing predation stress, or a control cue. We then gave them two behavioural assays, an open-field and a shoaling test, before performing whole-genome bisulfite sequencing on whole brains. Guppies exposed to alarm cue during development exhibited increased shoaling in adulthood compared to those exposed to the control treatment, but there were no impacts detected on boldness or exploratory behaviour. The shift in shoaling behaviour displayed by guppies exposed to alarm cue matches that seen in naturally occurring populations, suggesting that developmental plasticity could be used to adaptively program behaviour to match predation environment. Additionally, we identified stable shifts in DNAm in the brain in response to developmental alarm cue exposure. These shifts in DNAm were in genes involved in neural signaling and responses to stimulus. In conclusion, this study has shown how early-life predation stress can induce developmental plasticity and that shifts in neural DNAm could be an underlying mechanism of developmental behavioural plasticity.

463 - Novel insights into X-Chromosome Inactivation through Peromyscus Leucopus

Maria De Lourdes Andrade Ludeña

Maria DL Andrade LudeÃ±a , Sha Sun, Aqsa Motiwala, Anthony D Long, Alan G Barbour
University of California, Irvine, University of California, Irvine, University of California, Irvine, University of California, Irvine, University of California, Irvine

In placental mammals, X-chromosome inactivation (XCI) is necessary for ensuring gene expression balance between XY males and XX females by silencing one X chromosome in females. XCI is initiated by the long non-coding RNA (lncRNA), X inactive specific transcript (Xist). However, the exact mechanism of Xist remains debated. The rodent Peromyscus leucopus has recently emerged as a model organism given its high levels of genetic diversity within the species. We identified an Xist homolog in P. leucopus that shares both human and mouse Xist features. Using RNA FISH, we visualized the localization of Xist lncRNA in the nucleus of P. leucopus fibroblast

cells. Contrary to current knowledge that Xist is exclusively expressed in female cells, we detected low levels of Xist expression in male P. leucopus fibroblasts. This suggests a different mechanism of X-inactivation where the single X chromosome in males may undergo gene silencing. We performed multi-species sequence alignments to define the six RNA motifs (Repeats A-F) within Xist homologs and identified one, Repeat C, that heavily varies in copy number and nucleotide sequence across species. Repeat C is heavily degraded in P. leucopus and humans, but expanded in mice, indicating divergent Xist RNA functions during evolution. To test our hypothesis that Xist sequence changes underlie male Xist expression, we use CRISPR-mediated gene editing in male and female P. leucopus cells. Comparative analysis among rodent species will further our understanding of the evolutionary and functional differences in Xist regulation, and how lncRNAs evolve.

509 - TRIFORCE: A bioinformatic tool to identify TADs

Alejandro Díaz de la Vega

Alejandro DÃ-az de la Vega , Dulce Irene Valdivia, Luis Delaye, Katarzyna Oktaba
Departamento de IngenierÃ-a GenÃ©tica, Cinvestav Irapuato, Departamento de IngenierÃ-a GenÃ©tica, Cinvestav Irapuato, Departamento de IngenierÃ-a GenÃ©tica, Cinvestav Irapuato, Departamento de IngenierÃ-a GenÃ©tica, Cinvestav Irapuato

The accurate identification of Topologically Associating Domains (TADs) is crucial for understanding genome organization and gene expression regulation. However, existing tools for TAD annotation present inconsistencies and lack of reproducibility, limiting the attainment of precise and reliable results. To address this challenge, we present TRIFORCE, a bioinformatic tool that automates and enhances TAD annotation. TRIFORCE integrates two widely used methods, such as HiCFindTADs from HiCExplorer and Arrowhead from Juicer, to identify candidate TADs and analyze structural changes in TAD separation scores. This method identifies genomic regions with similar trends in contact frequency, revealing potential TADs. TRIFORCE employs a majority vote approach to integrate TAD calls from the aforementioned methods, ensuring greater robustness and accuracy in consensus TAD identification. TRIFORCE validation was performed using 5 kb and 10 kb resolution matrices with Drosophila melanogaster data. The results demonstrate that TRIFORCE significantly improves consistency and reproducibility in TAD annotation compared to existing tools. TRIFORCE is an innovative bioinformatics tool that automates and enhances TAD annotation, offering a comprehensive solution for TAD analysis with greater precision, robustness, and reproducibility. Its use promises a significant impact on understanding genome organization and gene expression regulation.

698 - Parent-of-origin specific DNA methylation and extreme sexual dimorphism in the absence of sex chromosomes

Tamsin Woodman

Tamsin Woodman, Hollie Marshall, Laura Ross
University of Edinburgh , University of Leicester, University of Edinburgh

Many organisms show sexually dimorphic phenotypes that are regulated by differential expression and gene splicing. The mealybug Planococcus citri has extreme sexual dimorphism whereby males and females differ dramatically in morphology, lifespan and the way they inherit genetic material. Mealybugs do not possess sex chromosomes and as such males and females are genetically identical, suggesting this extreme sexual dimorphism is epigenetically regulated. Males exhibit a form of genomic imprinting known as paternal genome elimination (PGE) whereby the haploid paternal genome is silenced in somatic cells and eliminated from the germline. The

molecular mechanisms behind this are still under speculation, however the involvement of DNA methylation seems likely. Mealybugs show distinct sex-specific DNA methylation profiles and considering the known involvement of DNA methylation in imprinting in mammals and plants, it is possible that it also plays a role in this species. In addition to this, mealybugs have independently evolved promotor methylation associated with gene silencing, more similar to the methylation profile seen in mammals than in other insects. To further understand the role of DNA methylation in PGE, we have combined RNA-seq and DNA methylation data with parental SNP information to ascertain whether differential methylation occurs on the maternal or paternal strand of DNA, and how this is linked to sex-biased gene expression as well as the differential behaviour of paternal chromosomes in males and females. Together these results will give insights into epigenetic regulation of sexual dimorphism and genomic imprinting in an insect with unique biology.

711 - Somatic-Mutation Based Deconvolution of 5MC in CpG and Non-CpG Context

Justin J.S. Wilcox

Justin J.S. Wilcox, Quentin Foucault, Toni GoÃŸmann
Technische UniversitÃ¤t Dortmund, Technische UniversitÃ¤t Dortmund, Technische UniversitÃ¤t Dortmund

Tissues represent a fundamental evolutionary interface at the junction of phenotype and genotype. Tissues differentiate by epigenetic mechanisms, which are of core interest to developmental and evolutionary biology. Epigenetic studies are however restricted by access to pure tissues, and underlying assumptions about the locations and distributions of epigenetics markers. Advanced surgical techniques cannot typically be applied to the foci of evolutionary studies in the field, nor can prior assumptions about the predisposition for and consistency of methylations in CpG context be safely made in most non-model organism. Epigenetic modifications are nonetheless linked to consistent patterns of somatic inheritance. As such, generalized evolutionary methodologies may allow for the efficient deconvolution of mixed tissues across eukaryotes. Here we apply somatic site-frequency information to deconvolute 5mC methylation patterns though analysis of mutational accumulation. We use simulations and archival reads to show that somatic mutations are common and detectable in next-generational sequencing data. We then use somatic-site-frequency changes to accurately derive the relative mixtures of bisulfite reads in in silico mixtures of known proportions. Using this information, we are able to perform unbiased and accurate deconvolution of mixed tissue methylations in CpG and non-CpG context. We are ultimately able to recover 20-30% of differentially-methylated sites, and approximately 50% of differentially-methylated CpG islands and gene bodies in any cytosine context at contamination levels up to 90%. Our findings highlights the utility of evolutionary theory across scales, and expand the accessibility of epigenetics to studies.

728 - Constructing a cattle and sheep pan-epigenome to elucidate epigenetic impacts upon gene expression and phenotypic variation

Stephanie McKay

Stephanie McKay, Shangqian Xie, Morgan Stegemiller, Gabrielle Becker, Kimberly Davenport, Katie Shira, Denise Konetchy, Patricia Villamediana, Brenda Murdoch
University of Missouri, University of Idaho, University of Idaho, University of Idaho, Washington State University, University of Idaho, University of Idaho, South Dakota State University, University of Idaho

Phenotypic variation is driven by both genetic and epigenetic variation, yet understanding the extent of epigenetic modifications and their impact on phenotype remains largely unexplained. Moreover, the onset of pangenomics remains disconnected from pan-epigenomics and hence inhibits a fully comprehensive view of diversity. In order to facilitate dissection of epigenetic contributions towards phenotypic variation, we have begun generation of a cattle and sheep pan-epigenome. Ultimately, an exhaustive characterization of epigenetic conservation and diversity will be examined with 5-methylcytosine, 5-hydroxymethylcytosine, histone marks (H3K4me3, H3K27ac,

H3K4me1 and H3K27me3), open chromatin, chromatin accessibility and N6-methyladenosine in three tissues from each of three different breeds of cattle and sheep. In this study we have performed Whole Genome Bisulfite Sequencing (WGBS) and Methylated RNA Immunoprecipitation Sequencing (MeRIP-Seq) to investigate the effects of DNA 5-methylcytosine and RNA N6-methyladenosine on gene expression. After characterization and comparative analysis, results indicate that RNA methylation has a greater influence on gene expression compared to DNA methylation. Subsequently, gene interaction networks were constructed to demonstrate the interactions of DNA methylation, RNA methylation and gene expression. Specifically, we identified two gene interaction networks whose gene expression was substantially impacted by RNA methylation. Interestingly, both of these gene interaction networks are known to influence transcription. These findings enhance our understanding of genome-wide dynamics and the complex interactions among DNA and RNA methylation with gene expression, offering new perspectives and opportunities to further epigenetics research.

755 - DNA methylation patterns in the placenta of viviparous fish

Yolitzi Saldivar Lemus

Yolitzi Saldivar Lemus, Constantino Macias Garcia, Leeban Yusuf, Michael G Ritchie

Texas State University, National Autonomous University of Mexico, National Autonomous University of Mexico, University of Saint Andrews, University of Saint Andrews

Sexual selection and the evolution of viviparity set the stage for a potential genetic conflict over parental investment between paternal and maternal alleles during offspring development, and possibly, between maternal and embryonic genes. This conflict has the potential to affect gene regulation in different ways. For instance, parentally-induced changes in epigenetic marks such as DNA methylation, can modulate or silence gene expression. We describe DNA methylation patterns in the trophotaenia (the embryonic component of the trophotaenial placenta of goodeids) of four species that differ in proxies for the strength of sexual selection, and therefore, potentially in the degree of sexual conflict. We analysed DNA methylation patterns of the four species in a CpG context, identified Differentially Methylated Sites (DMS) between species pairs and also looked at Allele-Specific Methylation patterns in the trophotaenia and muscle (control tissue) from adults of Girardinichthys multiradiatus, the most sexually dimorphic species. DNA methylation patterns vary across species and gene components, particularly in G. multiradiatus, which had the lowest proportion of genes with gene body methylation, but the largest number of intermediately methylated and unmethylated genes in the trophotaenia. Only a few genes were consistently methylated, intermediately methylated or unmethylated across species, which suggests a high turnover of gene methylation. Although from our preliminary data we cannot precisely determine what evolutionary forces are linked to interspecific variation in DNA methylation patterns, our results suggest that sexual selection/conflict could be, at least, partially playing a role.

793 - Defining the Role of Cis-Acting DNA Sequence in Evolutionary Divergence of the Bivalent Chromatin State in Mammals

Kimberly Griffin

Kimberly Griffin, Kira Marshall, Kenny Thai, Bluma Lesch

Yale, Yale, Yale, Yale

Changes in gene regulation play a major role in evolution. Bivalency is an important developmental gene regulatory state defined by the presence of the activating histone mark H3K4me3 and repressive mark H3K27me3 on the same nucleosomes. Bivalency is found in cells with stem-like characteristics and is located in the promoters of transcriptionally silent genes important for somatic differentiation during embryonic development. Bivalency is present in spermatogenic cells throughout mammalian evolution and is proposed to poise developmental genes to be efficiently activated or repressed after fertilization as cells differentiate in the embryo. It is currently not

understood how the bivalent state is regulated.

We are using a comparative evolutionary approach to determine what sequence changes are important in causing a gene to gain bivalency in a species lineage. The gene Traf6 is bivalent in mouse embryonic stem cells (mESCs) and spermatogenic cells, while it is only marked by H3K4me3 in other mammalian species. To determine what genomic sequence designates Traf6 as bivalent in mouse but not in other species, we generated a humanized reporter by replacing the promoter of mouse TRAF6 with the corresponding region from the human genome in mice and mESCs. We inserted GFP after the humanized sequence, allowing us to track transcriptional consequences of changes in chromatin state. We found that murine germ cells carrying the humanized Traf6 promoter recapitulate the human regulatory state, providing us with a system for dissecting how changes in promoter sequence affect bivalency at this model locus and potentially impact phenotype.

794 - Altering translation accuracy mediates bacterial adaptation to phages

Laasya Samhita

Aabeer K Basu, Laasya Samhita

Ashoka university, Ashoka university

Mistranslation, specifically the incorporation of incorrect amino acids into polypeptide sequences, increases phenotypic variability within clonal populations. Potentially, this helps the bacterium explore a range of non-mutational phenotypic space, although the variability is often non-heritable. Although error prone translation is known to provide specific benefits under stressful environments, its role in aiding host-pathogen interactions remains unexplored. Using plaque assays and by tracking growth dynamics, we tested if modifying translation accuracy in E. coli affects susceptibility to bacteriophage P1 infection. Preliminary results indicate that both increasing and decreasing the rate of mistranslation can confer absolute phage resistance to E. coli. The resistance observed is absolute and stable across generations, suggesting a genetic link between phage infection pathways and translation accuracy, rather than a phenomenon mediated by enhanced population variability. Intriguingly, the observed resistance is dependent on the identity of the mutation introduced to modify mistranslation rate, rather than the extent to which accuracy is altered. Additionally, enhancing bacterial mutation rates using mutator alleles does not impact resistance to phage, suggesting that mistranslation driven mutagenesis is unlikely to underlie our observations. We are presently examining molecular mechanisms that could explain our observations via microscopy and whole genome sequencing.

1049 - CTCF regulates a lncRNA to coordinate chromatin landscape and gene expression in hematopoietic cells

Hober Nelson Núñez Martínez

Hober Nelson NÃºÃ±ez-MartÃ-nez, Gustavo Tapia-UrzÃºa, Ã•ngel JosuÃ© Cerecedo-Castillo, Carlos Alberto Peralta-Ã•lvarez, Georgina Guerrero, Maite Huarte, Felix | Recillas-Targa

Institute of Cellular Physiology, UNAM, Institute of Cellular Physiology, UNAM, Institute of Cellular Physiology, UNAM, Institute of Cellular Physiology, UNAM, Institute of Cellular Physiology, UNAM, Center for Applied Medical Research, University of Navarra, Institute of Cellular Physiology, UNAM

The architectural protein CCCTC-binding factor (CTCF) has emerged as an essential regulator of gene expression in cell differentiation. However, the relationship and its regulatory effect on lncRNA genes in hematopoiesis remains elusive. We demonstrated that CTCF targets and regulates the lncRNA DUBR, which was highly expressed in human hematopoietic stem and progenitor cells (HSPC) but depleted in erythroblasts. By CRISPR-Cas9-mediated genome editing and genome-wide analysis of transcriptome as well as chromatin

landscape, we assessed that DUBR depletion enhances the expression of myeloid-erythroid cell differentiation genes by coordinating genome-wide enhancer activation and reducing the activity of the transcriptional repressor HES1 at gene promoters. Our results indicate that attenuation of lncRNA DUBR coordinates the establishment of a myeloid-erythroid gene expression program.

1059 - Unraveling the epigenetic mechanisms of heat stress adaptation in Heliconius erato lativitta: Insights from gene expression analysis of Hsp40 and Hsp90.

Ángel Andrés Ojeda

Angel AndrÃ©s Ojeda, Caroline Nicole Bacquet

Universidad Regional AmazÃ³nica Ikiam, Universidad Regional AmazÃ³nica Ikiam

The escalating occurrence and severity of extreme temperature episodes pose a substantial challenge for species' adaptation. Ectothermic organisms, such as insects, rely on external sources to regulate their body temperature. Among Lepidoptera, Heliconius erato lativitta, a butterfly species prevalent in Neotropical forests, has long been studied, revealing susceptibility to heat stress throughout its life cycle. While the induction of heat shock proteins (Hsps) aids in mitigating protein denaturation and maintaining cellular equilibrium, epigenetic mechanisms in this species remain unexplored. This study investigates the gene expression of Heat shock protein 40 kDa (Hsp40) and Heat shock protein 90kDa (Hsp90) in H. e. lativitta under heat stress conditions. Employing RT-qPCR and Livak's Î”Î”CT method, gene expression was assessed across head, thorax, and abdominal tissues in both male and female individuals. Our findings identified two Hsp40 family genes, DnaJ-1 and DnaJ-5, along with one Hsp90 family gene. Notably, DnaJ-1 and DnaJ-5 were underexpressed at 40 Â°C but overexpressed at 5 Â°C shock conditions, while Hsp90 exhibited overexpression at 40 Â°C. Abdominal and head tissues demonstrated the highest levels of overexpression across all treatments, while the thorax exhibited underexpression. Generalized linear models revealed that temperature treatments significantly influenced gene variability, with tissue types also playing a notable role, albeit no significant differences were observed between sexes. This study contributes valuable insights into the epigenetic mechanisms underlying the adaptation of tropical butterflies to climate change.

1098 - Evolutionary study on DNA methylation in the tree shrew brain and its role in X chromosome inactivation

Dongmin Son

Dongmin Son, Lei Shi, Soojin V. Yi

Department of Ecology, Evolution and Marine Biology, University of California Santa Barbara, Neuroscience Research Institute, University of California Santa Barbara, Key Laboratory of Genetic Evolution & Animal Models, Kunming Institute of Zoologyï¼ŒChinese Academy of Sciences, Key Laboratory of Animal Models and Human Disease Mechanisms of Yunnan Province, Kunming Institute of Zoology, Kunming College of Life Science, University of Chinese Academy of Sciences, Department of Ecology, Evolution and Marine Biology, University of California Santa Barbara, Neuroscience Research Institute, University of California Santa Barbara, Department of Molecular, Cellular and Developmental Biology, University of California Santa Barbara

Tree shrew is an emerging model organism to study human evolution due to several advantages, notably their evolutionary proximity to primates. Here we analyzed the first genome-wide, single-base-resolution methylome and transcriptome from tree shrew brains. We identified a clear correlation between gene expression and methylation level in both the promoter and gene body regions. We examined the X chromosome DNA methylation and observed a significant global reduction (hypomethylation) of DNA methylation across the entire X chromosome in females. This pattern resembled what was observed in koalas (a marsupial) and differed from that in humans. Our comparative analysis indicated that the degree of differential DNA methylation between the female and male promoters was dependent

on CpG contents. Conversely, differences in CpG contents between species could explain divergent patterns of DNA methylation. Moreover, we show that female X hypomethylation does not directly contribute to the silencing of the inactivated X chromosome, nor does it significantly drive sex-specific gene expression, suggesting the existence of a unique role of DNA methylation in XCI. We also newly annotated the Xist locus in tree shrew, and demonstrate that it is strongly regulated by DNA methylation, influencing its extremely sex-specific expression. This study provides a foundation for future investigations into the DNA methylation system in tree shrews, facilitating the exploration of evolutionary aspects of primate DNA methylation systems and the role of DNA methylation in the X chromosome.

49 - Whole-genome phylogenomics of the tinamous (Aves: Tinamidae): comparing gene tree estimation error between BUSCOs and UCEs illuminates rapid divergence with introgression

Lukas J Musher

Lukas J Musher, Therese A Catanach, Thomas ValquÃ-, Robb T Brumfield, Alexandre Aleixo, Kevin P Johndon, Jason D. Weckstein
The Academy of Natural Sciences of Drexel University, The American Museum of Natural History, The Academy of Natural Sciences of Drexel University, Universidad Nacional Agraria La Molina, CORBIDI-Centro de OrnitologÃ-a y Biodiversidad, Louisiana State University, Instituto Tecnologico Vale, University of Illinois Urbana-Champaign, Drexel University

Incomplete lineage sorting (ILS) and introgression increase genealogical discordance across the genome, which complicates phylogenetic inference. In such cases, identifying orthologs that result in gene trees with low estimation error is crucial because phylogenomic methods rely on accurate gene histories. We sequenced whole genomes for the tinamous (Aves: Tinamidae) to dissect the sources of gene and species-tree discordance and reconstruct their interrelationships. We compared results from: (1) coding genes (BUSCOs), (2) ultraconserved elements (UCEs) with short flanking regions, (3) UCEs with intermediate flanks, and (4) UCEs with long flanks. We hypothesized that orthologs with more phylogenetically informative sites would result in more accurate species trees because the resulting gene trees contain lower error. Consistent with our hypothesis, we found that long UCEs had the most informative sites and lowest rates of error. However, despite having many informative sites, BUSCO gene trees contained high error compared to long UCEs. Unlike UCEs, BUSCO gene sequences showed a positive association between the proportion of parsimony informative sites and gene tree error. Thus, BUSCO and UCE datasets have different underlying properties of molecular evolution, and these differences should be considered when selecting loci for phylogenomic analysis. Still, species trees from different datasets were mostly congruent. Only one clade, with a history of ILS and introgression, exhibited substantial species-tree discordance across the different data sets. Overall, we present the most complete phylogeny for tinamous to date, identify a new species, and provide a case study for species-level phylogenomic analysis using whole-genomes.

221 - Comparative Genomics of Glossophaga mutica (Phyllostomidae): Demographic History and Evolutionary Processes

Jesús Antonio Rocamontes Morales

Jesus Antonio Rocamontes Morales, Jorge Ortega, Enrique Ibarra Laclette, Gabriela Castellanos Morales
El Colegio de la Frontera Sur (ECOSUR), Unidad Villahermosa, Universidad AutonÃ³ma de Zacatecas, Escuela Nacional de Ciencias BiolÃ³gicas, Instituto PolitÃ©cnico Nacional, Instituto de EcologÃ-a, A.C. (INECOL), El Colegio de la Frontera Sur (ECOSUR), Unidad Villahermosa

The discovery of cryptic species is common in phyllostomid bats. At times, traditional taxonomic tools such as morphology fail to reveal the existence of cryptic species within species complexes. In contrast, comparative genomics allows for the analysis of divergence and

similarity patterns among genomes, and the comparison of the demographic history of a species from a single genome, to contribute to the resolution of taxonomic debates. For instance, within the genus Glossophaga, Glossophaga soricina was previously considered a species complex. Recently this species complex was separated into four species, one of which (Glossophaga mutica) is endemic to Mexico and Central America. Interestingly, G. mutica includes an insular population in the Islas MarÃ-as archipelago, near the coast of Nayarit in Mexico. Phylogenetic relationships between island and mainland populations are still unclear, as they exhibit differences in their morphological characteristics, suggesting that the insular population may have undergone a distinct evolutionary trajectory, which can be studied from a genomic perspective. In this project, we used a whole genome sequencing approach to compare the nuclear genomes of one G. mutica individual from the Islas MarÃ-as archipelago and one from the mainland to understand the differences in their evolutionary trajectories and demographic history to test the hypothesis of two distinct species in mainland and insular G. mutica. We present preliminary results of genomic assemblies and results of SNP discovery between the nuclear genomes of G. mutica from continent and island. Our results show 14,561,725 SNPs between G. mutica genomes.

242 - Integrating paleontological, morphological and molecular methods of telling evolutionary time - assessing the timescale of crocodilian evolutionary history.

Sidney Davies

Sidney Davies, Phil Donoghue, Davide Pisani
University of Bristol, University of Bristol, University of Bristol

In phylogenetics, the molecular clock represents the only viable means of converting the relative ages of species divergences into absolute times in order to provide a fully-dated phylogeny that is calibrated to geologic time. However, there are a diversity of methods available, both to account for rate heterogeneity across the tree and to incorporate information from the fossil record in order to provide reliable temporal constraints that are used to extrapolate the mutation rate and determine the time of divergence among species pairs. There is much debate over which methods are the most appropriate to employ in a variety of evolutionary scenarios, but despite this, very few studies have investigated the relative performance of different clock methods in a single, integrated study. We set out to achieve this goal using crocodilians as a model group to compare the ways in which various commonly-used methods used to incorporate fossil information into a molecular clock analysis behave in an empirical setting, such as node-dating, tip-calibration, total-evidence analysis and the fossilized birth-death process. This allows us to make direct comparisons on the behaviour of disparate clock methods in the context of a well-studied taxonomic group with a rich evolutionary history, and may provide insight on how we can expect various methods to perform in future empirical analyses.

246 - A Backbone Tree of Formicidae —— The First Step to Reveal Ant Evolution in the Cretaceous Terrestrial Revolution and Post K-Pg Extinction Era

Wentao Tao

Wentao Tao, Davide Pisani, Philip Donoghue, Donato Antonio Grasso, Enrico Schifani
University of Bristol, School of Biological Sciences, University of Bristol, School of Biological Sciences, University of Bristol, School of Earth Sciences, University of Bristol, School of Earth Sciences, University of Parma, Department of Chemistry, Life Sciences, and Environmental Sustainability, University of Parma, Department of Chemistry, Life Sciences, and Environmental Sustainability

With over 15,000 species, colonising nearly all terrestrial environments, ants are one of the most successful insects on the earth. However, despite genome-scale data have been used to explore their internal relationship, there are still some inconsistencies between different researches. Mysteries, such as the relationship of the subfamilies Leptanillinae and Martialinae with other living ants, as well as

the root age and divergence time of internal clades, are still waiting to be revealed.
Here, we build up the largest genome-scale dataset, containing 91 ant species and 10 outgroups. After going through a series of cleaning pipeline to remove hidden paralogs and long branch species, we build up trees under the best fitting mixture model, then calibrate them with 19 fossils following the best practice of fossil calibration so as to investigate the evolutionary history of ants.
Our phylogenetic analysis supports the hypothesis that the leptanillomorph clade (Leptanillinae and Martialinae) is the sister group to all other living ants. While the novel molecular clock analyses suggest that crown ants originated in the early-Cretaceous and diversified afterward with a younger age compared to the previous research.
Therefore, in the future, by adding more molecular and fossil data, combining with the latest tree-building and fossil-calibrating method, we can better know about (1) the diversification of crown ants in the late-Cretaceous under the background of the angiospermâ€™s explosion, and (2) how they occupied the niches previously belonging to stem group ants after the K-Pg event, finally achieved success today.

248 - Metagenomics for non-model species and museum collections pave way for studying changes in environmental microbiome.

Rafal Marek Gutaker

Rafal M Gutaker, Robin B Guevarra
Royal Botanic Gardens, Kew, Royal Botanic Gardens, Kew

Natural history museum collections are repositories of time-stamped samples for plant and animal species. Most collections were made in the last couple of centuries and have documented some of the most radical anthropogenic environmental changes. Analysing such collections through the lens of metagenomics would facilitate research on the impact of agriculture, land-use change and widespread use of antibiotics on microbiomes. Two major bottlenecks are: difficulties in distinguishing between bona fide microbiome and post mortem colonizers, and the assumed need to remove host genome for non-model species. Using three examples: cultivated plants (rice), wild plants (ragweed) and insects (beetles), we show that the lack of host reference genome does not preclude informative metagenomic assignments, even in historical samples. Furthermore, as exemplified by plant herbarium, we show that it is possible to build collection-specific reference panel for post mortem contaminations, which in turn facilitates microbiome comparisons in time. Finally, we show one application of such approach and the discovery of previously unknown strains of nitrogen-fixing bacteria associated with rice cultivation over 100 years ago. In summary, we showcase that metagenomics on historical biological samples is possible. Given its huge potential for understanding environmental change, we encourage more research in this direction.

338 - Ancient RNA expression profiles from the extinct woolly mammoth

Emilio Mármol Sánchez

Emilio MÃ¡rmol-SÃ¡nchez, Bastian Fromm, Elif Bozlak, Marianne Dehasque, Morteza Aslanzadeh, ZoÃ© Pochon, Nikolay Oskolkov, Tom van der Valk, Panagiotis Kalogeropoulos, Camilo ChacÃ³n-Duque, Inna Biryukova, Valeri Plotnikov, Albert Protopopov, BjÃ¶rn Andersson, Erik Ersmark, Marc R FriedlÃ¤nder, Love DalÃ©n
Department of Molecular Biosciences, The Wenner-Gren Institute, Science for Life Laboratory., Centre for Palaeogenetics , The Arctic University Museum of Norway, UiT, The Arctic University of Norway, Institute of Animal Breeding and Genetics, University of Veterinary Medicine, Vienna Graduate School of Population Genetics, University of Veterinary Medicine, Human Evolution Program, Department of Organismal Biology, Uppsala University, Department of Molecular Biosciences, The Wenner-Gren Institute, Science for Life Laboratory., Centre for Palaeogenetics, Department of Archaeology and Classical Studies, Stockholm University, Department of Biology, National

Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Lund University, Centre for Palaeogenetics, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Department of Molecular Biosciences, The Wenner-Gren Institute, Science for Life Laboratory, Stockholm University, Centre for Palaeogenetics, Department of Zoology, Stockholm University, Department of Molecular Biosciences, The Wenner-Gren Institute, Science for Life Laboratory, Academy of Sciences of Sakha Republic, Academy of Sciences of Sakha Republic, Department of Cell and Molecular Biology, Karolinska Institute, Centre for Palaeogenetics, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Department of Molecular Biosciences, The Wenner-Gren Institute, Science for Life Laboratory, Centre for Palaeogenetics, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Department of Zoology, Stockholm University

The analysis of ancient DNA has recently gained considerable momentum, allowing the study of extinct and extant organisms that lived up to 2 million years ago. This has enabled the reconstruction of genomes and historical ancestry of multiple extinct species, as well as revealing the complex nature of the bygone ecosystems where they once thrived. However, current DNA sequencing techniques cannot alone provide information about tissue identity, gene expression dynamics or transcriptional regulation â€“ all of which is encoded in the RNA fraction, and critical to fully understanding the biology of these now lost species. Here we report the oldest ancient RNA expression profiles ever recorded and dated to around 30,000 years-old from a permafrost-preserved woolly mammoth, representing the first instance of gene expression profiles characterized in an extinct species from the Late Pleistocene. We report tissue-specific gene expression patterns at both the coding and noncoding level, comparative analyses with mammoth ancient DNA data and genomic and transcriptomic traces supporting that a phenotypically identified female mammoth is actually a genetic male. Our findings provide evidence for the preservation of meaningful transcriptomes beyond preconceived temporal limits, doubling the age record registered so far. With our results, we anticipate a renewed interest in the study of ancient RNA molecules, pushing the field beyond DNA and fostering a new era of integrative paleogenomics and paleotranscriptomics studies.

428 - Optimizing Human Immunodeficiency Virus RNA Reverse Transcription from Archival Fixed Tissues

Tess Wilson

Tess Wilson , Melanie Kuch , Debi Poinar, Aisha Nazli , Charu Kaushic
McMaster University, Institute for Infectious Disease Research , McMaster University , McMaster University , Institute for Infectious Disease Research , McMaster University , Institute for Infectious Disease Research , McMaster University , Institute for Infectious Disease Research

The early evolutionary history of HIV remains relatively unknown due to a paucity of sequence data prior to the 1980s. Without ancestral HIV sequences, understanding of viral emergence is limited to results from phylogenetic models predicting past patterns from modern data. Accessing ancestral sequences from archival fixed tissues presents an opportunity to glimpse into the evolutionary past of the virus. The acquisition of viral sequences from archival tissue is, however, stymied by the low abundance of viral reads against the host background. To combat this methodological challenge, we compared four priming techniques during reverse transcription to optimize whole genome viral cDNA generation. We compared the use of random primers, gene-specific primers, and two combined approaches to optimize the coverage of viral targets. These approaches were tested on modern HIV-1 RNA spiked onto human RNA at various concentrations. The cDNA samples were then prepared for sequencing using single-stranded library preparation and sequenced. The mean coverage of the HIV-1 genome was similar across priming techniques. The copy number of specific amplicons was higher, however, in samples primed using gene-specific primers. These results demonstrate the potential utility of using a combined priming approach to generate near whole-genome data from viral species in archival samples. Understanding the emergence and spread of viruses affecting

modern and historical populations can be improved by accessing viral sequences, but this can only be done if robust techniques specific to viral RNA in fixed tissues are developed, as we have done here

522 - Herbarium seed embryos as alternative sources of ancient DNA

Hong Phuong Le

Hong Phuong Le, Yoon Kyung Lee, Rafal Gutaker
Royal Botanic Gardens, KEW, Royal Botanic Gardens, KEW, Royal Botanic Gardens, KEW

Natural history collections, which contain hundreds of millions of specimens classified in terms of time, space, and taxonomy, provide valuable resources for many fields of research. Since the first success of ancient DNA (aDNA) extraction in the 1980s, these repositories, including herbaria for plants, have been intensively used to study microevolutionary processes, especially genetic responses to anthropogenic activities over the past several centuries. Two crucial challenges of aDNA research comprise contamination by environmental DNA, and degradation. aDNA extracted from herbarium leaf specimens are short fragments, ranging from 50 to 100 bp due to a higher breakdown rate than that in bone assemblage. To maximise the amount of data retrieved and minimise destructive sampling, we have compared whole-genome sequencing libraries constructed for leaves and seed embryos in herbarium specimens. Additionally, damage patterns and the proportion of endogenous DNA in both samples have been evaluated. The results have shown that DNA isolated from embryos was less fragmented than that from leaves. Accordingly, large embryos in seeds are a very promising alternative source of genomic DNA from herbarium specimens.

593 - An unusual mode of chromosomal evolution in Mucoromycota fungi

Helle Tessand Baalsrud

Helle T Baalsrud, Thu-Hien To, Dana Byrtusova, Mariann Arnyasi, Lars GrÃ¸nvold, Juan Fu, Volha Shapaval, Simen Sandve
Norwegian University of Life Sciences, Norwegian University of Life Sciences, Norwegian University of Life Sciences, Norwegian University of Life Sciences, Norwegian University of Life Sciences, Norwegian University of Life Sciences, Norwegian University of Life Sciences, Norwegian University of Life Sciences

Decades of comparative genomics studies have revealed that gene content and order on chromosomes, known as synteny, can be conserved over hundreds of million years of evolution. Synteny will degrade with increasing evolutionary distance between species, however, this process is poorly studied in most lineages as it requires dense sampling of high-quality genomes. Using chromosome-scale assemblies from 25 species in the understudied Mucoromycota fungi we show that one lineage displays an unusual mode of chromosome evolution resulting in conserved mesosynteny. Mesosynteny entails conservation of gene content within homologous chromosomes, but with complete reshuffling of gene order. We also found several polyploidization events, with one species having mesosynteny between its subgenomes. Mesosynteny has previously been reported in a distantly related group of Ascomycota fungi, but its evolutionary significance is currently unknown. We found a unique link in Mucoromycota between mesosynteny and the loss of centromere proteins (CENP-A and C), which are conserved across all other Eukaryotes. This could have affected how new chromosomal rearrangements navigate meiosis as heterozygotes. The mesosyntenic species also have a significantly higher repeat content. The cause-and-effect relationship between repeat evolution, meiosis and chromosomal rearrangements is still an open question in biology. Our study sheds new light on constraints affecting synteny evolution by demonstrating that the mode of synteny evolution varies between lineages, possibly due to differences in the meiotic machinery and/or repeat landscape. Mucoromycota is thus an excellent new system to explore the evolutionary drivers and consequences of breaking up gene linkages.

616 - Genomic basis of the evolution and diversity of Lepidoptera

Peter Mulhair

Peter Mulhair, Asia Hoile, Liam Crowley, Owen Lewis, Peter Holland

University of Oxford, University of Oxford, University of Oxford, University of Oxford, University of Oxford

Lepidoptera represents one of the most diverse animal orders with over 160,000 described species. The diversity in form and behaviour within this clade can be attributed to adaptations at the genetic level. Studies of gene content and genome evolution in Lepidoptera have been limited to certain well-studies species, owing to the lack of representative, high-quality genomes. Recent construction of a large number of chromosome-level assemblies provides the opportunity to analyse the role of gene and genome evolution in the diversification of this order of insects. Here, we provide a genome-wide analysis of protein coding gene evolution across Lepidoptera. We assess order-wide rates of evolution, patterns of gene gain, duplication and loss, and conservation of gene microsynteny across 216 arthropod species, 182 of which are Lepidoptera. Comparing rates of evolution between families reveals variable patterns of genetic diversity, with the families Pieridae, Lycaenidae, and Geometridae displaying the highest overall genome-wide evolutionary rates. Assessing gene turnover reveals dramatic bursts of gene gain in the two nodes subsequent to the early diverging families Micropterigidae and Tineidae (i.e. Ditrysia minus Tineidae and Apoditrysia), with 1,623 and 1,444 gene families gained, respectively. Finally, we analysed the conservation of local gene synteny across Lepidoptera. Similar to genome macrosynteny, we find a high rate of conserved microsynteny across the order, but also family specific gene clusters, for example in Tortricidae (671 gene clusters), Lycaenidae (649 clusters), and Geometridae (418 clusters). These provide insights into family specific gene reshuffling, clustering, and regulation.

739 - An ancient horizontal gene transfer provided raw material in the evolution of light-producing organs in toadfishes

Emily Lau

Emily Lau, Todd H Oakley

University of California, Santa Barbara, University of California, Santa Barbara

Mutations, including gene duplication and horizontal gene transfer, provide raw genetic material for evolution. While gene duplication is considered a major mechanism for generating novel structures and functional innovations, the role of horizontal gene transfer in evolution, especially in eukaryotes, is underappreciated. Horizontal gene transfer is becoming more commonly attributed to functional innovations, but there is thus far little evidence of this mechanism in providing raw genetic material for the evolution of novel structures. Here we provide evidence supporting the role of horizontal gene transfer in providing raw genetic material for the evolution of novel light-producing organs, called photophores, in toadfishes. We perform genome sequencing, phylogenetics, and comparative genomics to identify and investigate the evolution of a structural protein in the photophore lens of toadfishes. Genes homologous to the lens protein are found in other bony fishes of Actinopteri, and separately in non-metazoans such as bacteria and fungi, but undetected in any other animal lineages. Our findings suggest this gene may have been transferred from bacteria to an ancestral Actinopteri fish, over 300 million years ago, and retained prior to being recruited as a structural protein during the evolution of bioluminescence in toadfishes 20 to 10 million years ago. Broadly speaking, our work highlights the important role of an ancient horizontal gene transfer event that provided raw genetic material for the evolution of novelty.

766 - Museomics improves and extends orang-utan mitochondrial phylogeny

Ana Agapito Vera

Ana Maria Agapito Vera, Viviani Mantovani, Samuel Wittwer, Irune Ruiz-Gartzia, NÃºria Albala Hermosilla, Esther Lizano, Logan Kistler, TomÃ s Marques-Bonet, Michael KrÃ¼tzen
University of Zurich, University of Zurich, University of Zurich, Universitat Pompeu-Fabra, Universitat Pompeu-Fabra, Universitat Pompeu-Fabra, Smithsonian Institute, Universitat Pompeu-Fabra, University of Zurich

The demographic history of orangutans (genus: Pongo) has been shaped by ancient geological events, recent human-driven forest exploitation, and a unique dispersal system with extremely strong female philopatry. Because of the latter, phylogenetic analyses based on mitochondrial (mtDNA) genomes allow for a comprehensive reconstruction of past and present genetic diversity of this genus. Previous work revealed reciprocally monophyletic clusters in Borneo and Sumatra, aligning with the three recognized species. However, previous work was hampered by the lack of samples from areas where orangutans could not be sampled or are now extinct.

In order to complement previous sampling efforts and potentially detect orangutan populations that might have been lost, we created a mitogenome phylogeny using 50 modern mitogenomes from previous studies and sequenced 61 mitogenomes from orangutanâ€™s museum specimens. Our analyses revealed a novel monophyletic subclade within Borneo, showing a geographic signal. Generally, divergence times were more recent than previous studies, but still within previous estimates.

Our work highlights the utility of museum samples for obtaining a more complete picture of past genetic diversity, especially in organisms that for logistic or political reasons are hard to sample. The advance of ancient DNA techniques also allows to create whole-genome data from museum samples, allowing more complete insights into how orang-utan genetic diversity responded to anthropogenic perturbations while at the same time advancing our understanding of orangutan evolution in the Holocene.

825 - Ex-Sporing Microsporidian Genomes: Polyploidy, Rearrangements, and Reduction

Amjad Khalaf

Amjad Khalaf, Kamil Jaron, Mark Blaxter, Mara Lawniczak
Wellcome Sanger Institute, Wellcome Sanger Institute, Wellcome Sanger Institute, Wellcome Sanger Institute

Microsporidia are eukaryotic, amitochondriate, spore-forming, obligately intracellular parasites with an extremely broad host range, infecting both metazoan and protozoan hosts. They have major public health and economic importance. For example, the microsporidian Nosema bombycis is the causative agent of the disease â€œpÃ©brineâ€• in farmed silkworms (Bombyx mori, Lepidoptera), which caused huge collapses in the silk industry across the globe in the 1800s. Additionally, the Anopheles-mosquito-infecting microsporidians Vavraia culicis and â€œMicrosporidia MBâ€• were shown to be associated with a reduction in Plasmodium transmission, forming a potential malaria control route.

Microsporidia remain difficult to culture, and isolating sufficient, clean genomic DNA from infected hosts remains equally challenging. As such, only around 50 out of nearly 2,000 microsporidian species have undergone whole-genome sequencing, typically from pooled infected hosts. The Darwin Tree of Life (DToL) offers an unrivalled opportunity to generate high-quality microsporidian genomes from a variety of under-sampled host species, as occasionally, infected hosts are sequenced. Here, I present over 30 new microsporidian genomes generated from single hosts sequenced by DToL, and microsporidian genomes generated from my independent sampling of 600 flying insects last year.

Using the genomes I generated, I showcase an assembly framework to deal with polyploidy and the complications of microsporidian

genomes, I explore the origins of polyploidy in Microsporidia, show the occurrence of structural rearrangements between the sub-genomes within individual microsporidians, and explore gene loss and gain (through innovation and horizontal gene transfer) across the group.

860 - Uncovering the genetic architecture of host plant shifts in Pierinae butterflies

Paula Fernandez-Begne

Paula Fernandez-Begne, Timothy O'Connor, Marcus Kronforst
University of Chicago, University of Chicago, University of Chicago

The evolutionary success of herbivorous insects is thought to be due to adaptation to novel chemical defenses in host plants, but the genetic architecture enabling host plant transitions is not well understood. Convergent host plant shifts, where two or more related lineages independently colonize the same plants, present a unique opportunity to associate genetic changes to dietary shifts. Pierinae butterflies are a diverse subfamily that rapidly diverged after the clade shifted to feed to Brassicales plants, which produce defensive glucosinolate secondary metabolites. Within Pierinae, multiple lineages independently colonized Brassicaceae, a Brassicales lineage that, in addition to ancestral glucosinolates, produces novel glucosinolates not found in close relatives. This replicated ecological transition allows us to investigate the genetic architecture of host plant adaptation through the lens of convergent evolution. To determine the number of independent transitions to Brassicaceae across Pierinae, we used a recently resolved Pieridae phylogeny and host-use records. We reconstructed proportional use of Brassicaceae across Pierinae and the probability of use vs non use at internal nodes. These analyses recovered two to four independent transitions to Brassicaceae within Pierinae. Although the dietary breadth of Pierinae is well characterized, taxonomically diverse genomic resources in this subfamily are lacking. To expand our dataset, we assembled and annotated highly complete draft genomes of 15 Pierinae and a close outgroup. The ancestral host use results and our expanded genomic dataset will allow us to leverage genomic scans to ask if convergence at the ecological level is underpinned by convergence at the genetic level.

1077 - REvolutionH-tl: Advancing Orthology Detection and Evolutionary Scenario Reconstruction

José Antonio Ramírez-Rafael

JosÃ© Antonio RamÃ-rez-Rafael, Maribel HernÃ¡ndez-Rosales
Center for Research and Advanced Studies, Irapuato Unit, University of Leipzig, Center for Research and Advanced Studies, Irapuato Unit

Understanding evolutionary relationships between genes and species is pivotal in deciphering genetic diversity, adaptation, and functional innovation across organisms. Orthology detection, crucial for inferring evolutionary scenarios, traces genes from common ancestors through speciation events. Accurate orthology detection is essential for inferring gene function and reconstructing biological processes' evolutionary trajectories. Here we present, REvolutionH-tl, a tool that addresses challenges in orthology detection and evolutionary scenario reconstruction. It efficiently identifies orthogroups, orthology, and paralogy relationships using a graph-based approach. By bypassing the need for gene and species trees, REvolutionH-tl maximizes computational efficiency and scalability. In this work, we showcase REvolutionH-tl's speed, accuracy, and versatility. Its graph-based methodology leverages sequence similarity for efficient orthology inference without resource-intensive procedures. Comparative analyses demonstrate REvolutionH-tl's superiority, particularly in orthology assignments and evolutionary scenario reconstruction. Ongoing developments aim to enhance REvolutionH-tl's capabilities and address current limitations. In conclusion, REvolutionH-tl offers great potential for applications in comparative genomics, evolutionary biology, and functional genomics, making it valuable across disciplines. Its implementation represents a significant advancement in bioinformatics, empowering researchers to unravel gene evolution and function complexities. Revolutionh-tl is freely available at

https://pypi.org/project/revolutionhtl/.

191 - How ecological interactions shape microbial mutation rates: a story of collective detoxification

Rowan Green

Rowan C Green, Rok KraÅ¡ovec, Christopher G Knight
University Of Manchester, University of Manchester, University of Manchester

The ultimate source of genetic variation, de novo mutagenesis, is a highly dynamic trait responsive to multiple ecological factors. Evolution therefore depends on the ecological environment not only for selection but also in determining the variation available in a population. One such environmental dependency is the inverse relationship between mutation rates and population density in many microbial species. Using dynamical computational modelling and in vitro mutation rate estimation we show that the negative relationship between mutation rate and population density arises from the collective ability of microbial populations to control concentrations of hydrogen peroxide (H2O2). Our in silico modelling approach to hypothesis generation allows us to efficiently explore the idea-space of possible mechanisms for this mutation rate plasticity through altering model structures and parameters. This generates select hypotheses, which we test using high-throughput mutation rate measurements alongside direct measurements of environmental H2O2. We demonstrate a loss of this density-associated mutation rate plasticity in Escherichia coli populations deficient in the ability to degrade H2O2 however, the reduction in mutation rate in denser populations can be restored in peroxide degradation-deficient cells by the presence of wild-type cells in a mixed population. These model-guided experiments provide a mechanistic explanation for density-associated mutation rate plasticity, applicable across all domains of life, and frame mutation rate as a dynamic trait shaped by microbial community composition.

275 - Accelerating evolution of T6SS resistance during experimental evolution

Kathryn MacGillivray

Kathryn MacGillivray, Tahrima Jubery, Erica Drummond, William Ratcliff, Brian Hammer
Georgia Institute of Technology, Georgia Institute of Technology, Georgia Institute of Technology, Georgia Institute of Technology, Georgia Institute of Technology

In polymicrobial communities, bacteria fight over nutrients and space. One common weapon is the Type VI Secretion System (T6SS), a nano-harpoon which injects toxins into neighboring cells. While much is understood about mechanisms of T6SS-mediated toxicity, less is known about mechanisms of defense against this attack. We previously subjected Escherichia coli populations to 30 rounds of attack (500 generations) by Vibrio choleraeâ€™s T6SS, with a period of exponential growth in between attacks. During these 500 generations, E. coli evolved evolve moderate resistance to T6 attack, through two independent genetic pathways, but this came with a substantial cost to their growth rate. Here, we doubled the duration of the experiment. Surprisingly, at least one E. coli lineage adapted even more rapidly during generations 500-1000, with resistance increasing 4.7 times as much as in the first phase, reaching an ultimate survival of 2400-fold higher than the ancestor. Moreover, our highly resistant E. coli does not grow slower than earlier isolates, breaking the trade-off between greater T6SS-survival and reduced growth rates. These results suggest that trade-offs make the early evolution of T6SS resistance difficult for bacteria that do not often encounter T6SS attack, while more sustained periods of interaction lead to an easier gain in survival. Our experiment highlights an interesting non-additivity of bacterial adaptation to T6SS attack. T6SS, relative to many other strongly-selective antimicrobial agents, is surprisingly robust, requiring a greater degree of sustained selection prior to the evolution of low-cost resistance.

359 - Expanding the Plant Microbiome

Sheila Roitman

Sheila Roitman, Haim Ashkenazy, Detlef Weigel

Max Planck Institute for Biology, Max Planck Institute for Biology, Max Planck Institute for Biology

Eukaryotic organisms harbour large communities of microorganisms forming an holobiont, considered to be a single ecological and evolutionary unit. In recent years, bacterial community dynamics and their effect on the plant holobiont have been the subject of many studies. However, little is known regarding the role that bacteriophages play in shaping those bacterial communities. In my work, I intend to set the basis for understanding the role of the microvirome in plant colonization and development, by studying Arabidopsis thaliana associated bacteria and phages, in laboratory and natural settings. I isolated lytic viruses of Pseudomonas viridiflava, an environmentally relevant plant-associated microbe. These bacteriophages are generalists, infecting more than one strain of P. viridiflava, including pathogenic and commensal bacteria. Lysogenic phages are ubiquitously found in the genomes of these microbes, and are likely playing a role in defense against lytic phages and other microbes. These findings will serve as the basis for a better understanding of the impact phages have on the bacterial community assembly and evolution in the plant.

388 - Diazotrophic endophytic community in Cordia dodecandra phyllosphere according to management in agroforestry systems

Miguel Angel López-Garrido

Miguel A LÃ³pez-Garrido, Miriam M Ferrrer, Yuri J PeÃ±a-RamÃ-rez, Natalia Y LabrÃ-n-Sotomayor, JosÃ© R Pacheco-Arjona, Hector Estrada-Medina

Universidad AutÃ³noma de Yucatan , Universidad AutÃ³noma de YucatÃ¡n, ECOSUR , ECOSUR , CONAHCYT-Universidad AutÃ³noma de YucatÃ¡n , Universidad AutÃ³noma de YucatÃ¡n

The phyllosphere represents the most extensive terrestrial habitat, approximately twice the land-earth area. Bacteria dominate the phyllosphere, playing an essential role in nitrogen fixation and hostâ€™s growth. Since management affects the diversity of the microbiota, assessing the composition of diazotrophic endophytic communities in trees within systems that differ in management practices can help us understand their role in nitrogen fixation. In Cordia dodecandra, the core microbiota comprises species with diazotrophic potential. However, the diversity of actors and their variation with management have not been assesed. For this purpose, microbial DNA was extracted from the leaves of two trees in a forestal system (reduced irrigation and weeding) and four in a silvopastoral system (goat grazing, star grass cover, besides irrigation and weeding). Additionally, the canopy cover was also measured. The DNA was amplified for the nifH gene by Qpcr, mass sequencing of the nifH genes, and V3-V4 region of the 16S rRNA. Alpha and beta diversity was analyzed for sequences. Similar Shannon and Chao1 index values were found in both systems. However, the communities composition differed between the systems. The relative abundance of nifH genes was positively correlated with canopy cover. Aureimonas, Quadrisphaera, Actinomycetospora, and Pseudokineococcus were enriched in the forestal system. Diazotrophic species identified (p > 99%) with nifH genes were Alcaligenes faecalis, Ensifer americanum, Methanosaeta concilii, Chroococcidiopsis thermalis, Hassallia byssoidea and Nostoc commune. A small proportion of the C. dodecandra phyllosphere microbiota is nitrogen-fixing, although the role of the majority of species systems remain unknown.

427 - Mixed Wolbachia infections resolve rapidly during in vitro evolution

Cade Mirchandani

Cade Mirchandani, Pingting Wang, Jodie Jacobs, Maximilian Genetti, Evan Pepper-Tunick, William T Sullivan, Russ Corbett-Detig, Shelbi Russell
Department of Biomolecular Engineering, University of California Santa Cruz, Department of Biomolecular Engineering, University of California Santa Cruz, Department of Biomolecular Engineering, University of California Santa Cruz, Department of Biomolecular Engineering, University of California Santa Cruz, Molecular Engineering Sciences Institute, University of Washington, Department of Molecular, Cell, and Developmental Biology, University of California Santa Cruz, Department of Biomolecular Engineering, University of California Santa Cruz, Department of Biomolecular Engineering, University of California Santa Cruz

Wolbachia pipientis is one of the most common endosymbionts found in arthropods, and it reached high prevalence in many taxa through its abilities to infect new hosts and their germlines. However, little is known about the ecological and evolutionary interactions that occur when a promiscuous strain colonizes an infected host. Here, we study what occurs when two strains come into contact in host cells following horizontal transmission and infection. We focus on the faithful wMel strain from Drosophila melanogaster and the promiscuous wRi strain from Drosophila simulans using an in vitro cell culture system. Mixing D. melanogaster cell lines stably infected with wMel and wRi revealed that wMel outcompetes wRi quickly and reproducibly. Furthermore, wMel was able to competitively exclude wRi, and we found that this was a nearly deterministic outcome, independent of the starting infection frequency. This competitive advantage was not exclusive to wMelâ€™s native cell background, as wMel also outgrew wRi in D. simulans cells. Overall, wRi is less adept at in vitro growth and survival than wMel and its in vivo state, revealing differences between cellular and humoral regulation. These attributes may underlie the observed low rate of mixed infections in nature and the relatively rare rate of host-switching in most strains. Our in vitro experimental framework for estimating cellular growth dynamics of Wolbachia strains in provides the first strategy for parameterizing endosymbiont and host cell biology at high resolution. This toolset will be crucial to our application of these bacteria as biological control agents.

476 - Adaptive Mechanisms of K1 Killer Toxin Resistance from a Sensitive-Killer Coevolution in Budding Yeast (S. cerevisiae)

Angelina G. Chan

Angelina G. Chan, Michelle Hays, Angela Hickey, Magdalena Pieczynska, Arjan de Visser, Gavin Sherlock
Department of Genetics, Stanford University, Department of Genetics, Stanford University, Department of Genetics, Stanford University, Institute of Genetics and Biotechnology, University of Warsaw, Laboratory of Genetics, Wageningen University, Department of Genetics, Stanford University

Killer yeasts, such as the K1 killer strain of S. cerevisiae, express a secreted anti-competitor toxin whose production and propagation requires the presence of two vertically-transmitted dsRNA viruses. Though killer yeasts are immune to their own toxin, infection results in a fitness cost in the absence of toxin-sensitive competitors. Previous experimental coevolution of killer yeast containing the K1 killer virus and sensitive (non-killer) yeast resulted in coadaptation, including emergence of toxin resistance in sensitive ancestor-derived populations. We identified a putative dominant gain-of-function missense mutation in the SSK1 gene, which is sufficient for toxin resistance and fit under laboratory conditions yet is present in a domain strictly conserved in natural isolates. Preliminary data suggest this mutation may impact interactions of Ssk1 with both the HOG (high osmolarity glycerol) and CWI (cell wall integrity) pathways. Ongoing analysis of two separate asymmetrically evolved populations, where the killer population was replaced with an ancestral stock every ~25 generations, suggests both the presence of alternate adaptive pathways independent of SSK1 and increased variability in the K1 resistance phenotype arising in the presence of a non-evolving competitor. Our future work aims to both further understand the mechanism(s) by which Ssk1 mediates toxin resistance, including in fungi associated with pathogenicity such as C. auris (which requires SSK1 for multidrug resistance), and the mechanisms of coadaptation or genetic conflict arising between viral and host genomes in

counterevolving killer yeast populations.

559 - Quantifying the stability of microbiomes and the timescale of antibiotic perturbation with FAVA

Maike L Morrison

Maike L Morrison, Katherine S Xue, Noah A Rosenberg

Stanford University, Stanford University, Stanford University

Most statistical methods for microbial data tend to analyze just one or two microbiome samples at a time, limiting their capacity to measure complex dynamics across multiple samples. For example, such methods are ill-suited to measuring the stability of microbial communities over long timescales, the heterogeneity of microbiomes across multiple host individuals, and the repeatability of microbial community assembly. To assess these aspects of microbial community dynamics, we need a measure of compositional variability across two or more microbiome samples at once. We introduce FAVA, an FST-based Assessment of Variability across vectors of relative Abundances. FAVA extends the diversity-partitioning framework of the population-genetic statistic FST, with microbiome samples playing the role of â€œpopulationsâ€• and microbial taxa playing the role of â€œalleles.â€• We demonstrate FAVA in a longitudinal analysis of the gut microbiomes of 22 healthy adults taking an antibiotic. We find that the antibiotic destabilizes microbial communities, using FAVA to quantify both the magnitude and the timescale of the antibiotic perturbation. In particular, we show that FAVA can identify a return to stability regardless of whether the community returns to its initial compositional state. We have implemented FAVA in an R package useful for quantifying the evolutionary and ecological dynamics of microbial communities over time, space, or host individuals.

590 - Artificial evolution of plant-associated synthetic communities

Niklas Kiel

Niklas Kiel, Ruben Garrido-Oter

Max-Planck Institute for Plant Breeding Research, CEPLAS, Max-Planck Institute for Plant Breeding Research, CEPLAS

Root-associated microbial communities are essential for plant health, providing beneficial functions such as nutrient mobilization and pathogen protection. Recent studies have highlighted the significance of microbial host preference in shaping these communities, with native bacteria often outcompeting those from different plant species. These findings suggest that microbial fitness is influenced by the host species. However, which plant factors and ecological or molecular mechanisms drive microbial host preference remain unclear. Here, we present the results of an artificial evolution experiment conducted over 16 successive plant generations, where synthetic communities are repeatedly exposed to model plants Arabidopsis thaliana and Lotus japonicus. Analyses of amplicon and shotgun sequencing after every host generation reveal that synthetic bacterial communities undergo adaptive changes in their genomes, resulting in an increased capacity for root colonization in a host-species specific manner. Our findings showcase the feasibility of experimental evolution approaches in-planta using complex synthetic communities, revealing a previously unknown capacity for rapid adaptation to their plant host. This evidence of adaptation highlights the evolutionary potential within root-associated microbiota and its implications for improving our understanding of the mechanisms driving plant-microbiota interactions. Not only does this study contribute to our understanding of microbial ecology but also has potential applications in agriculture, where leveraging microbial adaptation could improve crop resilience and productivity. Future research, will further elucidate the genetic basis of these adaptations, offering new insights into microbial evolution and broader implications for ecosystem dynamics and agricultural practices.

609 - The Salmon Microbial Genome Atlas enables novel insights into bacteria-host interactions via functional mapping.

Arturo Vera Ponce de Leon

Arturo Vera-Ponce de Leon, Matthias Hoetzinger, Tim Hensen, Shashank Gupta, Bronson Weston, Sander M Johnsen, Jacob A Rasmussen, Knut Rudi, Lars Snipen, Christian RenÃ© Karlsen, Morten T Limborg, Stefan Bertilsson, Ines Thiele, Torgeir R Hvidsten, Simen R Sandve, Phillip B Pope, Sabina Leanti La Rosa
Norwegian University of Life Sciences, Norwegian University of Life Sciences, Swedish University of Agricultural Sciences, University of Galway, Norwegian University of Life Sciences, University of Galway, Norwegian University of Life Sciences, University of Copenhagen, Norwegian University of Life Sciences, Norwegian University of Life Sciences, NA, Nofima, University of Copenhagen, Swedish University of Agricultural Sciences, University of Galway, Norwegian University of Life Sciences, Norwegian University of Life Sciences, Norwegian University of Life Sciences, Norwegian University of Life Sciences

Our knowledge about the salmon (Salmo salar) gut microbiota is limited, despite the importance of salmonid aquaculture as one of the most expanding food-production sectors worldwide. This hampers the study of functional interactions between host and bacteria, and the development of dietary strategies aimed at the manipulation of the gut microbiota to improve salmon welfare. Furthermore, the lack of salmon gut-derived microbial genome data makes it difficult to interpret functional omics dataset, as much of the data do not match the available reference bacterial genomes. Combining published and unpublished data, here we present the first version of the Salmon Microbial Genome Atlas (SMGA), originating from fish reared both in fresh water and saltwater. The SMGA comprises 211 high-quality bacterial genomes, recovered by cultivation (n=131) and gut metagenomics (n=80). These genomes were taxonomically assigned into 14 different orders, including 35 distinctive genera and 29 potentially novel species. Benchmarking the SMGA, as a database for metatranscriptomics and metabolomics experiments, functionally characterized key populations in the salmon gut that were detected in vivo. This included the ability to degrade diet-derived fibers and release vitamins and other exo-metabolites with known beneficial effects, which were validated by in vitro cultivation and untargeted metabolomics. Together, the SMGA enables high resolution functional insight into salmon gut microbiota with relevance for salmon nutrition and health.

619 - An Integrated Pipeline of Machine Learning Techniques for Analysing Microbial Communities

Adamandia Kapopoulou

Adamandia Kapopoulou, Ianis Vilela, Russel Jasper, Jessica Rieder, Stephan Peischl, Irene Adrian-Kalchhauser, Claudia Bank
Division of Theoretical Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics, Division of Theoretical Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, Division of Theoretical Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics, Division of Theoretical Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics, Department of Infectious Diseases and Pathobiology, Vetsuisse Faculty, Institute for Fish and Wildlife Health, University of Bern, Interfaculty Bioinformatics Unit, University of Bern, Swiss Institute of Bioinformatics, Department of Infectious Diseases and Pathobiology, Vetsuisse Faculty, Institute for Fish and Wildlife Health, University of Bern, Division of Theoretical Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics

The significance of microbial communities in maintaining natural ecosystems at equilibrium and human and other speciesâ€™ health is indisputable. With the emergence of numerous microbial datasets, there is growing interest in characterising and understanding the factors that influence microbial composition. Specifically, there is a need for robust methods that relate differences in relative microbial abundance to environmental factors.
Recirculating Aquaculture Systems (RAS) are controlled ecosystems that provide an ideal setting for studying microbial communities. They bridge the gap between lab-controlled environments and natural ecosystems. In RAS, we can trace the colonisation of different

compartments over time and try to link observed disease outbreaks to the presence of pathogenic microbes.
In a study of the RAS microbiome across space and time, we conducted ultra-deep sequencing of 496 samples from various compartments across six RAS farms at different time points. To analyze this complex dataset, we developed a statistical pipeline leveraging state-of-the-art machine learning tools. Our pipeline focuses on distinguishing between adjacent microbial communities and the environmental factors shaping them, starting with the microbial relative abundance data commonly obtained from metagenomic studies.
Our findings validate previous research indicating the significance of pH as a driving force behind microbial variability. Furthermore, our pipeline's versatility was demonstrated through successful testing on other compartmentalized microbial communities, such as from longitudinally sampled ruminant guts.

621 - Substrate drives densely connected microbial community assembly in culturally diverse, traditional fermented foods.

Arya Gautam

Arya Gautam, Rahgavi Poopalarajah, Anique R Ahmed, Bimarsh Rana, Tsedenia W. Denekew, Nakyeong Ahn, Lina Utenova, Yadhu Sudan Kunwar, Justin L. Sonnenburg, Aashish R Jha
Sabitri Sciences, New York University Abu Dhabi, New York University Abu Dhabi, Integrated Rural Health Development and Training Center, New York University Abu Dhabi, New York University Abu Dhabi, New York University Abu Dhabi, Integrated Rural Health Development and Training Center, Stanford University, New York University Abu Dhabi

The beneficial effects of fermented foods for human and planetary health is increasingly being appreciated. Throughout history, indigenous peoples around the globe have developed fermentation techniques for food preservation, improved digestibility, and flavor enhancement. In contrast to industrial fermentation which require starter cultures, traditional approaches rely on naturally-occurring microbes to spontaneously ferment foods. The process of microbial community formation in traditionally fermented non-European foods remains understudied. Here, we characterized the bacterial and fungal communities of diverse plant and animal-based fermented foods from Nepal, South Korea, Ethiopia, and Kazakhstan that were traditionally prepared for household consumption. Traditional fermented foods are reservoirs of diverse and densely interconnected networks of bacteria and fungi that reside on food surfaces. In addition to the canonical fermenters â€“ lactic acid bacteria (LABs), Bacillales, and Saccharomycetales (yeasts) â€“ that are commonly used in commercial fermentation, these foods comprise of additional microbes that contribute to fermentation process or in aesthetics (taste, flavor) as well as environmental microbes. Environmental bacteria in fermented foods vary by geography but substrates shape the bacterial community as well as bacteria-bacteria and bacteria-fungi interactions in traditional fermented foods. Plant-based ferments harbor bacterial communities that putatively degrade carbohydrates while animal-based fermented foods contain bacteria enriched with protein degradation pathways. Our results indicate that microbiome assembly in fermented foods is a highly dynamic process and this underscores the need to develop an ethical and equitable framework for investigating the microbial communities of traditional fermented foods and their therapeutic potential.

758 - Acetic acid bacteria: an evolutionary story of the quest for sugar

Ana Cuesta-Maté

Ana Cuesta-MatÃ©
Globe Institute, University of Copenhagen

Acetic acid bacteria are generally considered beneficial microbes for humans and used to produce several fermented foods and

beverages. They dominate the microbial communities in acetic fermentation, like vinegar or kombucha. Humans have historically domesticated these bacteria to obtain some of the ferments that we consume, but the question remains on how these microbes got there in the first place. Fermented food researchers find it difficult to sample these bacteria in natural habitats, but our work shows that some of these bacteria are found in social insect guts and environment, in the food storages of their nests, and in the flowers and fruits social insects visit. These bacteria are adapted to somewhat harsh environments (low pH, high osmolarity) where simple sugars are readily available, and so, colonized by many different types of microbes. How has life within insect microbiomes made these bacteria ideal to colonize host-controlled environments, like our nutritious ferments? This project focuses the adaptions of acetic acid bacteria to these dynamic communities, using their presence in fermented food and insect microbiomes to follow their evolution within microbial communities. Genomic adaptions have been observed in acetic acid bacteria domesticated by insects, namely, genome reduction and base compositional bias towards an AT-rich genome. Similar patterns can be observed in acetic acid bacteria in our ferments, but are there further adaptions to life within insect microbiomes that gave acetic acid bacteria an advantage to become our fermented food microbes? Could insect metabolic pathways be already trained for food fermentation?

773 - An interdisciplinary exploration of microbial responses to iodine metabolism in kelp

Maximiliana Bogan

Maximiliana Bogan, Cathy Pfister
University of Chicago, University of Chicago

Kelp-associated bacteria present a metabolic and evolutionary puzzle. Kelp are large brown seaweeds that serve as foundation species in numerous ecosystems, and healthy kelp possess dense, diverse microbiomes. However, kelp also efflux large amounts of iodine upon exposure to oxidative stress, and some of these iodinated compounds are likely potent antibiotics. So how do microbial communities persist, when their hostâ€™s iodine metabolism may put them in danger? Does the microbiome respond to this metabolism at a community level, an individual level, or both? By combining multi-omics techniques with targeted growth assays, we will characterize: 1) how overall microbial gene expression responds to iodinated compound emission from kelp, and 2) the direct fitness effects of iodinated compounds on individual members of the community. Dissecting the response of specific microbial taxa vs. the whole microbiome will help us understand whether bacterial communities respond to a specific environmental stressor as expected, based on individual phenotypes. We will first apply an oxidative stressor to freshly collected kelp blades, and sample for transcriptomic and metabolomic analysis. By comparing microbial transcript abundances across different concentrations of iodinated compounds, we will identify what genes are up- or down-regulated in the bacterial community response. We will then expose whole-genome-sequenced kelp-associated bacteria to iodinated compounds in the lab, to measure how bacteria with known genotypes are affected by environmentally relevant chemicals. These results will elucidate the genetic mechanisms that underlie a diverse microbial communityâ€™s response to the unique iodine metabolism of kelp.

800 - Expanded Geographic Distribution for Two Legionella pneumophila Sequence Types of Clinical Concern

JENNAFER HAMLIN

Jennafer A. P. Hamlin, Natalia A. Kozak -Muiznieks, Jeffrey W. Mercante, Lavanya Rishishwa, Emily T. Norris, Anna B. Gaines, Maliha K. Ishaq, Jonas M. Winchell, Melisa J. Willby
Division of Bacterial Diseases, Centers for Disease Control and Prevention, Division of Bacterial Diseases, Centers for Disease Control and Prevention, Division of Bacterial Diseases, Centers for Disease Control and Prevention, Applied Bioinformatics Laboratory, Applied

Bioinformatics Laboratory, Applied Bioinformatics Laboratory, Division of Bacterial Diseases, Centers for Disease Control and Prevention,, Division of Bacterial Diseases, Centers for Disease Control and Prevention,, Division of Bacterial Diseases, Centers for Disease Control and Prevention,

Legionella pneumophila serogroup 1 sequence types (ST) 213 and 222, closely related variants, were first detected in the early 1990s in the Midwest US and subsequently in the Northeast US and Canada. These STs are associated with both healthcare- and community-acquired Legionnaires' Disease (LD). We aimed to assess the trend in cases linked to these STs and identify any genetic distinctions. Analyzing data from the Centers for Disease Control and Prevention (CDC) spanning 1992 to 2020, we found a yearly increase of 0.15 cases of ST213/222-associated LD across the Mountain West to Northeast region. Most cases occurred in Michigan (26%), New York (18%), Minnesota (16%), and Ohio (10%). ST222 caused at least five LD outbreaks in the US between 2002 and 2021, while no outbreaks were attributed to ST213. Genetic analysis of 230 isolates revealed high nucleotide identity within each ST (99.92% for ST222, 99.9299.92% for ST213), indicating low diversity within and between these types with a minimum of 99.5% and a maximum of 99.87%. Despite identifying genomic features like plasmids and CRISPR-Cas systems, no specific association explained the increasing prevalence and spread of ST213 and ST222. However, our study highlights the expanded geographical distribution of these STs in the US.

836 - The influence of subsistence strategies on oral microbiome composition and function in ancient populations

Sarah J Johnson

Sarah J Johnson, Tanvi P Honap, Cara Monroe, Alan Leventhal, Monica V Arellano, Douglas J Kennett, Keith M Prufer, Cecil M Lewis
Laboratories of Molecular Anthropology and Microbiome Research, University of Oklahoma, Laboratories of Molecular Anthropology and Microbiome Research, University of Oklahoma, Laboratories of Molecular Anthropology and Microbiome Research, University of Oklahoma, Center for the Ethics of Indigenous Genomic Research , San Jose State University, Muwekma Ohlone Tribe, Muwekma Ohlone Tribe, University of California, University of New Mexico, Center for Stable Isotopes, Laboratories of Molecular Anthropology and Microbiome Research, University of Oklahoma, Center for the Ethics of Indigenous Genomic Research

Through the analysis of shotgun metagenomes recovered from archaeological dental calculus, we present an investigation into the effects of diet on oral microbiome composition and function in human populations from the pre-European Contact Americas. These include ancestral Ohlone individuals from California, USA, and ancestral Maya individuals from rock shelters in Belize occupied for over 10,000 years. MetaPhlAn4 and HUMAnN3 were used to estimate taxonomic composition and functional potential, respectively, and SourceTracker analysis and DNA damage patterns were used to authenticate recovery of ancient oral metagenomes. In the ancestral Maya, the transition from a pre-agricultural to maize-based diet is expected to mark a change in carbohydrate metabolism. In this study, we demonstrate that this shift in subsistence pattern was accompanied by a significant reduction in abundance of the oral bacterium, Pseudoramibacter alactolyticus, estimated to be the primary contributor to carbohydrate metabolism. Following maize adoption, several Streptococcus species, including known opportunistic pathogens implicated in infective endocarditis, became dominant contributors to carbohydrate metabolism. The ancestral Ohlone, primarily hunter-gatherers with a starch-rich diet of acorns and seeds, did not have a dominance of Streptococcus species contributing to carbohydrate metabolism, but did show a significantly higher abundance of microbial gene families involved in starch metabolism, such as alpha-amylase and starch phosphorylase, as compared to pre-staple maize ancestral Maya and other ancient agricultural populations. These data reveal a complex metabolic response to dietary changes highlighting the interplay between diet, microbial ecology, and health over the course of human evolution.

900 - How does nutrient environment and varying mutational constraint impact regulatory rewiring in

bacteria?

Louise M Flanagan

Louise M Flanagan, James S Horton, Tiffany B Taylor
University of Bath, University of Manchester, University of Bath, University of Bath

Gene regulatory networks are fundamentally important for organism development and survival, ensuring that the right complement of genes is expressed in the right place at the right time. Despite their critical role, they are also subject to change, and many factors drive their evolution.

Here I consider how environmental differences can pull evolution down different mutational paths even when selecting for the same phenotype, and the effects of varying mutational constraint and accessibility.

Using an established model system in the bacterium Pseudomonas fluorescens, where I re-evolve a lost swimming phenotype (due to flagellar regulator deletion) through starvation selection in different nutrient environments, I show how mutational spectra vary in complex and minimal nutrient environments. By introducing a strong mutational hotspot that restores swimming, I compare the effects of strong and reduced mutational constraint to regain the swimming phenotype and the consequences of this. Having a mutational hotspot can lead to faster restoration of a swimming phenotype, but it can prevent access to other viable mutations that produce faster swimmers. I also explore how the swimming phenotype varies between mutants and investigate the apparent loss of swimming when placed in an alternative nutrient environment, and how this links to regulatory architecture and gene expression.

This work highlights the importance of considering the environment and mutational accessibility in evolutionary studies. Incorporating these variables into predictive evolution models is crucial for gaining a comprehensive understanding of the evolutionary dynamics of regulatory networks.

964 - Phylo-pangenomics: read mapping and variant detection in huge microbial datasets guided by pandemic scale evolutionary history

Alexander Kramer

Alexander M Kramer, Nicolas M Ayala, Lily M Karim, Sumit Walia, Harsh Motwani, Alan Zhang, Yatish Turakhia, Russell Corbett-Detig
University of California, Santa Cruz, University of California, Santa Cruz, University of California, Santa Cruz, University of California, San Diego, University of California, San Diego, University of California, Santa Cruz, University of California, San Diego, University of California, Santa Cruz

Pangenomes, which encode all genetic variation in related genomes, are an increasingly powerful tool in evolutionary genomics. Current pangenome methods catalog variation among observed haplotypes, but neglect the evolutionary processes that generated this variation. In the era of pandemic-scale genomics, huge datasets reflecting microbial evolution are common and not well-served by existing pangenome methods. We present an alternative that, in addition to known variation, also encodes inferred molecular evolutionary history of samples in a phylogenetic tree. Tree pangenomes are compact and accommodate millions of complete microbial genomes, scaling to many more individuals than existing haplotype graph-based approaches. We also developed panmap, a variant caller and genome assembler guided by reference tree pangenomes. From short or long reads, panmap finds a close relative in a tree pangenome with k-mer sketching, aligns reads to that sequence or inferred ancestor, computes genotype likelihoods using a mutation spectrum prior, and optionally generates a consensus assembly of the input sample. We show that entire global phylogenomic datasets for diverse pathogens

like Mycobacterium tuberculosis and SARS-CoV-2 can be compressed into reference tree pangenomes and used with panmap to improve variant calling and assembly pipelines for these species. Incorporating evolutionary relationships into pangenomic analyses can significantly enhance the accuracy and efficiency of identifying genetic variants and assembling genomes. This approach holds promise for advancing our understanding of microbial evolution and improving high-throughput analyses across domains like clinical genomics and public health.

970 - Building a reference genomic database of oral microbes for high-resolution oral microbiome analyses

Gabriel Yaxal Ponce Soto

Gabriel Yaxal Ponce Soto, Nicolas Rascovan

Institut Pasteur, Universitel• de Paris CitÃ©, CNRS UMR 2000, Microbial Paleogenomics Unit, Institut Pasteur, Universitel• de Paris CitÃ©, CNRS UMR 2000, Microbial Paleogenomics Unit

The oral microbiome situates at the interface between the external environment and the host, exerting significant influence on both local and systemic health. Despite advancements in metagenomic studies, which have provided insights into the taxonomic composition of the oral microbiome, there remains a substantial gap in understanding its variation across various levels (including pangenomic, strain diversity, and functional) and scales (ranging from individual to population, temporal, and geographical). While k-mer-based approaches have been widely employed to classify sequencing reads taxonomically, their utility for more in-depth analyses at the above mentioned levels and scales is limited. To move forward into this direction, we constructed a comprehensive and non-redundant database of oral microbial genomes to be used with mapping-based strategies. Our database and approach allow to gather whole genome data of most species in the oral microbiome, which can be used to i) have better accuracy in detecting taxa with strain-level resolution, ii) perform phylogenomic and phylogeographic analyses, and iii) explore pangenome composition of each species and its genetic functional variation. A main advantage of this strategy is that it also allows analyzing ancient oral microbiomes (e.g., ancient DNA from archaeological dental calculus), which are often too degraded to perform de novo assemblies, thus providing a temporal depth to investigate changes over time. Through these efforts, we aim to deepen our understanding of the oral microbiome's dynamics and composition, to investigate the evolutionary trajectories of its associated species, and its implications for human health and disease.

989 - Spread or dead? Is gene regulatory rewiring to resurrect motility a quirk of our model system, or a more general occurrence across a broad range of microbes?

Mitchell T Reynolds

Mitchell T Reynolds, Tiffany B Taylor

University of Bath, University of Bath

Previous work in the Taylor lab has shown rapid and repeatable rewiring between the nitrogen and flagellar networks in Pseudomonas fluorescens. In the absence of the flagellar master regulator (FleQ), NtrC (response regulator associated with nitrogen uptake) is co-opted to resurrect flagellar motility. Both regulators are part of the same protein family and interact with ï•³54 to initiate transcription. My research aims to explore the role of regulatory network architecture, and the wider genome, in determining gene regulatory rewiring pathways, utilising the rescue of flagellar motility as a model system. We deleted the master regulator for flagella control in Pseudomonas aeruginosa (FleQ), Caulobacter crescentus (CtrA) and Salmonella enterica (FlhDC). I also plan to delete the flagella master regulators in Escherichia coli (FlhDC), Bacillus subtilis (SwrA) and Bordetella bronchiseptica (FliA). These microbial species cover a broad range of genetic backgrounds and environmental niches and regulation of flagellar motility is under the control of different sigma factors.

So how do these respond to being placed in a swim or die scenario? Initial results in close relative to P. fluorescens, P. aeruginosa, has demonstrated that motility resurrection occurs as soon as 48 hours after inoculation onto soft agar, and Sanger sequencing of candidate genes suggests that rewiring also occurs through NtrC. However, after being left on soft agar for 8 weeks, distant relative of P. fluorescens, C. crescentus, failed to resurrect motility. Using comparative analysis, I aim to determine the factors that facilitate or constrain rewiring of regulators across different genetic backgrounds.

1037 - The lipopolysaccharide specificity of mammalian TLR4/MD-2 complexes has fluctuated randomly over evolutionary time

Sophia Phillips

Sophia Phillips, Maggie Barry, Laurel Moneysmith
University of Oregon, University of Oregon, University of Oregon

The animal innate immune system continuously evolves to recognize and respond to microbes. Toll like receptor-4 (TLR4) and co-receptor MD-2 form a complex, found on the surface of immune cells, that mediates inflammation in response to lipopolysaccharides (LPS) from the outer membrane of gram-negative bacteria. Different bacteria produce LPS that vary in number of acyl chains, which allow them to activate or evade TLR4/MD-2. We hypothesized that host species maintain specificity for their unique bacterial landscapes via coevolution of TLR4/MD-2 that recognize these LPS molecules. To test this hypothesis, we characterized LPS specificity of TLR4/MD-2 complexes across mammals. We found that, while the ability to recognize LPS variants with six or seven acyl chains is strongly conserved, recognition of LPS with four or five acyl chains was repeatedly gained and lost during evolution. A simple explanation for this would be that natural selection altered specificity as microbial communities changed over time. Consistent with this hypothesis, the co-receptor MD-2 (where LPS binds) has a higher ratio of sites with elevated dN/dS compared to TLR4 (6.3% vs 2.7%, respectively). However, mutagenesis of sites in the MD-2 binding pocket experiencing diversifying selection were not sufficient to switch ligand specificity. Instead, the genetic background was a stronger predictor of effect than identity of mutation. This suggests that the primary evolutionary constraint on mammalian TLR4/MD-2 is to recognize larger LPS molecules, while the ability to recognize small LPS molecules fluctuates stochastically over evolutionary time.

1089 - Genomic Insights into the Evolutionary Trajectory of Antibiotic Resistance in Klebsiella spp. and Escherichia coli: A Study from Central Mexico (2020-2023)

Miriam Sarahi Lozano Gamboa

Miriam S Lozano, Brian E MuÃ±oz, Anastacio Palacios , Corina D CEAPÄ,
State Laboratory of Aguascalientes (LESPA), State Laboratory of Aguascalientes (LESPA), State Laboratory of Aguascalientes (LESPA), Aguascalientes Health Secretariat, National Autonomous University of Mexico

This study explored the evolutionary dynamics of multi-drug strains of Klebsiella spp. and Escherichia coli isolated from Aguascalientes hospitals between 2020 and 2023. We identified 65 multidrug-resistant strains. These strains were subjected to biotyping and sequencing to elucidate their genomic landscapes and antibiotic resistance profiles. Comprehensive antibiotic susceptibility testing was conducted using broth microdilution, including the detection of Î²-lactamases and carbapenemases. Our study revealed a significant increase in the diversity and prevalence of resistance genes among multidrug strains of Klebsiella spp. and Escherichia coli over the years, particularly affecting third- and fourth-generation antibiotics. Genotype-phenotype correlations identified discordances, indicating potential adaptation mechanisms in response to antibiotic exposure, whereas phylogenetic analyses demonstrated temporal divergence and evolutionary relationships within the strain populations. Comparisons with pre-pandemic samples emphasized the accelerated pace of

genetic evolution under antimicrobial pressure. Additionally, biosynthetic gene cluster (BGCs) analysis provided insights into secondary metabolite production. Our study provides a comprehensive investigation of the evolution of antibiotic resistance in a hospital environment, demonstrating how genetic makeup alters and resistance genes evolve. These findings underscore the importance of monitoring genetic alterations in bacteria to inform medical treatment strategies and to devise effective interventions. Our study highlights the pivotal role of genomic surveillance in shaping clinical strategies and in promoting the development of targeted interventions. By understanding the evolution of bacterial resistance, health care professionals can make informed decisions and take proactive measures to combat the escalating threat of antibiotic resistance.

1117 - Genomic insights into microbial community assembly and competitive interactions in a 78-strain model community

Marisol Navarro-Miranda

Marisol Navarro-Miranda, VÃ-ctor M Higareda-Alvear, Obed RamÃ-rez-SÃ¡nchez, IvÃ¡n Anguiano-Aguirre, Maribel Rosales-HernÃ¡ndez, Gabriela Olmedo-Alvarez

CINVESTAV, Solena Mexico, Solena, Mexico, CINVESTAV, CINVESTAV, CINVESTAV

Understanding the dynamics of microbial communities is crucial for elucidating the evolutionary processes shaping their composition and function. This study delves into the interplay between genomic features and ecological interactions within bacterial communities to discern their implications in microbial evolution. Taking advantage of our model community of 78 strains from sediments of a biodiverse lagoon system in Cuatrocienegas, Coahuila, Mexico, composed predominantly of Bacillota strains, we conducted a broad analysis integrating phylogenomics, ecological interactions network construction, and genomic characterization. Our research elucidated the important role of taxonomic features in shaping the ecological network structure, with distinct taxonomic groups exhibiting differential distribution patterns of Biosynthetic Gene Clusters (BGCs). Notably, the identified BGCs showed lineage specificity, underscoring the evolutionary constraints within microbial lineages. Furthermore, taxa identified as antagonists displayed potential for antibiotic activity, shedding light on their ecological roles as antibiotic producers and contributors to antibiotic resistance dynamics. By linking ecological and genomic data, our findings underscore the necessity of adopting a comprehensive approach to understand the evolutionary trajectories and functional implications of microbial communities. This integrative framework not only deepens our understanding of community assembly processes but also provides valuable insights into potential environmental resistance mechanisms and BGCs reservoirs. Through elucidating the intricate interplay between genomic attributes and ecological interactions, our study advances the comprehension of microbial evolution within the broader context of community dynamics. In doing so, we aim to tackle the fundamental question of whether the impact of individual microbial constituents surpasses that of the entire community.

83 - Hidden in Plain Sight; Staphylococcus argenteus as an obscured public-health risk.

Maitiú Marmion

Maitiu Marmion, Cheryl P Andam

Department of Biological Sciences, University at Albany State University of New York, Department of Biological Sciences, University at Albany State University of New York

Staphylococcus aureus is one of the best characterized bacterial species, due to its prevalence in the human holobiont and its role in antibiotic-resistant infection. However, it has become evident in recent years that Staphylococcus argenteus, which had previously been considered as part of Staphylococcus aureus Clonal Complex-75, poses a novel risk to human health. This study aimed to assess the global prevalence, phylogenetic relationships, and genomic features of S. argenteus. Ecological sources from which publicly available

sequences of S. argenteus has been isolated include sick and healthy humans, animals, food, and environment, with most samples found in Asia and Europe. We identified four major sequence types (ST2250,ST2198,ST2793,ST1223) which vary in the number and types of antibiotic resistance and virulence genes present per genome. 37 virulence genes and 6 antimicrobial resistance genes were detected in all genomes. Type 7 secretion systems were significantly more present in non-pathogenic isolates, while distinct cap8 capsular and set exotoxin profiles were found among different STs and ecological sources. The mecA gene, which confers broad resistance to beta-lactams, was detected in 27.86% of S. argenteus genomes. Time-calibrated phylogenetic analysis revealed the time of the most recent common ancestor was 281 C.E., followed by multiple diversification events. This study provides interesting insights into the spread and diversity of a lesser-known pathogen which has often been eclipsed by S. aureus. Our results highlight the need for broader and systematic surveillance efforts to identify reservoirs, resistance determinants, and clinical pathologies of S. argenteus.

107 - On the origins of the BA.2.86 ('Pirola') SARS-CoV-2 lineage

Marina Escalera-Zamudio

Marina Escalera-Zamudi, Cedric CS Tan, Lucy van Dorp, FranÃ§ois Balloux
UCL Genetics Institute, Department of Genetics, Evolution & Environment, University College London, END-VOC consortium, UCL Genetics Institute, Department of Genetics, Evolution & Environment, University College London, END-VOC consortium, UCL Genetics Institute, Department of Genetics, Evolution & Environment, University College London, END-VOC consortium, UCL Genetics Institute, Department of Genetics, Evolution & Environment, University College London, END-VOC consortium

Global SARS-CoV-2 surveillance enabled fine-scale monitoring of divergent variants that display unusually high number of mutations presumably conferring enhanced transmissibility and/or immune evasion, and that may lead to epidemic waves. One example is the Omicron-descending BA.2.86 sublineage ('Pirola'), first detected in mid-2023. With 34 'unique' non-synonymous mutations in the Spike protein, and a predominant circulation in the USA/Europe, it was classified as a variant of interest (VOI) in November 2023. By February 2024, the BA.2.86 diverged into multiple sublineages, and dominated worldwide (BA.2.86*/JN.1*).

Using multiple fine-scale evolutionary analyses applied to selectively subsampled large-scale datasets, we explore the evolutionary process underlying the emergence of the BA.2.86*/JN.1* lineages. Our results reveal a basal cluster of hybrid BA.2/BA.2.86-like genomes displaying some (but not all) BA.2.86-specific lineage defining mutations. Hybrid genomes displayed synchronous circulation in multiple countries prior to the emergence of the BA.2.86. Additionally, we find evidence for diversifying selection, and for putative BA.2 X BA.2.75 recombination in the branch leading to the BA.2.86*/JN.1* clade.

The dominant hypothesis explaining the emergence of divergent SARS-CoV-2 variants postulates that these may arise through accelerated evolution in chronic infections. Our findings point out to cryptic co-circulation and gradual evolution, incompatible with a single emergence associated with prolonged shedding. Our results contribute to understanding the evolutionary mechanisms behind the emergence of (some) divergent SARS-CoV-2 variants, and underscore the importance of adequate subsampling strategies when interpreting initial phylogenetic placements of emerging lineages.

146 - Phylogenetic analysis of Mycobacterium tuberculosis sequences from Mexico in the global situation

Ikuri Alvarez-Maya

Ikuri Alvarez-Maya, Manuel Garcia-Ulloa, Jaime Martinez-Urtaza

Medical and Pharmaceutical Biiotechnology, Center for Research and Assistance in Technology and Design of the State of Jalisco (CIATEJ),, Department of Genetics and Microbiology, Faculty of Biosciences, Autonomous University of Barcelona, Department of Genetics and Microbiology, Faculty of Biosciences, Autonomous University of Barcelona

The global spread of Mycobacterium Tuberculosis (MTB) remains a major public health issue. Since data on resistance to anti-tuberculosis drugs are collected, the country coverage of such data is much more limited in the country as Mexico. We study the population structure of Mexican populations, through the analysis of the introduction nature of genotypes, by searching for clusters and genomic variation indicating possible transmission events involved in spreading Mycobacterium tuberculosis multidrug resistant (MDR-TB) behavior in Mexico.
MTB raw Illumina short reads were collected from projects PRJEB30933, PRJNA824124, and PRJEB44165 on the website www.ebi.ac.uk/ena repository. Additionally, the scaffolds from project PRJNA751891 at ww.ncbi.nlm.nih.gov. were also collected. Multidrug resistance profiles, lineages, and sublineages were assigned to strains in silico with tb-profiler v4.3.0.
The alignment of all Mexican and worldwide strains yielded a total of 31,951 variable sites which were used for the phylogenetic reconstruction. Lineage 4 had the highest frequency, with 336 (66.8%) of the global MTB complete genomes belonging to this lineage. This was followed by lineage 2, which accounted for 89 (17.7%) of the global MTB complete genomes. Interestingly, our in-silico analysis of drug resistance worldwide revealed a high percent of strains belonging to lineage 2 with Pre-Extremely Drug Resistant TB phenotype.
Overall, Mexico has a percentage greater than 90% of L4 strains, as well a particular sublineage with 90% percent of MDR-TB was found endemic. Thus, our study contributes to clarifying the implications for the evolution of this pathogen in Mexico and worldwide.

218 - Quantifying virus-driven adaptation in Myotis genomes using the McDonald Kreitman test

Genavieve Kathryn Gray-Sandoval

Genavieve K Gray-Sandoval, M. Elise Lauterbur, Juan Manuel Vazquez, Peter Sudmant, Lucie Etienne, David Enard
University of Arizona, University of Arizona, University of California Berkeley, University of California Berkeley, CIRI, Ecole Normale SupÃ©rieure de Lyon, UniversitÃ© Claude Bernard Lyon 1, University of Arizona

Bats are a natural reservoir for a multitude of viruses including coronaviruses, filoviruses, henipaviruses, and retroviruses. Bats are thought to be a major host of past and future zoonoses, so it is important to understand their evolutionary history. We can develop a better understanding of this evolutionary relationship between bats and viruses by evaluating patterns of virus-driven adaptation in bat genomes. Myotis bats provide details of these adaptations due to their recent radiation allowing us to look for differences among and within species. We aim to detect ancient viral epidemics in Myotis bats through genomic adaptation analyses. We created reference genomes for 9 Myotis species in the western United States. This allows us to conduct an unbiased genome-wide approach to identify gene candidates that have undergone strong
selection. We have resequenced Myotis bat genomes assembled with PacBio HiFi using short reads and genome-wide variant calls to estimate protein adaptation using the framework of the McDonald Kreitman test. We then focused specifically on adaptation driven by viruses at bat genes that are known to interact with viruses. Viral interacting proteins (VIPs) are known to be substantial drivers of mammalian adaptation and can allow us to identify signals of adaptation. Positive selection at VIPs may indicate the impact of viruses on ancient bat populations. Our results highlight patterns of virus-driven adaptation in previously largely unexplored bat genomes.

229 - An introgressed KIR haplotype enabled positive selection on HLA-A, enhancing Natural Killer cell function in Oceania

Neus Font-Porterias

Neus Font-Porterias, Liyen Loh, Philippa M Saunders, Camilla Faoro, Neda Nemat-Gorgani, Genelle F Harrison, Suraju Sadeeq, Luca Hensen, Shu Cheng Wong, Jacqueline Widjaja, Shiying Zhu, Gonzalo Montero-Martin, Katherine M Kichula, E Birdie Clemens, Marcelo Fernandez-Vina, Lisbeth A Guethlein, Julian P Vivian, Jane Davies, Adrian Miller, Alexander J Mentzer, Stephen J Oppenheimer, William Pomat, Peter Parham, Jamie Rossjohn, Steven YC Tong, Katherine Kedzierska, Andrew G Brooks, Paul J Norman

Department of Biomedical Informatics, University of Colorado School of Medicine, Department of Immunology and Microbiology, University of Colorado School of Medicine, Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Infection and Immunity Program and Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Department of Structural Biology and Department of Microbiology and Immunology, Stanford University, Department of Biomedical Informatics, University of Colorado School of Medicine, Department of Biomedical Informatics, University of Colorado School of Medicine, Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, , Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Infection and Immunity Program and Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Stanford Blood Centre, Department of Pathology, Stanford University, Department of Biomedical Informatics, University of Colorado School of Medicine, Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Stanford Blood Centre, Department of Pathology, Stanford University, Department of Structural Biology and Department of Microbiology and Immunology, Stanford University, Infection and Immunity Program and Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Menzies School of Health Research, Royal Darwin Hospital, Department of Infectious Diseases, Royal Darwin Hospital, Indigenous Research Unit, Griffith University, Wellcome Centre for Human Genetics, University of Oxford, Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Institute of Social and Cultural Anthropology, School of Anthropology and Museum Ethnography, University of Oxford, Papua New Guinea Institute of Medical Research, Department of Structural Biology and Department of Microbiology and Immunology, Stanford University, Infection and Immunity Program and Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Institute of Infection and Immunity, Cardiff University School of Medicine, Victorian Infectious Diseases Service, The Royal Melbourne Hospital, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Department of Infectious Diseases, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Department of Immunology and Microbiology, University of Colorado School of Medicine, Department of Biomedical Informatics, University of Colorado School of Medicine, Department of Structural Biology and Department of Microbiology and Immunology, Stanford University

Pathogen exposure and demography diversify immunity-related genes across human populations. In particular, admixture and local adaptation enhance diversity of the killer cell immunoglobulin-like receptors (KIR) and their ligands, human leukocyte antigens (HLA) class I. HLA class I presents protein fragments derived from infecting pathogens at the cell surface, which are recognized by KIR to modulate immune responses. Enabling the unique analysis of the combinatorial diversity of KIR and HLA class I, Oceanians show complex admixture patterns and evidence of archaic introgression. We analyzed genome-wide SNPs together with HLA and KIR sequencing data from 1200 Oceanians including 80 First Nations Australians. The most prevalent KIR ligand in Oceania is HLA-A*24:02, previously associated with poor prognosis following influenza virus infection. The high frequency of HLA-A*24:02 is coupled with evidence for positive selection and low sequence diversity. We also identify KIR3DL1*114, an allele of the inhibitory KIR3DL1, uniquely characterized by positively-selected Phenylalanine at HLA-contacting residue 166. KIR3DL1*114 is restricted to Oceanians, having highest frequency in

Papuans, and D-statistics and haplotype divergence indicate likely Denisovan introgression. Functional analyses of KIR3DL1*114+ donor immune cells, molecular and crystallography data all show that KIR3DL1*114 binds more strongly than other allotypes to HLA-A*24:02. KIR3DL1*114 also has differential specificity for the peptides presented by HLA class I, including those derived from influenza A virus. Together, our results suggest co-evolution of KIR with HLA maximizes innate immune cell function in Oceanians.

264 - PhyloAges of personal progenitors of blood cells for health, disease, and longevity

Jack M Craig

Jack M Craig, Sayaka Miura, Sudhir Kumar

Temple University IGEM, Temple University IGEM, Temple University IGEM

An individual's chronological age does not always correspond to the health of tissues in their body, especially in cases of disease. Therefore, estimating an individual's chronological age is a useful tool to diagnose disease and its progression. We present novel metrics based on the phenomenon of clonal hematopoiesis, which is associated with blood-related diseases, to quantify the loss of phylogenetic diversity in hematopoietic stem cells (HSCs), the precursors of differentiated blood cells. These metrics capture the increase in blood cancer incidence with age, establishing a model which estimates the phylogeny-based physiological age of HSCs in an individual, which we call â€œPhyloAge.â€• We will present novel approaches for estimating PhyloAge from high resolution data like single cell sequences of HSCs, or from bulk sequencing of blood that produces single nucleotide polymorphisms in HSCs. For both types of datasets, we are able to accurately estimate the PhyloAges of healthy individuals within 10 years. We will also present results for individuals suffering from blood cancers, which we hypothesize to have HSC PhyloAges exceeding their chronological ages. We will also test the hypothesis that the presence of blood cancer drivers is associated with longevity, as there is a reported overabundance of these drivers in individuals over the age of 90. Thus, PhyloAge represents a complement to conventional biomarker-based methods of estimating the physiological age of blood and assessing disease risk. Its effectiveness for bulk sequencing, in addition to single cell sequencing datasets, would enable its adoption in research and clinical settings.

331 - Molecular evolutionary mechanisms of acquisition of anticancer drug-resistance in lung cancer.

Yosuke Seto

Yosuke Seto, Naoya Fujita, Ryohei Katayama

Japanese Foundation for Cancer Research, Japanese Foundation for Cancer Research, Japanese Foundation for Cancer Research

Lung cancer is the leading cause of cancer death in the world. Activating mutation of epidermal growth factor receptor (EGFR) which induces constitutive activation of EGFR signaling pathway is a major oncogenic driver mutation of lung cancer. Inhibitors against EGFR tyrosine kinase (EGFR-TKIs) dramatically improved the prognosis of EGFR mutation-positive lung cancer. However, acquired resistance against EGFR-TKI after the initial clinical response is widely observed. To date, multiple mechanisms of EGFR-TKI resistance including bypass pathways (e.g., MET and HER2 signaling pathway) and secondary resistance mutation (e.g., T790M) have been reported. Especially because the de novo T790M resistance mutation emerges in initially T790M-negative drug-tolerant cancer cells during EGFR-TKI treatment, it is important to elucidate the process of the emergence of resistance mutations in the drug-torelant cancer cells during the EGFR-TKI therapy. To elucidate how the cancer cells survive under EGFR-TKI treatment, we conducted single-cell RNA-seq and single-cell ATAC-seq analyses using EGFR mutation-positive lung cancer patient-derived cells (PDCs). We observed the epithelial-to-mesenchymal transition (EMT)-like morphological change under EGFR-TKI treatment. Consistent with the observation, our single-cell analyses showed EGFR-TKI treatment-induced dynamic cell differentiation, and differential gene expression and chromatin accessibility analyses revealed induction of anti-apoptosis response and NF-kappa B mediated inflammatory response. Combined with

the study of EGFR-TKI selection, we found that EGFR-TKI-induced cytidine deaminase plays a critical role in the emergence of resistance mutations.

406 - ASSESSING EVOLUTION OF PRIMATE ATHEROSCLEROSIS USING IPSCS

Eva Briem

Eva Briem, Wolfgang Enard

Anthropology & Human Genomics, Faculty of Biology, Ludwig-Maximilians-UniversitÃ¤t, Anthropology & Human Genomics, Faculty of Biology, Ludwig-Maximilians-UniversitÃ¤t

Atherosclerosis is a global disease characterized by a hardening of the arteries. This is caused by a plaque accumulation, primarily composed of foam cells derived from macrophages that take up oxidized low-density lipoprotein (oxLDL). To understand the disease in detail, mainly mouse models and in vitro human models are used. To bridge the evolutionary gap between mice and humans, non-human primates (NHPs) are also used as a model system for biomedical research. However, we know very little about the molecular similarities and differences among NHPs and humans in atherogenesis. In particular, it has been suggested that NHPs are less likely to form the typical plaques in response to a high-fat diet, making it especially important to elucidate common and distinct mechanisms in atherosclerosis among NHPs.
Here I present an in vitro model of foam cell formation from induced pluripotent stem cell (iPSC)-derived macrophages. Our methodology allows the differentiation of NHP iPSCs into macrophages and their uptake of oxLDL. In turn, this allows to compare the transcriptional profiles between human and NHP macrophages and foam cells to identify conserved and diverged regulatory networks of primate atherogenesis.

432 - An optimized COVID-19 booster strategy informed by molecular evolution and seasonal adaptation

Hayley B. Hassler

Hayley B. Hassler, Alex Dornburg, Jeffrey P. Townsend

Yale University, Georgia Institute of Technology, University of North Carolina, Charolotte, Yale University

Understanding the interplay between vaccine- and infection-mediated immunity and endemic viral transmission is paramount for public health decision-making. However, without multiple years of vaccination and infection data paired with appropriate controls, guidance regarding interventions such as the optimal timing of booster vaccination is dictated largely by expert opinion. Leveraging information from the temporal and evolutionary divergence of coronaviruses, we integrate molecular evolutionary tools with data-driven phylogenetic models to analyze the probability of infection given waning antibody levels under endemic conditions. Using the consequent distribution of likely times to breakthrough infection, we assess booster effectiveness at a range of intervals and empirical spatiotemporal infection trends. We find that, in response to viral evolution and seasonal surges, the benefit of booster administration varies by both date and geography. Booster effectiveness can be maximized by preceding expected peaks in virus transmission (e.g. New York City: September 15; Stockholm: November 6; Israel: February 21). Vaccinating on or around these dates provides tailored protection when it is most needed. We further assess how these optimal dates change markedly in the event that someone experiences a breakthrough infection during the year. These results underscore the value of evolutionary medicine and phylogenomic comparative methods to improve the efficacy of booster programs, minimize the personal and societal impact of COVID-19, and adapt public health responses to the dynamic nature of viral transmission worldwide.

451 - Directed evolution of a broadly neutralizing antibody precursor

Berk A Alpay

Berk A Alpay, Angela M Phillips, Thomas Dupic, Michael M Desai

Harvard University, University of California, San Francisco, Harvard University, Harvard University

High-throughput methods have enabled exploration of protein genotype-phenotype maps by comparing effects of many mutations on a background, as well as the paths by which proteins traverse these effects as they evolve in function. The former is typically done by deep mutational scanning and the latter by assaying a large combinatorial library of <20 substitutions separating two proteins. Thus, deep mutational scans are broad across the protein but shallow in the number of mutations made, while assays of combinatorial libraries are narrow but deep.

Here we investigate genotype-phenotype maps in both breadth and depth, using directed evolution combined with high-throughput phenotyping to explore mutational effects along separate evolutionary paths. The protein we use is the germline form of the antibody CR9114, whose somatic form neutralizes a very broad range of influenza by binding a conserved epitope of the hemagglutinin antigen. Antigenic conditions and molecular pathways leading to the broad neutralization of flu are of interest to vaccine development and more generally to understanding dynamics of protein evolution during affinity maturation.

We perform four rounds of directed evolution by yeast-displaying error-prone PCR mutants of germline CR9114, incubating with various hemagglutinin strains and combinations thereof, and selecting better-binding mutants by fluorescence-activated cell sorting. We then measure binding affinity at each round and under each antigenic condition via Tite-Seq. We thereby quantify genotype-phenotype maps and compare molecular patterns of affinity maturation across tens of replicates and under different antigenic conditions.

465 - Searching for adaptations to malaria parasites in blood group genes across non-human primates

Alejandra Herbert Mainero

Alejandra Herbert Mainero, Ellen Leffler

Department of Human Genetics. The University of Utah, Department of Human Genetics. The University of Utah

Malaria parasites have influenced the evolution of multiple blood group genes in humans, presumably by altering the receptors these parasites use to gain access to red blood cells. Despite the presence of closely related parasites across the primate clade little is known of how they have impacted blood group evolution and if similar selective pressures have led to convergent adaptations. Here we applied selection tests to 44 genes that underlie human blood groups and expanded the number of primate species to identify putative adaptations to malaria-causing parasites across the clade. Our results show that 31.8% of the blood group genes show some evidence of positive selection (dN/dS>1) using PAML. We used BEB analysis with P>99% to identify positively selected sites within these genes. Unsurprisingly, in five genes the selected sites coincide with regions that may be accessible to these parasites. In the other nine genes, two possess no sites with P>99%, and seven either have no positively selected sites in extracellular regions or do not encode a surface protein, suggesting Plasmodium may not be the main selective pressure. We also noted a negative relationship between the number of antigens in the human blood group system and dN/dS>1 values, indicating possible ongoing diversifying selection. Therefore, we tested for balancing selection in primates with sufficient polymorphism data using betascan and found evidence in some cases of both selection across timescales. These results indicate new targets of selection across primate blood group evolution, some of which are consistent with a malaria selection.

470 - The genetic architecture and the evolutionary consequences of the human pelvic form

Liaoyi Xu

Liaoyi Xu, Eucharist Kun, Marianne Brasil, Tarjinder Singh, Vagheesh Narasimhan
The University of Texas at Austin, The University of Texas at Austin, Western Washington University, Columbia University , The University of Texas at Austin

Since diverging from chimpanzees, pelvic shape has evolved significantly in the human lineage and is thought to be affected by the obstetrical dilemmaâ€”an evolutionary mismatch between the increasing infant brain size and a narrowing birth canal. To elucidate the genetic underpinnings of pelvic morphology, we applied a deep learning model to 31,115 DXA scans from the UK Biobank, extracting a set of seven pelvic proportion (PP) phenotypes. These phenotypes are highly heritable (25-40%) and linked to 179 independent loci, indicating significant genetic determinants of pelvic shape. Unlike other skeletal proportions, the subpubic angle, associated with birth canal morphology, exhibits a genetic correlation between sexes significantly less than 1, aligning with its reproductive significance. Phenotypic and genetic association analyses revealed that larger birth canal phenotypes were associated with reduced walking pace, decreased risk of back pain, and increased risk of hip osteoarthritis, all phenotypes linked to locomotor efficiency. We also observed that narrower birth canals, are associated with a reduced risk of pelvic floor disorders but an increased risk of emergency cesareans delivery and obstructed labor due to insufficient dilation, reinforcing the obstetrical dilemma's impact. Moreover, the genetic correlation between birth canal width and infant brain size (proxied by birth weight) supports the co-evolution of these traits, but no link was found between pelvic phenotypes and gestational duration. This comprehensive study provides significant insights into a longstanding debate on human evolution, emphasizing the obstetrical dilemma's role in the co-evolution of the human brain and pelvis.

478 - Mutation and fitness landscape of SARS-CoV-2 reveal strong selection on synonymous mutations

Jean-Francois Gout

Jean-Francois Gout, Jori Symons, Claire Chung, Bert M. Verheijen, Sarah Shemtov, Dorien de Jong, Gimano Amatngalim, Monique Nijhuis, Marc Vermulst
Mississippi State University, Translational Virology, Department of Medical Microbiology, University Medical Center, University of Southern California, Leonard Davis School of Gerontology, University of Southern California, Leonard Davis School of Gerontology, University of Southern California, Leonard Davis School of Gerontology, University of Southern California, Leonard Davis School of Gerontology, Translational Virology, Department of Medical Microbiology, University Medical Center, Regenerative Medicine Center Utrecht, Division Child Lung Diseases, University Medical Center, Translational Virology, Department of Medical Microbiology, University Medical Center, University of Southern California, Leonard Davis School of Gerontology

New variants of SARS-CoV-2 are a constant threat to current vaccine and treatment strategies for COVID-19. Because these variants are created by mutations, we obtained precise mutation rate and spectrum of five SARS-CoV-2 variants and quantified the fitness effect of almost 4,000 mutations using ultra-deep circular RNA consensus sequencing (CirSeq). We identified several virus and host-specific features that influence SARS-CoV-2 evolution. Most notable, we discovered that the secondary structure of the viral RNA molecule protects the SARS-CoV-2 genome against RNA damage-derived mutations. We also show that the preservation of these secondary structures is a major driver of fitness. This study improves our understanding of SARS-CoV-2 genome evolution and identifies key weakness of the virus which can help optimize the generation of treatments and vaccines.

502 - Unraveling the Interplay of Genetic and Epigenetic Evolution in Tumor Progression Through

Phylogenetic Mapping

Deyana Tabatabaei

Deyana Tabatabaei, Sayaka Miura, Sudhir Kumar
Temple University , Temple University , Temple University

Tumor evolution mirrors the principles of species evolution, involving cell division, genome replication, and the emergence of genetic and epigenetic alterations over time. These intricate changes drive the diversification of cancer cells, yet their respective contributions in this process remain inadequately characterized. Therefore, we hypothesized that we could uncover their relative contributions by mapping genetic and epigenetic changes on a phylogeny of tumors. We analyzed eleven liver and metastatic colorectal cancer datasets, where single-nucleotide polymorphisms (SNVs) and DNA methylations were profiled from each tumor sample. We inferred a phylogeny of cancer cells within patients using SNVs and mapped driver mutations on each phylogeny through ancestral sequence reconstruction. Similarly, we mapped epigenetic shifts along the same phylogeny using a method to detect trait shifts in species genomics (l1ou). Our findings unveiled a compelling trend: genes involved in cancer evolution seldom displayed genetic mutations and DNA methylation changes at the gene transcription start sites. Rather, DNA methylation alterations alone in cancer-associated genes were more prevalent than driver mutations. These results implicate epigenetic evolution as a major driving force in the differentiation of tumor cells throughout cancer progression. By advancing our understanding of the interplay between genetic and epigenetic changes in tumor evolution, this study lays a foundation for future investigations aimed at elucidating the complexities of cancer development.

503 - Beyond Driver Mutations: Evolutionary Clustering of Cancer Patients for Precision Medicine

Xavier Soler Sanchis

Xavier Soler-Sanchis, Jonathan Frazer, Mafalda Dias
Centre for Genomic Regulation, Centre for Genomic Regulation, Centre for Genomic Regulation

Cancer involves a number of different changes that lead to abnormal and uncontrolled cell growth. Analogous to Darwinian evolution in the origin of species, cancer development is based on two constitutive processes: the continuous acquisition of heritable genetic variation in individual cells by more or less random mutation, and natural selection acting on the resulting phenotypic diversity. This results in a high mutational, biochemical, and histological tumor heterogeneity that makes driver mutation identification very challenging.

The arrival of second-generation sequencing technologies has allowed us to produce a huge amount of data, including the whole genomes from cancer patients. Together with this trend, we are seeing the emergence of personalized medicine, which aims to maximize the value of data from individual patients. Historically, patients have been stratified according to their mutated driver genes or driver mutations but in many cases, this clustering does say little about their treatment response.

Our approach exploits the power of evolutionary and big tumor data to cluster patients by cancer aetiology. We use whole genome information combined with the functional impact of the mutations learned from genetic variation seen on different evolutionary timescales to underpin the different commonly affected biological processes. Furthermore, this has revealed clusters with different drug responses. In summary, this method enables us to identify the characteristics that account for the differences between tumors, suggest which drug or treatment is more likely to be effective for a given patient, and to explain why.

520 - Shifting mutational constraints in the SARS-CoV-2 receptor-binding domain during viral evolution

Ashley Taylor

Ashley Taylor, Tyler Starr
University of Utah, University of Utah

As SARS-CoV-2 evolves, mutations arising in new variants could potentially leverage epistasis to their advantage, increasing receptor-binding affinity or allowing the virus to escape antibodies. While large epistatic shifts occurred to enable the emergence of Omicron due to the N501Y substitution, it is unknown whether this epistatic trend will continue. Here, we aim to understand how epistasis is affecting the evolution of more recent viral variants. Using a yeast display system for deep mutational scanning, we constructed libraries of recent Omicron BQ.1.1 and XBB.1.5 spike receptor-binding domains (RBDs) containing all possible single amino acid mutations and single-codon deletions within the RBD, and used FACS-seq titrations to measure the impact of each mutation on ACE2 receptor-binding affinity. Patterns of mutational effects were compared against those previously measured for earlier SARS-CoV-2 variants to identify changes in mutational effects over time. We observed that many of the amino acid mutations and deletions were well tolerated, some even enhancing binding to ACE2, with many sites consistent in mutational impact across variants. However, we did identify epistatic interactions between the R493Q reversion that separates BQ.1.1 and XBB.1.5 from early Omicron variants and mutations at positions that later emerged in the more recent XBB.1.5-derived EG.5 lineage. Our results demonstrate the influence of epistasis on SARS-CoV-2 variant evolution. Our ongoing studies investigate the more recent BA.2.86, EG.5 and HK.3 SARS-CoV-2 variants, data that we hope will inform viral monitoring and forecasting.

587 - Evolutionary informed polygenic risk scores

Federico Billeci

Federico Billeci, Mafalda Dias, Jonathan Frazer
Center of Genomic Regulation, Center of Genomic Regulation, Center of Genomic Regulation

Polygenic disorders are a broad class of diseases caused by multiple genetic mutations, often involving noncoding and regulatory variants. The most common approaches to predicting disease phenotype from genotype are scores computed with linear regression models but their potential as diagnostic tools is still heavily debated. Indeed, these models still exhibit severe overfitting to the training cohort, with performance deteriorating as one moves to more distantly related cohorts. They also provide limited insight into disease aetiology. These limitations are thought to be largely due to current approaches being restricted to considering only common variants and the difficulty in identifying the causal relations between variants and their pathways. To overcome these issues we propose a generative modelling approach which leverages evolutionary constraints. Indeed, growing evidence shows that genomic models trained on complementary evolutionary scales offer a very accurate guide for disease-causing variants in severe developmental disorders (Orenbuch et al, 2024; Fiziev et al 2023). Still, the potentiality of evolutionary information in complex polygenic disorders remains mostly unexplored. Our approach provides a unified framework for predicting the joint impact of both common and rare variants and revealing novel disease sub-classes. We apply this modelling approach to type 2 diabetes, a disease with large genetic heterogeneity, for which a clear diagnostic criteria is still lacking.n

638 - HLA class I escape drives the evolution of SARS-CoV-2 in human population

Ekaterina Riumina

Ekaterina D Riumina, Evgeniia I Alekseeva, Galya V Klink, Georgii A Bazykin3
Skolkovo Institute of Science and Technology, MSU, Skolkovo Institute of Science and Technology, A.A. Kharkevich Institute for

Information Transmission Problems of the Russian Academy of Sciences, A.A. Kharkevich Institute for Information Transmission Problems of the Russian Academy of Sciences

SARS-CoV-2 evolution is shaped by human adaptive immunity, with mutations that allow escape from the B-cell response conferring selective advantage and spreading in the population. Meanwhile, the role of the escape from T-cellular cytotoxic response remains controversial. Here, we study the origin and spread of SARS-CoV-2 variants that allow escape from presentation by the HLA class I alleles that are common in human populations. We find that 35% of mutations that are characteristic of the variants of concern, and 34% of mutations that have reached high (>5%) frequencies in viral populations, facilitate escape from T-cellular alleles. Mutations associated with escape from common HLA alleles reach higher frequencies than those that allow escape from less common HLA alleles. Moreover, viral escape mutations reach higher frequencies in those countries where the causal HLA alleles are more frequent, indicating that viral escape is driven by the local genetic composition of the human host population. Together, these data indicate that CTL escape is a major driver of SARS-CoV-2 evolution and an epidemiological concern, and reveal a novel facet of selection on this virus.

667 - Malaria in African monkeys: recovering the genomes of Plasmodium and Hepatocystis parasites from a co-infected monkey

Hend Abu-Elmakarem

Hend Abu-Elmakarem, Reuben W. Nowell , Weimin Liu, Lindsey J. Plenderleith, Beatrice H. Hahn , Paul M. Sharp
Institute of Ecology and Evolution (IEE), University of Edinburgh, Institute of Ecology and Evolution (IEE), University of Edinburgh, Biological and Environmental Sciences, University of Stirling, Department of Medicine, University of Pennsylvania, Institute of Ecology and Evolution (IEE), University of Edinburgh, Department of Medicine, University of Pennsylvania, Philadelphia, Institute of Ecology and Evolution (IEE), University of Edinburgh

Plasmodium species of Old World monkeys form a clade with human-infecting parasites for which an African origin has been suggested. However, most members of that lineage infect Asian monkeys, while only two genetically characterized species are known to infect African monkeys, P. gonderi and P. mandrilli (formerly P. sp. DAJâ€•2004), and these infections are rare. In contrast, African monkeys are often found to be infected with strains of Hepatocystis; relatives of Plasmodium that are transmitted through Culicoides biting midges rather than Anopheles mosquitoes. Genomic sequences of these African monkey parasites are scarce; in addition to P. gonderi, there is one genome available for a strain of Hepatocystis.
We used Selective Whole Genome Amplification to obtain genome sequences from a dried blood spot, from an African monkey (Cercopithecus pogonias), known to contain Plasmodium DNA. Analysis revealed a co-infection, with both P. mandrilli and a Hepatocystis parasite. We first present a bioinformatic pipeline to disentangle the genomes of these closely related species. Next, we conduct a comparative analysis of the evolution of Plasmodium species infecting African monkeys, Plasmodium species infecting Asian monkeys, and Hepatocystis species infecting African monkeys. Analyses of the rates of synonymous and non-synonymous substitutions of orthologous genes are used to look for signatures of adaptive evolution in the parasite genomes. We also characterise genomic differences between Hepatocystis and Plasmodium that parasitize African monkeys, particularly in regards to species-specific genes families, their host invasion mechanism and their possible role in the variation in the parasitesâ€™ prevalence levels.

768 - Investigating the dynamics of protein constraint across the tree of life with deep Bayesian hierarchical models.

Charlie W J Pugh

Charles WJ Pugh, Jonathan Frazer, Mafalda Dias

Centre for Genomic Regulation, Centre for Genomic Regulation, Centre for Genomic Regulation

Generative probabilistic models of protein sequences are transforming protein design, pathogen forecasting and clinical variant annotation. In the latter case, provided certain criteria are met, these models can now be considered strong evidence for missense variant pathogenicity and are an invaluable diagnostic tool. Crucially to their success, these expressive models are able to capture constraints on protein sequences by learning from variation across the tree of life. However, current variant effect prediction approaches model the distribution of broad protein families or, in the case of protein language models, the entire protein universe in order to make predictions in a specific species. Modelling all sequences as from a single distribution in this way overlooks the changes in constraint on a protein across the tree of life. To address this issue we have developed a novel modelling strategy to capture the distribution of sequences conditioned on the position in the phylogeny. We achieve this with a deep Bayesian hierarchical model based on the phylogenetic structure of large multiple sequence alignments. Our model learns from diverse sequences across the phylogeny while also capturing constraints that have evolved more recently. This has enabled us to obtain state-of-the-art scores on the ProteinGym benchmark. Our approach has allowed us to trace the evolutionary history of constraints in a selection of disease relevant human proteins. With this we have identified differing evolutionary signatures in gain-of-function and loss-of-function variants, giving promise that our approach can help classify disease aetiology.

769 - Transcriptional, epigenetic and evolutionary mechanisms shaped the acute inflammatory response of endothelial cells across mammals.

Mohamed Bayoumi Fahmy Hawash

Mohamed B.F. Hawash, Liangxi Wang, Kumaragurubaran Rathnakumar, Sana Akhtar Alvi, Tiegh Taylor, Huayun Hou, Azad Alizada, Liis Uuskula-Reimand, Suvojit Hazra, Lina Anotounians, Sara Knaack , Erika Da-Inn Lee, Sushmita Roy, Jennifer Mitchell, Jason Fish, Michael Wilson

Genetic and Genomi biology department, The hospital for sich children (Sickkids), Genetic and Genomi biology department, The hospital for sich children (Sickkids), Paris sud, Cincinnati Children's, Genetic and Genomi biology department, The hospital for sich children (Sickkids), University of Toronto, Genetic and Genomi biology department, The hospital for sich children (Sickkids), Genetic and Genomi biology department, The hospital for sich children (Sickkids), Genetic and Genomi biology department, The hospital for sich children (Sickkids), Genetic and Genomi biology department, The hospital for sich children (Sickkids), Wisconsin institute of discovery, University of Wisconsin-Madison, Genetic and Genomi biology department, The hospital for sich children (Sickkids), Wisconsin institute of discovery, University of Wisconsin-Madison, Wisconsin institute of discovery, University of Wisconsin-Madison, Wisconsin institute of discovery, University of Wisconsin-Madison, University of Toronto, Princess Margaret Cancer Centre, University Health Network, Genetic and Genomi biology department, The hospital for sich children (Sickkids)

Endothelial cells line all blood vessels and play a key role in maintaining physiological homeostasis. Augmented inflammation of ECs is a known risk factor for cardiovascular disease (CVD) and bleeding disorders. However, regulatory circuits that govern the inflammatory response of ECs remain poorly understood. Here, we used comparative and functional genomics to explore regulatory landscape that controls the acute inflammatory response. We characterized the inflammatory response of aortic EC from five mammals namely, human, mouse, rat, dog, rat and pig before and after simulation with TNFÎ± after 45 minutes using RNA-seq, ChIP-seq and ATAC-seq. We found number of inflammatory effectors to be conserved across all species such as CCL2 and CSF1. We also observed a markedly stronger response in mouse compared to other mammals. Transcription factor (TF) footprinting identified NF-ÎºB related motifs (e.g. REL) to be dominant TF observed after acute TNF treatment in all species except in mouse where we found AP1-related TF (e.g. c-JUN) footprints at similar level to the NF-ÎºB footprints. Consistently, REL and c-JUN ChIP-seq data for human and mouse EC treated with TNF for 45

minutes identified both TFs to be differentially accessible at similar levels after treatment in mouse, unlike human. Moreover, we found accelerated substitutions in the ape lineage enriched in the regulatory elements closest to genes of the highest response in human. In conclusion, we found various mechanisms that have shaped the inflammatory response of EC in mammals.

770 - Investigating the robustness of response to environmental perturbations across primate cardiomyocytes

Emma M. Pfortmiller

Emma M. Pfortmiller, Michelle C. Ward

The University of Texas Medical Branch at Galveston, The University of Texas Medical Branch at Galveston

Susceptibility to complex disease, such as cardiovascular disease (CVD), is influenced by a poorly-understood balance between genetics and environment. Given that CVD-associated genetic variants are largely located within non-coding regions of the genome, their effects are likely mediated through the regulation of gene expression (GE). GE regulation is context-dependent and impacted by environmental stimuli. Due to the sequence similarity between humans, it can be challenging to identify inter-individual differences in GE and therefore susceptibility to disease. To magnify GE changes, we take an evolutionary approach comparing humans to our closest evolutionary relative. We developed an in vitro model using induced pluripotent stem cell-derived cardiomyocytes from humans and chimpanzees to study robustness of response to stimuli. We exposed cells for 24 hours to eight stimuli known to induce the unfolded protein response (UPR), DNA damage response (DDR), innate immune response (IMR), as well as manmade chemicals where the response mechanism is unknown (MCR). Across species, UPR, DDR, and MCR treatment classes decrease viability at high concentrations, while IMR does not. We selected sub-lethal concentrations of each treatment to measure the transcriptional response to stimuli. We identified an up-regulation of known response genes including HSPA5 (Tunicamycin-UPR), CDKN1A (Nutlin-3-DDR), IL-6 (TNFÎ±-IMR) and PPARA (PFOA-MCR) in both human and chimpanzee cells. With this model, we are measuring global GE changes to gain insight into species-specific molecular mechanisms induced following perturbation, which may contribute to CVD susceptibility.

805 - Characterisation of haploinsufficient tumour suppressor genes across human cancers

Elle Loughran

Elle Loughran, Aoife McLysaght

Trinity College Dublin, Trinity College Dublin

Tumour suppressor gene (TSG) losses are canonically recessive, following Knudsonâ€™s seminal two-hit model of retinoblastoma occurrence. However, it has become clear that some important tumour suppressor genes, including TP53 and PTEN, are haploinsufficient: loss of one copy, or in the case of PTEN even mild decreases in activity, are enough to promote tumorigenesis. These haploinsufficient TSGs are potential targets for activation therapy, and their study may additionally offer insights into the functional impacts of copy number alterations in general.

In this project, we aim to systematically identify dosage-responsive tumour suppressor genes across cancer types from the Cancer Genome Atlas and the Cancer Cell Line Encyclopedia and characterise features that distinguish dosage-responsive and non-responsive TSGs. We use a combination of approaches to screen for altered tumorigenicity after partial TSG loss, including survival analysis, inference of cell proliferation rate from gene expression data, and information from DepMap RNAi knockdowns. We then search for associations between TSG dosage-responsiveness and features including germline dosage sensitivity and dosage compensation. Finally, we will compare the loss patterns of dosage-responsive and nonresponsive TSGs in the context of polyploid tumours, where additional

â€˜backupâ€™ copies of TSGs make full knockout less likely.

1013 - THE EFFECTS OF SEASONALITY ON THE EXPRESSION AND EVOLUTION OF ANTIMICROBIAL PEPTIDES IN THE SLENDER ANOLE (Anolis apletophallus)

Yanileth Fernanda Lopez-Tacoaman

Yanileth Fernanda Lopez-Tacoaman, Karla AlujeviÄ‡, Claire E Williams, Christian L Cox, Michael L Logan, W Owen McMillan
Smithsonian Tropical Research Institute, Smithsonian Tropical Research Institute, University of Nevada, Smithsonian Tropical Research Institute, University of Nevada, Florida International University, Smithsonian Tropical Research Institute, Smithsonian Tropical Research Institute, University of Nevada, Smithsonian Tropical Research Institute

Understanding the impact of seasonality on the expression of antimicrobial peptides (AMPs) is key to understanding the evolutionary dynamics and immunology of many vertebrates. I leveraged the recently published genome of the slender anole and other closely related lizard species to explore the evolutionary history and dynamics of AMPs, and to assess their expression across different tissues under varying environmental conditions. I used a combination of morphological, physiological, and molecular data to answer two primary research questions: 1) What are the evolutionary dynamics of antimicrobial peptides in the slender anole? 2) What are the expression patterns of key antimicrobial genes across different tissues and environmental conditions? To do this, I collected 120 adult lizards during wet and dry seasons at a lowland tropical forest site in central Panama. Each individual was cataloged, measured for morphological and physiological traits, and gene expression was quantified for known AMP genes using quantitative PCR. This study sheds light on the adaptability of the slender anole's immune system to dynamic tropical environments and potential correlations between AMP-related gene expression and seasonality. My research contributes to understanding of host-pathogen interactions in tropical ecosystems, with implications for biodiversity conservation under climate change. Finally, my findings lay the foundation for broader investigations into AMPs across diverse reptile species to gain insight into evolutionary dynamics and conservation.

1017 - Comparative analysis of allele age estimators for rare variants in personalized genomics

Alyssa Pivirotto

Alyssa M Pivirotto, Noah Peles, Jody Hey
Temple University, Temple University, Temple University

As the costs of sequencing continues to decrease, personalized genomics is becoming increasingly accessible. This increase in sequencing has led to the discovery of more rare variants, enabling the identification of their functional impacts and correlation with disease phenotypes. One way to characterize a variant is to estimate the time it entered the population. However, allele age estimators such as Relate, Genealogical Estimator of Variant Age, and coalescence time, were developed for the use with whole genome sequence dataâ€”the methods have not been tested on exome data from personalized medicine sequencing efforts. We examined the performance of each of these estimators on simulated exome data under a neutral constant population size model and found that each method provides usable estimates of allele age from whole-exome datasets. To test the robustness of these methods, analyses were undertaken to simulate data under a population expansion model and background selection. Relate performs the best amongst all three estimators for both the neutral constant (r=0.64) and expansion population model (r=0.68) with a 17% and 15% drop in accuracy between whole genome and whole exome estimations. Relate is also best able to parallelize to yield quick results with little resources. However, Relate is only able to scale to thousands of samples; Relate is unable to match the hundreds of thousands of samples being currently released. While more work is needed to expand the capabilities of current methods of estimating allele age, these methods estimate the age of mutations with a modest decrease in performance.

1047 - Dissecting the Spatial Dynamics of Pseudomonas aeruginosa Persistence in Cystic Fibrosis under CFTR Modulator Therapy

Tongqiu Jia

Tongqiu Jia, Sam Durfey, Pradeep Singh, Alison Feder

University of Washington, University of Washington, University of Washington, University of Washington

During chronic infections in people with cystic fibrosis (CF), Pseudomonas aeruginosa (Pa) diversifies into genotypically and phenotypically distinct subpopulations across lung regions. While CFTR modulator therapies reshape lung environments and reduce bacterial loads, they often fail to eliminate Pa. Through whole genome, spatially-resolved, longitudinal sequencing of nine people with CF harboring Pa lung infections, we sought to resolve the basis of Pa persistence through CFTR modulators spatially and genotypically. We found that the number of surviving bacterial lineages through therapy varied among participants (2-36 lineages), suggesting widely varying bottleneck sizes. To disentangle environmental factors associated with persistence, we applied Bayesian phylogeographic approaches to infer time-calibrated phylogenetic trees jointly with bacterial intra-lung migration histories. We fit logistic regression models to explain how characteristics of the pre-treatment lung environment (bacterial density, lung damage, inflammation, and spatial position) associate with lineage persistence and found that regions with higher damage pre-treatment were more likely to possess persistent lineages. Notably, persistent lineages were not spatially confined and frequently recolonized other putatively cleared regions. For regions with multiple persistent lineages, surviving bacteria were genotypically diverse, suggesting that environment might drive lineage survival above genotype. Our findings highlight the role of pre-treatment lung environment for predicting Pa persistence, and suggest migration between lung regions shapes the identity and structure of Pa populations post-treatment. Overcoming these spatial dynamics are essential for developing targeted therapeutics to better manage infection and underscore the importance of spatial factors in pathogen responses to therapy.

1050 - Discovery of a virulent HIV variant and implications for the evolution of virulence

Luca Ferretti

Chris Wymant, Luca Ferretti

University of Oxford, University of Oxford

The evolution of virulence is a classical example of life-history tradeoffs with major implications for public health. For HIV, the tradeoff has been quantitatively described from epidemiological evidence. We show that phylogenetic inference coupled with clinical measurements of set-point viral load reveals a single-peaked viral fitness landscape, providing a direct demonstration of the trade-off. However, evolution of the virus is not restricted to this landscape. We discovered a highly virulent variant of subtype-B HIV-1 spreading in the Netherlands in 1998-2014. Infected individuals exhibit a 3.5-5.5 fold increase in viral load, as well as CD4 cell decline twice as fast as individuals with other subtype-B strains. Epidemiological features suggest that the increased virulence is attributable to the viral strain. Genetic sequence analysis suggests that this variant arose in the 1990s from de novo mutation, not recombination, with increased transmissibility and an unfamiliar molecular mechanism of virulence. This surprising finding illustrates the difficulties in predicting the evolution of virulence and the need for genetic monitoring of viral variants.

1062 - Evolution of Antibiotic Resistance in Pristine Oasis Bacteria: Insights into Mechanisms Resembling Modern Human Resistance Strains

Maribel Hernández-Rosales

Maribel HernÃ¡ndez-Rosales, Diana BarcelÃ³-Antemate, Gerardo Amorez-Ruiz, Norberto GarcÃ-a-Miranda, IvÃ¡n Anguiano-Aguirre, Luis Francisco SalomÃ©-Abarca, Alejandra Rodriguez-Verdugo, Adam Olshen , JesÃºs Silva-SÃ¡nchez, Gabriela Olmedo-Alvarez
Center for Research and Advanced Studies, Irapuato Unit, Center for Research and Advanced Studies, Irapuato Unit, Center for Research and Advanced Studies, Irapuato Unit, Center for Research and Advanced Studies, Irapuato Unit, Center for Research and Advanced Studies, Irapuato Unit, Center for Research and Advanced Studies, Irapuato Unit, Center for Research and Advanced Studies, Irapuato Unit, University of California, University of California, Instituto Nacional de Salud PÃºblica, Center for Research and Advanced Studies, Irapuato Unit

Antibiotic resistance is commonly linked to human activities such as antibiotic overuse and misuse, yet it is crucial to recognize its presence in microbial communities predating human intervention. Bacteria have evolved within environments where antibiotic production and resistance mechanisms coexist intricately. This study aimed to identify antibiotic resistance genes (ARGs) in sediment communities from an oasis devoid of human influence, with a focus on Bacillota phylum strains. Utilizing the Comprehensive Antibiotic Resistance Database (CARD) platform, comparative genomic analysis of 78 bacterial strains unveiled a diverse array of ARGs previously characterized in human populations. Employing Hidden Markov Models facilitated the identification of specific resistance genes within bacterial genomes. Additionally, growth screening on twelve antibiotics enabled comparison of resistance patterns with CARD's identified mechanisms.
Results revealed a notable correlation between taxonomic classification and antibiotic resistance profiles, suggesting an evolutionary influence on resistance mechanisms. Phenotypic resistance screening showcased extensive resistance to specific antibiotic classes, aligning with identified ARGs. Furthermore, an analysis of Î²-lactamase genes provided insights into the genetic underpinnings of resistance, revealing various classes while missing prevalent clinical mutations.
This research offers insights into the ancient origins of antibiotic resistance and underscores the importance of studying environmental reservoirs of resistance genes to deepen our understanding of the evolutionary dynamics of antibiotic resistance in microbial communities and provide clues on resistance strategies applicable to human health.

1116 - Unraveling the Impact of Genomic Mutations on SARS-CoV-2 Severity in a Mexican Population

Brian Ernesto Muñoz Gómez

Brian Ernesto MuÃ±oz GÃ³mez, Miriam Sarahi Lozano Gamboa, Anastacio Palacios Marmolejo, Corina-Diana Ceapa
State Public Health Laboratory of the State of Aguascalientes, State Public Health Laboratory of the State of Aguascalientes, State Public Health Laboratory of the State of Aguascalientes, School of Medicine, Cuauhtemoc University Campus Aguascalientes, National Autonomous University of Mexico

The SARS-CoV-2 virus has created a global health crisis, and genomic investigations have become vital in understanding the mutations and their impact on disease severity. A study conducted by experts aimed to explore the relationship between genomic variants and the severity of SARS-CoV-2-induced pneumonia. The study analyzed 150 deceased patients with confirmed COVID-19 in Mexico, collected between 2021 and 2023, spanning diverse age groups and genders. Of the cohort, 77% exhibited at least one comorbidity, with hypertension being the most prevalent.
The study discovered that the Delta variant was the predominant lineage associated with fatalities during 2021-2022, underscoring its virulence. However, currently, the XBB lineage prevails in Mexico, requiring further investigation into its unique characteristics. The study also identified two novel mutations, L95M and L46M, in ORF 8 and ORF 9, respectively. L95M may confer resistance to immune system evasion, potentially influencing disease severity. L46M, located within the TOM70 link site, warrants scrutiny due to its potential functional implications. Remarkably, these mutations were exclusive to the studied population and did not appear in sequences across Mexico or

worldwide.

The findings provide valuable insights into the evolutionary dynamics of SARS-CoV-2 within specific populations, ultimately paving the way for tailored treatments and personalized medicine. The research contributes to the global fight against COVID-19, offering a beacon of hope in these challenging times.

1126 - Selection drives the evolution of genes involved in the response of the human vaginal epithelium to Lactobacillus crispatus

Omar E Cornejo

Omar Eduardo Cornejo
University of California Santa Cruz

Differences in the vaginal microbiome are associated with differences in risk to sexually transmitted diseases, pre-term birth and bacterial vaginosis. Five main types of microbial communities are widely recognized in the literature, four of which are enriched with different species of Lactobacillus, and one of them is comprised of a diverse assemblage of species found at intermediate abundances. Microbial communities enriched by Lactobacillus crispatus are commonly associated with positive health outcomes, while diverse communities are commonly associated with increased risk to disease. We have performed experiments co-culturing immortalized vaginal epithelial cells with Lactobacillus crispatus bacteria. Our experiments have shown that the presence of L. crispatus drives the upregulation of genes involved in the keratinization and cornification of the stratum corneum as well as genes involved in the innate immune response. Because of the association between Lactobacillus enriched communities and good health outcomes, we further explored if genes involved in the response to Lactobacillus crispatus present signatures of positive selection. We find that a small fraction of genes differentially regulated by Lactobacillus are under selection in human populations, narrowing the set of genes that we anticipate might have been driven by selection in a microbe-mediated manner. Our work shows that selection shapes the evolution of genes involved in microbe mediated risk to disease.

1127 - Unlocking Ancient Mechanisms: Insights into Mollusk Regeneration and Evolution

Luis F Hernández Ramírez

Luis F HernÃ¡ndez-RamÃ-rez, Wilbert GutiÃ©rrez-Sarmiento, Alfredo Varela-EchavarrÃ-a, Maribel HernÃ¡ndez-Rosales
Cinvestav, Irapuato Unit, Institute of Neurobiology, UNAM, Institute of Neurobiology, UNAM, Cinvestav, Irapuato Unit

Regeneration, the ability to replace lost or damaged tissues, has implications ranging from developmental biology to regenerative medicine. This work delves into the molecular intricacies of tail regeneration through RNA-seq analysis in the mollusk Deroceras leave, a novel gastropod study model.

A de novo transcriptome was assembled and differential expression analysis was conducted, revealing transcripts that were significantly altered post-amputation. The differentially expressed transcripts were used to predict proteins and construct a protein-protein interaction network. Centrality measures were then applied to the network to identify key proteins involved in regeneration, including proliferation, migration, apoptosis, angiogenesis, remodeling, and differentiation.

Orthology search across multiple species revealed high conservation of these proteins. Notably, we found proteins implicated in neurodegenerative diseases. Specifically, the expression of proteins involved in regulation of multivesicular bodies (MVBs), which contribute to the generation of extracellular vesicles (EVs). These EVs express surface molecules facilitating targeting and internalization

into cells via receptor-ligand interactions, endocytosis, or phagocytosis. Alternatively, they can fuse with cell membranes, delivering their contents into the cytosol. Thus, these proteins likely regulate protein secretion into the extracellular matrix.

We also identified orthologs of VPS4, involved in MVB regulation and vesicle trafficking, present in organisms ranging from procaryotes of the Asgard group to unicellular eukayotes such as Dictyostelium discoideum.

In conclusion, these results suggest that multicellular organisms may have repurposed ancient mechanisms to facilitate tissue regeneration, enriching our understanding of mollusk regeneration and its evolution.

98 - Ancient DNA screening at Denisova Cave to explore Bos/Bison admixture

Alexandre Gilardet

Alexandre Gilardet, Katerina Douka, Peter Heintzman, Love DalÃ©n
Centre for Palaeogenetics, Department of Evolutionary Anthropology, Faculty of Life Sciences, University of Vienna, Centre for Palaeogenetics, Centre for Palaeogenetics

Nuclear genomes can usually provide a more nuanced story than mitogenomes. Previously, mitochondrial sequences have shown that the yak (Bos grunniens) clusters with the American bison (Bison bison) while the European bison (Bison bonasus) clusters with cattle (Bos taurus) (Marsolier-Kergoat et al. 2015, Zeyland et al. 2012). However, nuclear data showed the American and European bison to be closer to each other than other relatives (Verkaar et al., 2004). Additional and more ancient nuclear data from both yak and bison can potentially resolve these discordant phylogenies. Denisova Cave lies in Siberia and both Bos and Bison sp. remains have been morphologically identified there. This site is thus a good candidate to resolve the common evolutionary history of Bos/Bison. The available bone fragments collection spans several glacial and interglacial periods up to the Middle Pleistocene and could therefore hold answers about the possible climate-mediated admixture history of Bos/Bison. We are now screening the collection for those samples with good endogenous DNA preservation. Already, we are able to retrieve mitogenomes at a high enough coverage for species assignment. The later generation of high coverage ancient nuclear genomes will help test the hypothesis of ancient hybridization and gene flow between the extinct steppe bison (Bison priscus) and wild yak (Bos mutus).

116 - Exploring evolutionary dynamics in Chorthippus grasshoppers: insights from an integrative approach using morphology and genetics

Dörte S Neumeister

DÃ¶rte S Neumeister, Diana C GonzÃ¡lez, Sarah Gaugel, Tunca Deniz YazÄ±cÄ±, Joachim Haug, Holger Schielzeth, Richard I Bailey, Ricardo J Pereira
LMU Munich, International Max Planck Research School for Biological Intelligence, LMU Munich, University of Hohenheim, Natural History Museum of Stuttgart, LMU Munich, LMU Munich, Friedrich Schiller University Jena, University of Lodz, Natural History Museum of Stuttgart

Investigating the evolution of behavioral isolation is fundamental to understanding the process of speciation, as behavioral barriers are thought to be first to arise during species formation. Here, we explore the interplay between song, wing morphology and genomic divergence during the radiation of sympatric species of Chorthippus grasshoppers, including a hypothesized hybrid form. First, using morphometric methods we found that wing shape correlates with song differences, despite high interspecific variation, providing a reliable proxy for behavioral isolation. Secondly, using genomic markers we found that four taxa form discrete groups without evidence of

ongoing hybridization. Interestingly, the taxon with intermediate song and wing morphology is genetically similar to C. brunneus. Subspecies of C. mollis that have similar song and wind morphology are genetically distinct. Such discordance between phenotypic traits and genetic differentiation suggests that song is not evolving neutrally, and that it can be facilitated either by introgression of a few genes or convergent evolution of similar songs. These results shed light on the complex dynamics of behavioral isolation, suggesting that it can evolve quickly and in a different direction from neutral evolution, facilitating species formation.

123 - Frequent hybridisation between parapatric lekking bird-of-paradise species

Filip Thörn

Filip ThÃ¶rn, Ingo A. MÃ¼ller, AndrÃ© E. R. Soares, Elizah Nagombi, Knud A. JÃ¸nsson, Mozes P. K. Blom, Martin Irestedt
Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Department of Zoology, Stockholm University, Leibniz Institut fÃ¼r Evolutions- und BiodiversitÃ¤tsforschung, Museum fÃ¼r Naturkunde, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Department of Zoology, Stockholm University, Stockholm, Leibniz Institut fÃ¼r Evolutions- und BiodiversitÃ¤tsforschung, Museum fÃ¼r Naturkunde, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, New Guinea Binatang Research Centre, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Leibniz Institut fÃ¼r Evolutions- und BiodiversitÃ¤tsforschung, Museum fÃ¼r Naturkunde, Department of Bioinformatics and Genetics, Swedish Museum of Natural History

Hybridisation has sometimes been viewed as an evolutionary dead end due to the purported reduced fitness of hybrids. Presently, interspecific hybridisation is viewed as an evolutionary conduit in which adaptive or neutral variants can be exchanged. Hybridisation has been reported in a wide range of taxa, including organisms with some of the most extreme forms associated with sexual selection: Lek-mating. However, whether occasional hybridisation can overcome the reproductive barriers in such systems and lead to the establishment of hybrid zones has not been well studied. In the present study, we aim to address these questions by focusing on one of the most well-known avian radiations where lekking has resulted in an extraordinary assemblage of phenotypic diversity; the birds-of-paradise (Aves: Passeriformes: Paradisaeidae). We find multiple signals of historical gene flow between and within two genera of birds-of-paradise using re-sequencing genomic data from museum specimens. In addition, we present the first empirical genomic evidence of a hybrid zone between two lekking bird-of-paradise species that differ substantially in sexually selected traits and behaviour. We explore this unique system by quantifying the genome-wide distribution of introgression and discuss the relevant patterns in relation to the expected recombination patterns and sexual selection. Our findings of recurrent historical and contemporary hybridisation are evidence of weak reproductive barriers that allow gene flow between these lekking species.

137 - Differential Introgression in a bunting hybrid zone illuminates the genes underlying early genetic barriers between incipient species

Niloofar Niloo Alaei Kakhki

Niloofar Alaei Kakhki, Sajad Noori, Ricardo J. Pereira
Stuttgart State Museum of Natural History, Stuttgart State Museum of Natural History, Stuttgart State Museum of Natural History

Studying hybrid zones offers unique insights into the evolution of reproductive barriers and their persistence amidst gene flow. Hybridization has diverse outcomes, such as the breakdown of reproductive isolation, emergence of adaptive phenotypic variability and the adaptive introgression. We investigated a narrow moving hybrid zone between the black-headed bunting (Emberiza melanocephala) and red-headed bunting (E. bruniceps) located in the northeast of Iran, where species boundaries are putatively maintained by behavioral isolation based on male coloration and new phenotypes arise through hybridization. Our ecological niche models indicate that this hybrid

zone is older than the Last Glacial Maximum and reveals strong divergence in environmental preferences between parental taxa. Whole genome sequencing data reveals substantial divergence in mitochondrial lineages, contrasting with reduced average divergence in the nuclear genome, consistent with extensive introgression upon secondary contact. A clear non-random association between mtDNA lineages and male plumage coloration suggests a strong female preference for conspecific males maintaining this hybrid zone. Chromosome painting demonstrates genome-wide admixture, except in genomic areas characterized by high differentiation peaks, consistent with localized barriers to gene flow or regions of low recombination. Spatial sampling will assess the hybrid zone's movement in the last 50 years, possibly due to adaptive introgression of coloration genes. Our results suggest that these sister species have exchanged genes over thousands of generations, making this hybrid zone an ideal natural laboratory to pinpoint the barriers against the gene flow and understand the role of behavioral isolation in the dynamics of species boundaries.

156 - Global, asynchronous sweeps at multiple insecticide resistance genes in Aedes aegypti mosquitoes

Thomas L Schmidt

Thomas L Schmidt
University of Melbourne

Aedes aegypti (yellow fever mosquito) is a highly invasive pest that confers the majority of the global dengue burden. This species successfully invaded much of the worldâ€™s tropical and subtropical regions between the 16th and 19th centuries. Insecticide-based management beginning in the mid-20th century led to the evolution of insecticide resistance in Ae. aegypti, and identical point mutations at the voltage-sensitive sodium channel gene (VSSC) on chromosome 3 have been observed across its global distribution. This study uses a genomic analysis of selective sweeps to infer how this global distribution of mutations was reached, and found a combination of adaptive introgression and repeated substitutions at the same nucleotide position. Surprisingly, a second locus on chromosome 2 was found to harbour variation segregating in line with VSSC genotype, and this locus showed similar patterns of positive selection to those observed at the VSSC gene. This locus contained 15 resistance-associated glutathione S-transferase (GST) epsilon class genes, and we detected three distinct homozygous haplotypes that had introgressed across the Indo-Pacific region, the Americas, and Australia. North American Aedes aegypti showed signs of significantly higher copy number of these genes. VSSC and GST sweeps had similar broad geographical patterns, but local patterns and linkage networks indicate these likely spread at different times. These findings highlight the strong patterns of selection acting on resistance genes globally and the likely key importance of GST genes in resistance evolution.

212 - Unveiling speciation mechanisms in brown algae: Insights from chromosomal and population-level approaches

Cécile Molinier

CÃ©cile MOLINIER, Lauric Reynes, Frank Chan, Marek Kucka, Sabrina Carli, Jerome Coudret, Fabian Haas, RÃ©my Luthringer, Susana Coelho, Myriam Valero, Agnieszka Lipinska
Max Planck Institute for Biology, The Sherwood Laboratory, School of Life Sciences, EDITERRANEAN INSTITUTE OF OCEANOGRAPHY, Station Biologique de Roscoff, Groningen Institute for Evolutionary Life Sciences (GELIFES),, Friedrich Miescher Laboratory of the Max Planck Society, Friedrich Miescher Laboratory of the Max Planck Society, Max Planck Institute for Biology, Station Biologique de Roscoff, Max Planck Institute for Biology, Max Planck Institute for Biology, Max Planck Institute for Biology, Station Biologique de Roscoff, Max Planck Institute for Biology

Numerous examples demonstrate that reproductive barriers between species are not always absolute, leading to varying degrees of

hybridization and gene flow. However, intrinsic post-zygotic isolation mechanisms frequently manifest through reduced hybrids fitness. This decline in hybrid fitness is often attributed to incompatibilities arising between diverged parental genomes, particularly involving sex chromosomes.

Here, we combine a detailed comparative analysis of the genome architecture of two brown algal species of the genus Ectocarpus with population genetic data from two hybrid zones in different geographic regions. Brown algae present an intriguing model for studying reproductive isolation due to their complex life histories, which involve alternation between haploid gametophytes and diploid sporophytes, with sex determined by haploid sex chromosomes. The roles of both haploid sex chromosomes and haploid-diploid life cycles in speciation have remained relatively unexplored.

We employed high-quality chromosome-level assemblies to examine specific genomic features such as sex chromosomes and chromosomal rearrangements, evaluating their potential roles as barriers to gene flow. Furthermore, by utilizing linked-read sequencing (haplotagging), we analyzed hundreds of individuals from natural populations to elucidate how genome-wide introgression varies across distinct geographical contexts and evolutionary histories, and how it is influenced by genome architecture.

Through the integration of chromosomal and population-level approaches, this study delves into key questions in speciation research, including the relative importance of gene flow, the impact of chromosomal rearrangements, and the roles of haploid sex chromosomes in reproductive isolation. Our findings offer valuable insights into the mechanisms driving speciation dynamics in brown algae.

260 - Different measures of reproductive isolation reflect different stages of species formation

Linda Hagberg

Linda Hagberg, DÃ¶rte Neumeister, Linde N Wieringa, JosÃ© L Bella, Ricardo J Pereira

Naturkundemuseum, Ludwig-Maximilians-UniversitÃ¤t, Ludwig-Maximilians-UniversitÃ¤t, Universidad AutÃ³noma, Naturkundemuseum

Speciation is often thought of as complete reproductive isolation (RI), characterised by the accumulation of reproductive barriers counteracting gene flow. Despite the widespread view of a forward-in-time speciation process, the accumulation of reproductive barriers is not linear or even irreversible, as numerous studies focusing on the admixing effects of hybridisation have found. Instead, the strength of reproductive barriers can fluctuate in cycles of e.g. geographic isolation (blocking gene flow) and secondary contact (allowing gene flow). Tools used to identify RI vary in their underlying assumptions and thus in their implications. Here, I compare the implications of three different methods commonly used for inferring RI with data from two pairs of hybridising populations of the meadow grasshopper Pseudochorthippus parallelus. I focus on RI inferred using (i) divergence time and effective gene flow, (ii) hybrid zone width, and (iii) heterogeneity of introgression across independent genomic regions. I find the strength of RI to be extremely variable among the methods and suggest that the differing results may simultaneously be correct but represent windows into different spatio-temporal and genomic scales of species formation.

282 - Differentiating mechanism from outcome for ancestry-assortative mating in admixed populations

Dashiell J Massey

Dashiell J Massey, Zachary A Szpiech, Amy Goldberg

Duke University, Pennsylvania State University, Duke University

Non-random mating is an important source of genetic structure in natural populations. Empirical studies across multiple species have reported positive correlations in trait values between mates, potentially confounding genome-wide association studies, selection scans, and demographic inference. Within several recently admixed human populations, empirical genetic studies have reported a correlation in

global ancestry proportion between spouses, referred to as ancestry-assortative mating. Here, we combine genetic data analysis and benchmarking of common parental-ancestry inference methods with forward genomic simulations to link observed correlations in ancestry between human spouses to the underlying mechanistic mate-choice process. Indeed, much of the theoretical work on assortative mating derives from single-trait models in non-human species, whereas ancestry proportion is a genome-level trait and human mate-choice is likely shaped by numerous sociocultural, as well as biological, factors. We consider the impacts of potential mate-choice models, parental-ancestry inference methods, and analysis strategies on the identification of evidence for ancestry-assortative mating in human populations. For instance, we find that multiple mate-choice models can produce identical correlations in ancestry proportion between spouses; however, we highlight alternatives to the common parental ancestry correlation plots in which these potential models may be well distinguished. With this work, we aim to sound a note of caution about interpreting correlations in empirical data as evidence for a particular model of human mating practices â€“ as well as to offer suggestions toward development of new best practices for analysis of human assortative mating.

284 - The co-evolution of phylogenomic signal and recombination rate during complex speciation in placental mammals

Nicole Foley

Nicole M Foley, William J Murphy
Texas A&M University, Texas A&M University

Recombination-aware phylogenomic studies have reshaped our understanding of the evolutionary histories of select clades across the Tree of Life. However, the widespread application of these transformative methods is limited by the availability of recombination maps that correspond to chromosome-level genome assemblies. Here, we characterized the progressive genome-wide re-patterning of phylogenomic signal that occurs with discordance due to introgression or incomplete lineage sorting across an old vertebrate lineage, placental mammals. We selected 10 mammal clades for which there was 1) a chromosome-level genome, 2) phylogenomic data, and 3) population-genomic data, representing mammalian superordinal clades. Population genomic data was used to generate recombination maps using a machine learning approach, ReLERNN, for 20 mammalian species. Discordant chromosomal patterns of phylogenomic signal generated from genome-wide samples of ML trees could be assigned to four groups displaying progressive discordance patterns. We repeatedly observed an enrichment of the species tree on the X with introgression. With more pervasive introgression, this pattern is highly restricted to a multi-megabase low recombining desert at the center of the X chromosome conserved across multiple mammalian families and orders. Using genome-wide PhyloP scores from a 240 mammalian species alignment we show this region is characterized by mild selective constraint and accelerated evolution. We hypothesize the enrichment for species-specific ampliconic genes and genetic components of the X chromosome inactivation center may act as recurrent components of a supergene in placental mammals.

447 - Tracing the impact of admixture in signatures of selection in the Mexican Biobank.

Ram González-Buenfil

Ram GonzÃ¡lez-Buenfil, Fernanda MirÃ³n-ToruÃ±o, Santiago G. Medina-MuÃ±oz, Austin W. Reynolds, AndrÃ©s Moreno-Estrada
UGA-LANGEBIO, CINVESTAV, Department of Anthropology, Baylor University, UGA-LANGEBIO, CINVESTAV, Department of Anthropology, Baylor University, UGA-LANGEBIO, CINVESTAV

Events of admixture (i.e., gene flow between populations) have occurred throughout the human lineage due to migration and other demographic movements that exposed individuals to new environments and different selective pressures. Present-day Mexico is home to one of the greatest admixture events in recent history primarily between Indigenous American and Europeans, which had evolved

independently from each other for thousands of years until ~500 years ago. This composition represents a unique setting to study the effects of admixture on adaptation. In Latin America particularly, the extent to which admixture has shaped adaptation processes remains poorly explored. To investigate this, we are analyzing genomic data from the Mexican Biobank (MXB), consisting of 6,057 individuals across all 32 states in Mexico and genotyped for 1.8 million genome-wide SNPs. To detect signatures of selection we are using two novel statistical models, Adaptmix and OHANA, which model allele frequencies in admixed populations to identify selection and distinguish whether it has occurred before or after admixture. We are also employing phenotypic data from MXB to correlate our candidate regions with potential traits under selection. Preliminary findings suggest the presence of both ancestral and post-admixture signatures of selection, with implications for the study of evolution in diverse populations from the Americas. This study leverages the largest dataset of its kind for the Mexican population and marks a significant advance in understanding the genetic basis of adaptation in admixed populations.

468 - Nonhuman Primate Models for Hominin Introgression and Genetic Admixture

Alaina L Brenner

Alaina L Brenner, Timothy D Weaver, Brenna M Henn

University of California, Davis, University of California, Davis, University of California, Davis

Hybridization and genetic admixture occur across many ecological taxa past and present, including isolated lineages of late Pleistocene hominins. However, little is known about the dynamics of these genetic exchanges in the first few generations of admixture. Therefore, this project investigates the mechanisms of hybridization and admixture in the genomes of contemporary nonhuman primates as a model to understand hominin introgression events in human evolution. Here, we use 8 generations of low coverage whole genome sequencing (WGS) to assess local and global ancestry of an admixed colony of 140 macaques once housed at the California National Primate Research Center (CNPRC). Indian and Chinese rhesus macaques are especially relevant due to their similar divergence time in generations to humans and Neandertals. These macaque subspecies have accumulated substantial genetic variation since the divergence and subsequent isolation of the two lineages ~160,000 years ago. During our first step in this ongoing project, we modified the GATK variant calling pipeline for macaques and identified 52,207,528 non reference alleles on average per individual. After assessing local and global ancestry in the WGS, we observed large discrepancies between the WGS admixture estimates and the previous estimates based on the CNPRC pedigree and limited microsatellite data. Specifically, ~40% of the sample have ancestry from both subspecies which is lower than expected under random mating. We can then use simulations to test whether this pattern is consistent with purifying selection against hybrids. This model will be used to illuminate human and Neandertal admixture within the first 10 generations.

568 - Detecting basal aurochs-derived tracts in the genome of African taurine cattle

Donghee Kim

Donghee Kim, Kwondo Kim, Choongwon Jeong

School of Biological Sciences, Seoul National University, Jackson Laboratory for Genomic Medicine, Department of Agricultural Biotechnology and Research Institute of Agriculture and Life Sciences, School of Biological Sciences, Seoul National University, Institute of Data Innovation in Science, Seoul National University

The African taurine cattle, the earliest domestic cattle in Africa, have captivated researchers for their diverging genetic profile from the Eurasian taurine as well as their phenotypes potentially adaptive to the environments of sub-Saharan Africa. Especially, itâ€™s been hypothesized that now extinct local aurochs populations may have contributed to the distinct genetic profile and adaptations of the African taurine. Consistent with this hypothesis, previous studies detected a significant amount of contribution (up to ~20% of ancestry)

from a basal taurine lineage, possibly representing the now extinct African aurochs. However, due to scarcity of ancient African aurochs genomes, the identity of this basal taurine ancestry in the African taurine still remains elusive. In this study, we identify genomic segments likely introgressed from a deeply divergent source population in the genome of African taurine using a method, Skovâ€™s Hidden Markov Model (HMM), not requiring an explicit model of the demographic history nor genome sequences representing the source. These identified segments cover ca. 15% of the callable genome of African taurine per individual genome, largely coinciding with our previous estimate of the amount of the basal taurine contribution. Also, we show that the identified genomic segments are from a source that are more deeply branching than the available Eurasian auroch genomes, while not closely matching to the available Moroccan aurochs genomes. Our study corroborates our previous report on the basal lineage introgression in the African taurine and suggests the yet-to-be-identified genetic diversity among the extinct African/Middle Eastern aurochs populations.

615 - Genome-wide demographic analyses of balaenid whales revealed complex history of gene flow associated with past climate oscillation

Bai-Wei Lo

Bai-Wei Lo
Max Planck Institute for Molecular Genetics

The balaenid whale, comprising three species of right whales and the bowhead whale, represents an ancient and highly endangered lineage of marine mammals. To unravel the evolutionary history of balaenid whales with respect to gene flow, a comprehensive analysis based on whole-genome data was conducted for all species within this group. Employing population genomic methodologies, we revealed the polytomic nature of extant right whales, identified passage of historical trans-equatorial migration, and provided estimates to the age of the group. Furthermore, we investigated the impact of glacial cycles on the connectivity of bowhead whale populations. By employing multiple complementary approaches to detect gene flow, we identified and characterized gene flow events from bowhead whales to North Atlantic right whales, offering detailed insights into the process. Lastly, we assessed the phenotypic consequences of interspecies gene flow. The outcomes of our study shed light on the intricate evolutionary history of modern balaenid whales, which have been profoundly shaped by ancient climate events.

671 - The great puzzle of the Great Lakes: uncovering the complex hybrid origins of the Lake Malawi cichlid radiation

Sophie Gresham

Sophie Gresham, Milan Malinsky, Hannes Svardal
University of Antwerp, University of Bern, University of Antwerp, Naturalis Biodiversity Center

With the increasing prevalence of admixture in the animal and plant kingdoms, describing patterns of biodiversity and understanding the process of speciation is becoming ever more complex. The large and spectacularly diverse adaptive radiations of cichlid fishes in the African Rift lakes, with both hybrid origins, recent cases of admixture and high success rate for hybridisation in aquaria, are a prime example of the complexities of species origins and provide a unique model system to study the evolutionary effects of gene flow.
The Lake Malawi radiation, with ~800 ecologically and phenotypically diverse species that have evolved within 800KY, is the largest of all African cichlid radiations. However, despite large species numbers, genetic diversity across Malawi species is remarkably low, potentially indicating the high importance of admixture during diversification. Here, we chart the history of admixture within the Malawi radiation, using a large-scale whole-genome sequencing dataset consisting of 276 Lake Malawi cichlid species and 72 cichlid species from surrounding African riverine and lake systems. Using genome-wide tests of excess allele sharing (D-statistics) and window-based

analyses, we discovered evidence of gene flow from three independent riverine lineages into the Malawi radiation. Surprisingly, present-day representatives of these lineages are genetically and geographically distant to the Malawi radiation. Assessing the fate of introgressed material, we discovered differential levels of introgression across Malawi lineages and signatures of selection. Overall, our results reveal an even more complex history of admixture and speciation in an already challenging and complex species system.

737 - Disentangling Admixture Histories in Recent and Complex Demographies

Christian D. Huber

Matthew P. Williams, Christian D. Huber

Pennsylvania State University, Pennsylvania State University

Recent shifts in ancient DNA (aDNA) research towards the recent human past have unveiled new challenges in the field of archaeogenetics. In particular, the performance of paleogenomic methods in answering questions of historical and archaeological importance amidst the increased demographic complexity and decreased genetic differentiation within the historical period remains an open question. Here, we present an in-depth simulation-based evaluation of qpAdm and the "admixture" f3-statistic, highlighting their limitations in differentiating between closely related source populations under typical aDNA conditions. We simulated data representing simple and complex demographic models, including conditions mimicking typical aDNA like SNP ascertainment, data missingness, and pseudo-haploidization, to assess these methods' performance. We show that under conditions resembling historical populations and aDNA conditions, qpAdm can identify a small candidate set of true sources and populations closely related to them. However, qpAdm is unable to further distinguish between them, limiting its utility for resolving fine-scaled archaeogenetic hypotheses. Furthermore, we introduce a novel heuristic guide that integrates qpAdm models with p-values and f3-statistics to distinguish between multiple plausible models, significantly improving archaeogenetic analysis. Finally, we demonstrate a significant improvement in qpAdm power and accuracy when using whole-genome branch-length f-statistics, emphasizing the potential to achieve more accurate historical reconstructions with more precise f-statistics. The findings offer significant insights into the strengths and limitations of admixture reconstruction methods, promising to refine the narrative of human demographic history and serving as a critical resource for researchers in archaeogenetics.

746 - Disentangling admixture and selection in human populations across the Pacific Islands

Sven M Kasser

Jan Sokol, Obed A Garcia, Sven M Kasser, Mark A Perjueli, Javier Blanco-Portillo, Feiyang Liu, Samuel Sacco, Kekoa Taparra, Santiago G Medina-MuÃ±oz, Charleston G Chiang, Patrick V Kirch, AndrÃ©s Moreno-Estrada, Chris Gignoux, Genevieve Wojcik, Keolu Fox, Marcus W Feldman, Alexander Ioannidis

Institute for Stem Cell Biology and Regenerative Medicine, Stanford University, Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Centre for Biological Diversity, University of St Andrews, Biology Department, NYU Abu Dhabi, Department of Biology, Stanford University, Institute for Computational and Mathematical Engineering, Stanford University, Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Department of Radiation Oncology, Stanford University, Unidad de Genomica Avanzada (UGA-LANGEBIO), Centro de Investigacion y Estudios Avanzados del IPN (Cinvestav), Department of Quantitative and Computational Biology, University of Southern California, University of Hawaiâ€˜i, Manoa, Unidad de Genomica Avanzada (UGA-LANGEBIO), Centro de Investigacion y Estudios Avanzados del IPN (Cinvestav), Division of Biomedical Informatics and Personalized Medicine, University of Colorado, Denver, Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Department of Anthropology and Global Health, University of California, San Diego, Department of Biology, Stanford University,

Department of Biomedical Data Science, Stanford Universiry, Department of Biomedical Engineering, University of California, Santa Cruz

Admixed populations present both a unique challenge and unique opportunity for understanding selective processes. Traditional selection scan methods often inadequately account for admixture patterns that may mask selection signatures. Conversely, incorporating admixture information can significantly enrich our understanding of selection, including its timing and the source of adaptive genetic variation.

In the present study, we analyze SNPs from understudied, highly admixed human populations in HawaiÊ»i, Central Eastern Polynesia, Micronesia, and Melanesia. We identify multiple new genetic risk loci of potential medical significance associated with these populations. Notably, we discover an ancestry-associated selection signature linking GDF5 polymorphisms with osteoarthritis in Polynesians, potentially introduced through historical contact with Native Americans. We also show that genotypes related to immunity and lipid metabolism vary in frequency among Pacific Islander populations, providing evidence of local adaptation potentially driven by differentiated patterns of subsistence and disease. Finally, incorporating admixture timing information allows us to date some selective events prior to European contact.

Our results reveal complex novel selection patterns in understudied populations spanning the remarkable maritime habitats of the open Pacific and reiterate the importance of properly accounting for admixture when studying the evolution of natural populations.

754 - Divergent ancestry in the TSHR gene is associated with changes in seasonal reproduction during dog domestication

Lauren Hennelly

Lauren M Hennelly, Katia Bougiouri, Laurent Frantz, Cindy Santander, Xin Sun, Nuno FG Martins, Frank Hailer, Renzo FF Balboa, Nucharin Songsasen, Thomas Gilbert, Klaus P Koepfli, Mikkel HS Sinding, Shyam Gopalakrishnan
University of Copenhagen, Smithsonian's National Zoo and Conservation Biology Institute, University of Copenhagen, Ludwig Maximilian University Munich, University of Copenhagen, University of Copenhagen, University of Copenhagen, Cardiff University, University of Copenhagen, Smithsonian's National Zoo and Conservation Biology Institute, University of Copenhagen, Smithsonian's National Zoo and Conservation Biology Institute, University of Copenhagen, University of Copenhagen

Domestication has altered many traits in dogs compared to their wild ancestor, the gray wolf. Among these changes is in reproduction, most notably, the loss of reproductive seasonality. To better understand the genetic basis of reproductive changes, we investigated genomic differences in key reproduction-related genes between dogs and wolves. We find one of the most extreme genome-wide signals of differentiation between dogs and wolves is at the thyroid stimulating hormone receptor (TSHR) gene, a key gene related to seasonal reproduction and thyroid hormone regulation. We pinpoint this differentiation at a ~18kb region containing highly divergent ancestry and a ~3kb LINE1 insertion found in dogs. Using ~2,000 canid genomes, we find the divergent haplotype is nearly fixed in dogs, however, it is at low frequency in Australasian dogs. While nearly absent in modern and ancient gray wolves, this divergent haplotype is most prevalent in wolves from India and Southwest Asia. Ancient canid genomes reveal it appeared early in dog domestication ~11,000 years ago and is associated with the secondary source of western dog ancestry. Our findings are consistent with recent work proposing that as dogs expanded into Southwest Asia, gene flow with local wolves acted as a second source of ancestry, which likely introduced new genetic variation for which selection to act upon. Although further work is necessary to quantify the functions of this ~18kb region, our work is the first to detect divergent ancestry at a locus that potentially plays a role in core traits related to domestication.

779 - Dobzhansky-Muller incompatibilities and adaptive introgression facilitate explosive speciation of Lake Baikal amphipods

Valentina Burskaia

Valentina Burskaia, Ilia Artiushin, Dmitry Fedorov, Yakov Korobitsyn, Georgii Bazykin, Hannes Svardal
University of Antwerp, University of Antwerp, Skoltech, Lomonosov Moscow State University, Kharkevich Institute for Information Transmission Problems, University of Antwerp

The adaptive radiation of amphipods in Lake Baikal has brought forward more than 340 species (20% of the worldâ€™s freshwater amphipods), making it one of the largest species flocks after the famous African cichlid radiations. We analyzed multi-species transcriptomic (60 species) and genomic (15 species) alignments of Lake Baikal Amphipoda. Our results indicate that hybridization between two ancient independent lineages of Baikalian amphipods preceded the period when fast speciation started. It was followed by differential fixation of introgressed loci in all main clades of the species flock. We found intriguing signals of Dobzhansky-Muller incompatibilities as well as positive selection on introgressed loci. These findings provide evidence that hybridization likely facilitated adaptive radiation of Lake Baikal amphipods. Both processes have been shown to occur in other adaptive radiations, but their functional role in rapid diversification is still debated: hybridization may also randomly coincide with speciation. In our set of hybridization tests and selection scans we demonstrate, that introgressed material undergoes different types of selection, which is unlikely to be a consequence of neutral processes.

789 - The role of admixture in the ongoing speciation in the adaptive radiation of sailfin silverside fishes in the Malili lakes (Indonesia)

Els Lea R De Keyzer

Els Lea R De Keyzer, Fabian Herder, Astrid BÃ¶hne, Hannes Svardal
Universiteit Antwerpen, Leibniz Institute for the Analysis of Biodiversity Change, Leibniz Institute for the Analysis of Biodiversity Change, Universiteit Antwerpen

Hybridisation before or early in an adaptive radiation can fuel speciation. But what if genetic exchange does not cease after species formation? Intuitively, we expect that species which continually admix will eventually collapse into a single species. Yet, we propose that ongoing gene flow may, in some cases, promote ecological divergence, when new genomic combinations facilitate niche adaptation. To shed light on these mechanisms, we characterise the genomic substrate for rapid ecological adaptation and species formation in the adaptive radiation of sailfin silverside fishes (genus Telmatherina) from the Malili lake-system in Sulawesi (Indonesia). The Malili lakes silverside radiation consists of ~15 species covering a variety of ecological niches. Besides presumed hybridisation in their early divergence, sailfin silversides also show strong signs of ongoing gene flow. This makes the silverside radiation ideally suited as a study system to obtain an integrated picture of the effects of ancient and ongoing gene flow in adaptive radiation. Using whole-genome sequences, we reconstructed the phylogeny of sailfin silversides. We detected signals of hybridisation at the root of the species tree and identified recent admixture between lacustrine species. In addition, we found that river-dwelling Telmatherina show within-population variation in their genetic relationships with the lacustrine species groups and both elevated D-statistic and f-branch statistics confirmed gene flow between riverine and lacustrine species groups. We aim to explore how this ongoing gene flow from the riverine group influences the diversification among the incompletely reproductively isolated lacustrine species.

953 - Genomic stability over 65 million years promotes interspecific hybridization in polyploid Potamogeton

Nikita Tikhomirov

Nikita Tikhomirov, Alison Scott, Alexander Bobrov, Zdenek Kaplan, Polina Novikova
Max Planck Institute for Plant Breeding Research, Max Planck Institute for Plant Breeding Research, Papanin Institute for Biology of Inland Waters RAS, Institute of Botany CAS, Max Planck Institute for Plant Breeding Research

Three recurrent whole-genome duplication events produced three species-rich lineages in Potamogeton, the largest aquatic plant genus. Despite diverging millions of years ago, the species of the same ploidy can hybridize and produce fertile offspring.

To study the effect of polyploidy and admixture on the evolution of Potamogeton, we generated 14 (and counting) PacBio HiFi-based genomes and resequenced 500 samples covering all main lineages in the genus.

We dated two whole-genome duplications by synteny-guided analysis of paralog divergence, yielding ages of 65 Mya (hereafter the â€œancientâ€• event) and 13 Mya (the â€œrecentâ€• event). The ancient duplication is shared by all Potamogeton taxa, and the recent by half of the extant Potamogeton species. The third polyploidization is even more recent and is shared by only two species.

The ancient polyploids have undergone substantial fractionation, as evidenced by the single-copy state of BUSCO genes and the high divergence of paralogs. Nevertheless, we found long self-collinear blocks formed by paralogs lasting from the ancient whole-genome duplication. The recent polyploids still exhibit many duplicated BUSCO genes and also demonstrate little structural variation. Furthermore, the recent genome duplication has likely happened through the hybridization of different species, resulting in the combination of two dissimilar subgenomes.

Despite the expectation that genomic rearrangements follow polyploidization, Potamogeton maintained a highly conserved genome structure. We hypothesize that this genomic stability might enable admixture between species, which could be beneficial under low population sizes and/or mutation rates.

965 - Exploring recent evolution of matrotrophy in swordtails (Xiphophorus)

Nemo Valentin Robles

Nim Robles, Theresa Gunn, Cheyenne Y. Payne, Derek Ly, David Reznick
Stanford University, Centro de Investigaciones CientÃ-ficas de las Huastecas â€œAguazarcaâ€•, A.C., Stanford University, Centro de Investigaciones CientÃ-ficas de las Huastecas â€œAguazarcaâ€•, A.C., Stanford University, Centro de Investigaciones CientÃ-ficas de las Huastecas â€œAguazarcaâ€•, A.C., University of California, Santa Cruz, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, University of California, Riverside, University of California, Riverside

Xiphophorus malinche and X. birchmanni are sister species of swordtail fishes residing at various elevations in river systems in MÃ©xico, which have extensive naturally-occurring hybrid zones. Unlike most teleosts, Xiphophorus bear live young (viviparity). One of the evolutionary benefits of viviparity is it allows for post-fertilization maternal nutrient provisioning, known as â€œmatrotrophyâ€•. Matrotrophy increases offspring survival, but can result in reduction of parent survivorship, frequency of reproduction, and lifetime number of offspring. Temperature, predation, and resource availability are known to influence the evolution of matrotrophy, which vary with elevation. Recent work by Payne et. al. (in prep), found that lab crosses between male X. malinche, a high elevation species, and female X. birchmanni, a low elevation species, are largely inviable while the reciprocal cross produced viable offspring. X. malinche live at 1200 m in cold, resource-poor environments. We hypothesize that they evolved matrotrophy to help their young overcome these

challenges. X. malinche fry are born much larger than other related species, while X. birchmanni fry are similar in size to other Xiphophorus, suggestive of a recent evolution of matrotrophy within X. malinche. Additionally, X. malinche mothers upregulate prolactin-related genes in ovarian tissue, a key hormone in pregnancy, during embryonic development. With this information, we plan to study matrotrophy during embryonic development in different environments to determine whether matrotrophy may be an adaptation to high elevation. X. malinche provides an exciting opportunity to study ecological and genetic drivers of matrotrophy and its role in interspecies reproductive isolation.

979 - Ecological genomic structure of a complex of Myotis bats in the Baja California peninsula and western Mexico

Laura Alejandra Najera Cortazar

Laura A Najera Cortazar, Sergio Ticul Alvarez Castaneda, Ian Carr, Kathryn Maher, Helen Hipperson, Winifred Frick, Simon J Goodman
BIOPOLIS Association - Research Centre in Biodiversity and Genetic Resources (BIOPOLIS-CIBIO), University of Leeds, Centro de Investigaciones Biologicas del Noroeste S. C. (CIBNOR), University of Leeds, University of Sheffield, University of Sheffield, University of California, Santa Cruz, University of Leeds

Cryptic species diversity within bats of the genus Myotis creates challenges in understanding their evolution and ecological interactions. Using mitochondrial and genome-wide SNP data, we investigated the population structure and species boundaries of a complex of Myotis bats in the Baja California peninsula, and two continental sites in Mexico. Mitochondrial haplotypes showed that individuals phenotypically identified as M. peninsularis, mostly form a distinct haplogroup closely related to M. velifer. However, phylogenetic analysis of mitochondrial sequences returned paraphyletic clades for the species. ddRAD SNP data results also did not reliably partition individuals, and identified evidence of introgression between the two in bats sampled in mid-peninsula and a continental population. Our mitochondrial phylogenetic analysis identified two other potential cryptic Myotis lineages that will need further investigation. Bayesian structure and cluster analyses showed high levels of genetic differentiation and population structure within M. californicus populations, detecting two different clusters, but with evidence for both local and long-distance dispersal. In contrast, admixture was detected among M. yumanensis individuals at several sites, reflecting gene flow with the other three Myotis species in the study. There was weak genetic structure among M. yumanensis individuals, with conflicts between nuclear and mitochondrial population assignments, suggesting female philopatry and long-distance male mediated dispersal. Our study represents an important step forward for the understanding of species boundaries in Myotis bats from the American continent that can inform development of taxonomic assessments and conservation strategies for North American bats.

1024 - Selection on hybrid incompatibility in baboons demonstrates Haldane's rule on time scales similar to human-Neanderthal divergence

Kasper Munch

Erik SÃ¸rensen, Kasper Munch
BiRC, Aarhus University, BiRC, Aarhus University

With divergences between 700,000 and 1.3 million years, the six baboon species mirror the relationship between members of the Homo genus. Baboons thus present an opportunity to study the emerging reproductive barriers similar to those present when non-African modern humans intermixed with Neanderthals. Exploiting the diversity across 225 full genomes of the six baboon species, we use local ancestry inference to identify admixture tracts along the individual genomes of yellow and olive baboons. Admixture proportions in Tanzania decline rapidly along a population transect away from the hybridization zone, from 21% in the Tarangire population closest to

the hybridization zone to 4% in the Serengeti population furthest away. The population admixture proportion in 100kb windows correlates with recombination rate, suggesting that the admixture gradient is maintained by negative selection. Following Haldane's rule for reproductive incompatibility, this correlation is strongest on the X chromosome. The population admixture proportion in 100kb windows further correlates with the diversity in unadmixed populations, a proxy for the joint effects of background selection and reduced incomplete lineage sorting. The strength of this correlation increases with distance from the hybridization zone as the selection has had more time to act on introgressed segments. In Ethiopia, autosomal admixture between Olive and Hamadryas baboons (only 700,000 years diverged) show similar evidence of selection, but here the correlation to recombination rate is an order of magnitude stronger for X chromosome admixture. Our results thus demonstrate hybrid incompatibility following Haldane's rule in primates diverged no more than humans and Neanderthals.

1051 - Do species share the same evolutionary fate? Comparing population history to explain variation in geographic range and genetic structure across Amazonian lizards

Ivan Prates

Ivan Prates, Miguel T Rodrigues, Daniel L Rabosky

Lund University, University of Michigan, Universidade de Sao Paulo, University of Michigan

Differences in population history might explain the tremendous variation among species in properties like geographic range size and levels of genetic structure. Alternatively, different species might experience a shared trajectory of origination from a parent species, range expansion, and eventually splitting into new species. To examine whether species share a common fate, we compare patterns of genetic diversity and structure, historical demography, and range expansion across dozens of sympatric lineages of Amazonian lizards. Our sampling comprises 24 reduced-representation genomic datasets including more than a thousand specimens from 8 taxonomic families. To ensure we are comparing equivalent units and given taxonomic knowledge gaps, we begin by delimiting candidate species based on phylogenetic and genotypic clustering analyses. Based on more than 80 resulting candidate species, we find remarkable variation in range size and levels of genetic diversity and structure. Large-ranged species tended to have shallower genetic isolation-by-distance (IBD) and experience demographic and geographic expansions. By contrast, more narrow-distributed species tended to be demographically stable and have steeper IBD. Population history and structure also tended to differ across clades; for instance, arboreal iguanian lizards overall had lower genetic diversity and IBD than small terrestrial geckos. Taken together, these results suggest that population trajectories partly explain variation in species ranges and genetic composition. Species with similar properties often shared evolutionary trajectories, though also following clade-specific trends. This analysis provides insights into the factors behind patterns of species richness and endemism in Amazonia, one of Earthâ€™s most biodiverse regions.

1100 - Genetic effects on complex traits and diseases are similar across segments of different continental ancestries in the Mexican Biobank

Yuridia Selene Posadas García

Yuridia Selene Posadas GarcÃ-a, MarÃ-a J. Palma MartÃ-nez, Valeria AÃ±orve Garibay, Lourdes GarcÃ-a GarcÃ-a, AndrÃ©s Moreno Estrada, Hou Kangcheng, Pasaniuc Bogdan, Mashaal Sohail

Centro de Ciencias GenÃ³micas de la UNAM, Centro de Ciencias GenÃ³micas de la UNAM, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano de la UNAM, Center for Computational Molecular Biology, Instituto Nacional de Salud PÃºblica, National Laboratory of Genomics for Biodiversity (LANGEBIO), CINVESTAV, Bioinformatics Interdepartmental Program, UCLA, Bioinformatics Interdepartmental Program, UCLA, Department of Pathology and Laboratory Medicine, David Geffen School of Medicine at UCLA,

Department of Computational Medicine, David Geffen School of Medicine at UCLA, Department of Human Genetics, David Geffen School of Medicine at UCLA, Centro de Ciencias GenÃ³micas de la UNAM

A hotly debated area in human genomics is the role of diversity in medicine. In particular, how similar are genetic effects on complex traits and diseases among different ancestral backgrounds? Recently a novel computational approach (admix-kit) was developed to compute correlation of causal effects (radmix) across local ancestries. Published results showed similar effects across ancestries in African-European admixed individuals sharing the same environment, concluding that genetic causal effects have limited contribution to reduced accuracy of polygenic scores. Here, we estimated correlation of effect sizes for 9 binary and 11 quantitative complex traits and diseases among Indigenous, African, and European ancestries by applying admix-kit to 5,833 individuals from the Mexican Biobank. Individual genotypes were imputed using Topmed and local ancestries were inferred using Gnomix. We compared segments of each ancestry against an aggregate of the other two ancestries. We showed that for all analyzed traits and diseases, radmix across ancestral backgrounds was not significantly different from one ($p > 0.0025$). Meta-analyzing across traits, we estimated radmix values of 0.88 (95% CI: 0.7290-1) for African segments, 0.977 (95% CI: 0.937-1) for Indigenous segments and 0.943 (95% CI: 0.896-0.982) for European segments. To corroborate our conclusions, we will assess the prediction performance of methods integrating GWAS from different populations in a Bayesian approach such as PRS-CSx compared to those estimating effect sizes for different local ancestral backgrounds separately. Our findings hold significant implications for determining the best approach to estimate polygenic scores in admixed individuals.

135 - Comparative single-cell regulome reveals evolutionary innovations in neural progenitor cells during primate corticogenesis

Xin Luo

Xin Luo, Yuting Liu, Yiming Sun, Cheng Li, Bing Su

Kunming Institute of Zoology, Peking University, Peking University, Peking University, Kunming Institute of Zoology

The cellular and genetic mechanism underlying the human-specific features of cortex development remains unclear. We generated a cell-type resolved atlas of transcriptome and regulome of the developing macaque and mouse prefrontal cortex, and conducted evolutionary analyses with the published complementary human data. We discovered a primate-specific expansion of two neural progenitor subclasses, glia-committed radial glia (RG) and truncated RG. Specifically, the human neural progenitors show extensive transcriptional rewiring in the growth factor and extracellular matrix pathways. Expression of the human-specific progenitor marker ITGA2 in the cortex of fetal mouse promotes progenitor proliferation and an increased upper-layer neuron proportion. We demonstrate that these transcriptional divergences are primarily driven by the activity changes of the distal regulatory elements in the genome. Markedly, the chromatin regions with human-gained accessibility enrich the human-fixed sequence changes, as well as sequence polymorphisms associated with intelligence and neuropsychiatric disorders. Our results uncover evolutionary innovations in neural progenitors and gene regulatory mechanism during primate cortex evolution.

186 - Evolutionary origin of the chordate nervous system revealed by amphioxus developmental trajectories

Yichen (Serena) Dai

Yichen Dai, Yanhong Zhong, Rongrong Pan, Yongheng Fu, Liang Yuan, Juan Du, Meng Li, Xiao Wang, Huimin Liu, Chenggang Shi, Gaoming Liu, Pingfen Zhu, Sebastian Shimeld, Xuming Zhou, Guang Li

Institute of Zoology, Chinese Academy of Sciences, Xiamen University, Xiamen University, Xiamen University, Xiamen University, Xinjiang

Normal University, Institute of Zoology, Chinese Academy of Sciences, Institute of Zoology, Chinese Academy of Sciences, Institute of Zoology, Chinese Academy of Sciences, Xiamen University, Xiamen University, Institute of Zoology, Chinese Academy of Sciences, Institute of Zoology, Chinese Academy of Sciences, University of Oxford, Institute of Zoology, Chinese Academy of Sciences, Xiamen University

Amphioxus is phylogenetically well positioned for advancing our understanding of vertebrate evolution. To understand chordate cell type evolution and examine hypotheses of vertebrate neural cell type origins, we developed and analyzed a single-cell RNA-sequencing dataset from seven amphioxus embryo stages. Numerous new amphioxus cell types were identified, including homologs to vertebrate aorta-gonads-mesonephros, hypothalamus, and neurohypophysis, rooting the evolutionary origin of these structures. Ancestor-descendant reconstruction of cell trajectories of amphioxus and other species then allowed us to deduce expression dynamics of transcription factor genes throughout embryogenesis, identifying three ancient developmental routes forming chordate neurons. This also served as a framework for characterizing cell specification at the mechanistic level and, exploiting this, we used gene knockout to examine the function of Lhx3/4, Msxlx, FoxQ2a, Brachyury-1 and Brachyury-2, key transcription factors involved in neural specification. Our results demonstrate three developmental origins for the vertebrate nervous system: an anterior FoxQ2 dependent mechanism that is deeply conserved with invertebrates, a less-conserved route leading to more posterior neurons in the vertebrate spinal cord, and a mechanism for specifying neuromesoderm progenitors that is restricted to chordates. The evolution of neuromesoderm progenitors may have led to a dramatic shift in posterior neural and mesodermal cell fate decisions and the body elongation process in a stem chordate.

347 - Single-cell RNA sequencing for phylogenomic analyses of uncultivable microbial eukaryotes - A case study with marine planktonic protists (Protista, Ciliophora, Oligotrichea)

Luciana Santoferrara

Shahed U. A. Shazib, Ragib Ahsan , Marie Leleu, George McManus, Laura Katz, Luciana Santoferrara

Purdue University, Smith College, University of Massachusetts Amherst, Smith College, University of Connecticut, Smith College, University of Massachusetts Amherst, na, Hofstra University

Protists comprise most branches of the eukaryotic tree of life, still their astonishing diversity and
billion-year-old evolutionary history remain poorly understood. Phylogenomics is a powerful tool to study evolution, but poses multiple challenges for protists (e.g., large and complex genomes, lack of dedicated databases and bioinformatic tools) in addition to the standard ones (e.g., collecting, storing, and analyzing large amounts of data). We have developed a phylogenomic workflow based on single-cell RNA sequencing data, integrating key steps from cell isolation to species tree inference. We assessed the effectiveness of our pipeline using publicly available and newly generated transcriptomes (11 and 28, respectively) from the Oligotrichea, a diverse group of ciliated protists. This group has been relatively well-studied based on ribosomal RNA gene (rDNA) markers, which we reconstructed by read mapping of transcriptome sequences and used both to supplement species identification and to compare rDNA phylogenies and phylogenomic inferences. We also compared phylogenomic analyses from well-curated orthologs (single-copy genes) and in the presence of paralogs (multi-copy genes), and evaluated the effect of missing data. For this, we analyzed multiple subsets of up to 1,014 conserved protein-coding genes by both concatenated gene trees and Asteroid species tree inferences. Both approaches yielded similar results, and most of the phylogenetic relationships among the Oligotrichea were consistent and strongly supported. Our analyses demonstrate that Asteroid provides robust phylogenetic support, while simplifying curation steps and maximizing the number of gene families represented in the inferences.

505 - Comparing the transcriptomes of Daphnia and zebrafish for the evolutionary conservation of a

gene network for fatty liver disease

Shaleen Glasgow

Shaleen Glasgow, Yavor Hadzhiev, Ferenc Mueller, John K Colbourne
University of Birmingham, University of Birmingham, University of Birmingham, University of Birmingham

Genes that govern fundamental biological processes are deeply rooted in evolutionary history. Comparative evolutionary biology demonstrates that disease-associated genes are evolutionarily conserved across metazoans. The peroxisome proliferator-activated receptor Î± (PPARÎ±) is a major transcription factor that regulates hepatic fatty acid and lipid metabolic processes in vertebrates. These metabolic processes can become perturbed by environmental chemicals, leading to fatty liver disease, fibrosis and hepatocellular carcinoma. Environmental pollution is a global health crisis, and rodents have been the primary models for human chemical safety testing due to our shared mammalian biology. However, the importance of mammalian systems is being challenged by ethical and scientific considerations. Comparative transcriptomics facilitates our understanding of evolutionarily shared responses to toxicants and broadens the utility of alternative model species in predicting adverse human and environmental health. Despite the absence of a liver in invertebrates, the molecular networks governing liver fatty acid metabolism may be evolutionarily conserved and possibly predictive of adverse health in humans. The micro-crustacean Daphnia possesses amino-acid sequence orthologs of human genes for 70% of the PPARÎ± signalling pathway. Moreover, Daphnia signals human pathways regulated by PPARÎ± when exposed to a PPARÎ± agonist. Here, we use comparative single nuclei transcriptomics of Daphnia and Zebrafish embryos to investigate the tissue-specific response to a PPARÎ± agonist and identify homologous liver-like tissues in Daphnia. This research shows promise for the reduction and replacement of mammals, and through conserved molecular biology, it proposes Daphnia as an alternative chemical safety testing of liver toxicity.

514 - Cell type differences underlying color vision diversification in Heliconius butterflies

Wei Lu

Wei Lu, Nicholas W VanKuren, Marcus Kronforst
University of Chicago, University of Chicago, University of Chicago

Animals exhibit a rich diversity of behaviors and perceptual capacities to interact with the complex environment. Mate preference is one such rapidly evolving adaptive behavior during the speciation process. Closely related but reproductively isolated species often evolve unique mating cues and the mate preference for conspecific mating cues.

Heliconius butterflies represent a classic example of adaptive radiation with extraordinary color pattern divergence. These diverse wing colorations are also used as mating cues for male mate choice. Previous studies have shown that Heliconius males with divergent mate preferences differ at multiple layers of the peripheral visual system, including photoreceptor peak sensitivity and inter-photoreceptor connections. To better understand the transcriptomic changes underlying this color vision diversification, we conducted single-nuclei RNAseq (snRNAseq) and profiled more than 80,000 nuclei from both adult male and female Heliconius butterflies, spanning from a polymorphic population to species of subgenus-level divergence (~14 Mya). We found that while most cell type abundances are similar across species and sex, a group of glial cells is highly variable in males and its abundance is associated with the inter-photoreceptor connection variation. Surprisingly, this glia cell type is marked by sens-2 expression, which is a top candidate gene in the genome-wide association study (GWAS) for male mate preference in Heliconius.

Overall, our snRNAseq results provide a comprehensive cell atlas for butterfly retina and enable us to identify cell type abundance and regulatory network changes in a group of butterflies undergoing rapid color vision evolution.

533 - Cell-type specific and temporally dynamic genetic regulatory effects in response to environmental perturbations

Wenhe Lin

Wenhe Lin, Mingyuan Merlin Li, Alexis Battle, Yoav Gilad

University of Chicago, Johns Hopkins University, Johns Hopkins University, University of Chicago

Understanding the molecular effects of genetic variation across different cellular contexts is crucial for elucidating the mechanisms underlying disease-associated loci. Cellular contexts encompass not just cell types but also developmental stages and environmental exposures. Despite significant advancements in genomics, the intricate interplay between genetic variants and their regulatory impact on gene expression within a multi-dimensional cellular context remains underexplored.

To explore genetic effects on gene regulation across a broader array of cellular contexts, we leveraged heterogeneous differentiating cultures (HDCs). HDCs are 3D-aggregates derived from human iPSCs and comprise asynchronously developing cells from all three germ layers. We generated control HDCs and HDCs exposed to nicotine, caffeine, and ethanol separately from 54 individuals. Using single-cell RNA-sequencing data from over 1.8 million cells, we identified 58 distinct cell types and characterized the effects of each exposure on gene expression, gene regulatory networks, cell type abundance, cell state transition, and cell-cell communication within each cell type. We discovered six main cellular developmental trajectories in our data, allowing us to characterize continuous and transient genetic regulatory effects in response to each exposure.

In our preliminary eQTL analyses, we mapped steady-state eQTLs and exposure response eQTLs, some of which appear to be cell-type-specific. We also identified dynamic gene-environment interactions that emerge temporally along specific developmental trajectories. We anticipate that these context-specific gene regulatory effects will help to explain disease-associated loci that have thus far failed to display regulatory potential in steady-state tissues.

618 - Single-cell and spatial transcriptomics of diverse organisms at the Swedish National Genomics Infrastructure (NGI)

Helena Parra-Acero

Helena Parra-Acero

Scilifelab/KTH

At NGI, we provide sequencing services, library preparation, and bioinformatic support to researchers in Sweden and abroad. One of our areas of focus is transcriptomics, including single-cell and spatial methods. These methods primarily target human or mouse samples, but we have successfully processed samples from other species including salamander, cockroaches, and sponges.

Single-cell RNAseq methods allow the characterisation of gene expression of hundreds/thousands of cells per sample, while retaining cellular resolution. Starting from a cell suspension, each cell is compartmentalised into a droplet via microfluidics (Chromium 10X) or FACS-sorting in a 384-well plate (Smart-seq3) and the transcriptome is barcoded individually. These methods can be used on non-human/mouse samples if quality requirements are met.

Spatial transcriptomic methods add a spatial dimension to a tissue's sequencing data. The 10x Visium assay for fresh-frozen tissue samples relies on capturing the polyadenylated RNA from a tissue section which has been placed on a slide coated with barcoded probes. The library obtained from this sample can be mapped spatially and combined with a histological image to gain spatial resolution of the tissue. Among the Visium assays available, only this one can be used for non-human/mouse samples.

The quality of the sample is vital for obtaining high-quality data. When sequencing non-standard samples, it is essential to understand the challenges that each specific method may pose, so discussing possible limitations with us in the beginning of a project can be of great benefit.

707 - Exploring single-cell gene expression of immune adaptations in Indigenous Communities from México and Colombia

Octavio Zambada-Moreno

Octavio Zambada-Moreno, Jose Antonio Corona-Gomez, Aaron Espinosa-Jaime, Carlos Humberto Ortiz-Ramirez, Andres Moreno-Estrada, Marisol Espitia-Fajardo, Angel Zamudio-Jacome, Juan Pablo PeÃ±a-Sanchez, Yuberlis Gonzalez-Ipuana, Guillermo Barreto

Centro de InvestigaciÃ³n y de Estudios Avanzados del IPN, Centro de InvestigaciÃ³n y de Estudios Avanzados del IPN, Centro de InvestigaciÃ³n y de Estudios Avanzados del IPN, Centro de InvestigaciÃ³n y de Estudios Avanzados del IPN, Centro de InvestigaciÃ³n y de Estudios Avanzados del IPN, Universidad del Valle, Universidad del Valle, Universidad del Valle, Universidad del Valle, Universidad del Valle

Understanding the genetic basis of adaptive phenotypic variation has been a long goal in evolutionary biology. In humans, this has been particularly challenging because genomics studies and resources have been focused mainly on European ancestries, missing most of the available genetic variation in Latin American populations. To help close this gap, we are profiling both genetic and regulatory variation at the level of single-cell gene expression in indigenous Mexican and Colombian populations, representing diverse geographic and cultural regions.

Since the immune system is one of the most responsive to rapid adaptation events, we will analyze gene expression changes in immune cell types as a function of diet, environment, and thus, different pathogen exposure. We are genotyping and comparing transcriptional profiles across 300 individuals from communities in MÃ©xico and Colombia through a single-cell RNAseq approach. To reach communities in remote and isolated locations, we have successfully implemented a new protocol for the isolation and cryopreservation of peripheral blood mononuclear cells (PBMCs) on-site, eliminating the need for specialized equipment and minimizing infrastructure constraints, providing efficiency and flexibility during fieldwork.

We aim to leverage both genomic and transcriptional information to identify expression quantitative trait loci (eQTLs) at the cell-type level. This will enable enhanced resolution in quantifying ancestry-specific abundances of crucial molecules controlling immune function, such as receptors and immune response activators. This work is part of a broader project led by the LatinCells consortium, which aims to create a comprehensive human cell map of Latin American diversity.

841 - How one becomes another: exploring the role of exaptation on the evolution of gene regulation

Cauã Antunes Westmann

CauÃ£ Antunes Westmann, Leander Goldbach, Andreas Wagner

Department of Evolutionary Biology and Environmental Studies, University of Zurich, Swiss Institute of Bioinformatics, Department of Evolutionary Biology and Environmental Studies, University of Zurich, Swiss Institute of Bioinformatics, Department of Evolutionary Biology and Environmental Studies, University of Zurich, Swiss Institute of Bioinformatics, The Santa Fe Institute

Exaptation, the co-option of existing traits for new functions, is a central process in Darwinian evolution. However, the molecular changes leading to specific exaptations remain unclear. Here, we investigated the potential of bacterial transcription factor binding sites (TFBSs) to evolve exaptively, examining three global transcription factors (TFs) from Escherichia coli: CRP, Fis, and IHF. Using a massively parallel

single-cell reporter assay, we mapped three combinatorially complete adaptive landscapes, encompassing all intermediate sequences between three pairs of strong binding sites for each TF. Our results revealed that these landscapes are smooth and navigable, with a monotonic relationship between mutations and their impact on gene regulation. Starting from a strong TFBS for one of our TFs, Darwinian evolution can thus create a strong binding site for another TF through a small number of individually adaptive mutations. Notably, most intermediate genotypes are prone to transcriptional crosstalk â€“ gene regulation mediated by both TFs. Because our landscapes are smooth, Darwinian evolution can also easily create binding sites that show such crosstalk whenever it is adaptive. Our study presents the first in vivo evidence that new TFBSs can evolve exaptively through multiple small and adaptive mutational steps. They also highlight the importance of regulatory crosstalk for the diversification of bacterial gene regulation.

932 - Lineage-resolved analysis of embryonic gene expression evolution in the Caenorhabditis nematodes

Christopher Ryan Livingston Large

Christopher RL Large, Rupa Khanal, LaDeana Hillier, Chau Huynh, Connor Kubo, Junhyong Kim, Robert H Waterston, John I Murray
University of Pennsylvania, University of Pennsylvania, University of Washington, University of Washington, University of Washington, University of Pennsylvania, University of Washington, University of Pennsylvania

What constraints govern the evolution of gene expression patterns across development remains a fundamental question of evolutionary biology. Single-cell sequencing can help define these constraints through systematic profiling of homologous cells across different organisms. The nematode, C. elegans is a well-studied model for embryonic development, and has an invariant lineage that is conserved with other Caenorhabditis species allowing for the direct comparison of gene expression between homologous progenitor and terminal cell types across evolution. This conserved and defined lineage will enable the Caenorhabditis species to act as a gold-standard for troubleshooting computational challenges in comparative single-cell biology such as batch correction, cell type annotation transfer accuracy, and metrics to describe divergence in cellular transcriptomes. Here, we have measured the spatiotemporal divergence of gene expression across embryogenesis by collecting, annotating, and comparing the transcriptomes of homologous embryonic cell types, using a dataset comprising >200,000 C. elegans and >190,000 C. briggsae cells. We find a high level of similarity in gene expression programs between the species despite tens of millions of years of evolutionary divergence. Even still, thousands of genes show divergence in their cell-type specific expression patterns, and these are enriched for categories involved in environmental response and behavior. Overall, our work identifies likely constraints on the evolution of developmental gene expression and provides a powerful resource for addressing diverse evolutionary questions. Further work will investigate the interplay between genomic rearrangements, protein sequence evolutionary rate, and gene regulatory network structure in cell type transcriptome conservation.

1018 - Improved characterization of single-cell RNA-seq libraries with paired-end avidity sequencing

John Chamberlin

John Chamberlin, Austin Gillen, Aaron Quinlan
University of Utah Department of Biomedical Informatics, University of Colorado Department of Medicine, University of Utah Department of Human Genetics, University of Utah University of Utah Department of Biomedical Informatics

Here, we present an unconventional paired-end alignment strategy enabled by the improved homopolymer handling of avidity sequencing (Element Biosciences). Sequencing past the poly(T) primer with the forward read allows direct inspection of the putative site at which the primer hybridized to the RNA molecule. We show that this increases confidence in annotation optimization strategies by ruling out internal priming or other artifacts as well as directly assigning the recovered reads to polyadenylation sites. We also share important adaptations

to the alignment workflow needed for accurate paired end alignments.

1086 - Evaluation of a deeply conserved vertebrate signal of neuronal activation in plethodontid salamanders

Apoorva Ganesh

Apoorva Ganesh, Stella M. Lai, Damien B. Wilburn
Ohio State University, Ohio State University, Ohio State University

Chemical signals for intraspecies communication are important to the reproductive strategies of many vertebrate taxa and likely evolve by sexual selection. Vertebrate olfactory systems possess many sensory neurons tuned to detect species-specific pheromones emitted by conspecifics, but biochemically identifying pheromone-receptors pairs is challenging. Consequently, little is known about the molecular co-evolution of pheromones and their receptors. Plethodontid salamanders are a valuable model system where multiple families of male protein pheromones can be recombinantly synthesized and alter female mating behavior via stimulation of olfactory neurons. The physiological response of neuronal activation in salamanders is still under research. Further, verified methods to measure this activation are limited. The current best approach requires harsh fixation methods that prohibit further isolation of genetic material from activated neurons to identify putative pheromone receptors. Here we evaluate the phosphorylation of Ribosomal Protein S6 (RPS6) as a potential probe for neuronal activation in plethodontid salamanders. C-terminal phosphorylation of RPS6 was the first discovered protein phosphorylation and has historically been associated with neuronal activation in other vertebrate taxa. Phylogenetic comparisons of plethodontid RPS6 support deep evolutionary conservation of the C-terminus including key serine residues. Compared to established methods, biochemical and immunohistochemical experiments with commercial anti-pRPS6 show similar levels of neuronal activity in response to pheromone treatment. Immunohistochemical labeling with anti-pRPS6 requires relatively gentle formaldehyde fixation of olfactory tissue that will enable future studies to specifically isolate pheromone-activated neurons for identification of receptor genes.

1092 - LatinCells: empowering research capacity in Latin America through community engagement, infrastructure development, and single-cell RNA profiling of diverse tissues for the Human Cell Atlas.

Adolfo Rojas

Domenica Marchese, Adolfo Rojas-Hidalgo, Ricardo Verdugo, Alanis HuaÃ±aco, Alicia Colombo, Gerardo Donoso, Katherine Marcelain, Marcelo Fonseca, Diego PÃ©rez-Stuardo, Emiliano Vicencio, SebastiÃ¡n Leiva-Navarrete, Guillermo Barreto, Carla Gallo, Patricia Severino, Patricia Possik, Hugo Guerrero-CÃ¡zares, LucÃ-a Spangenberg, Daniela Robles-Espinoza, JosÃ© Antonio Corona-GÃ³mez, Carlos Ortiz-RamÃ-rez, AndrÃ©s Moreno-Estrada, Vinicius Maracaja-Coutinho
Universidad de Chile, Universidad de Chile, Universidad de Talca, Universidad de Talca, Universidad de Chile, Universidad de Chile, Universidad Arturo Prat, Universidad de Chile, Universidad de Chile, Universidad de Chile, Universidad del Valle, Universidad Peruana Cayetano Heredia, Hospital Israelita Albert Einstein, Brazilian National Cancer Institute (INCA), Mayo Clinic, Institut Pasteur de Montevideo, Universidad Nacional AutÃ³noma de Mexico, Advanced Genomics Unit (UGA-LANGEBIO), Advanced Genomics Unit (UGA-LANGEBIO), Advanced Genomics Unit (UGA-LANGEBIO), Universidad de Chile

Latin America boasts unparalleled ethnic and genetic diversity, a result of centuries of admixture between indigenous populations and other continental ancestries. However, this richness remains largely unrepresented in the Human Cell Atlas, impeding the development of tailored medical interventions and perpetuating healthcare disparities. The LatinCells initiative seeks to rectify this by constructing comprehensive cellular maps with samples from diverse indigenous and admixed populations across Brazil, Chile, Colombia, Mexico, Peru, Uruguay, and among U.S. Latinos. We have implemented a multifaceted approach, which includes actively engaging with indigenous

communities in remote and rural locations, deploying advanced technological solutions for scRNAseq data generation from immune cells and gallbladder samples, and developing bespoke computational analysis pipelines. We are also building local capacity through the creation of training materials and workshops focused on single-cell RNA sequencing analysis, tailored to the needs of researchers in Latin American countries. Our preliminary results are focused on pilot samples for which we have single-cell expression atlases of immune cells derived from various Latin American ancestries. Through LatinCells, we aim not only to bridge the gap in representation within the Human Cell Atlas but also to foster a collaborative network that empowers local communities and researchers to strengthen research capacity across Latin America.

Keywords: Single-Cell; Immune System; Transcriptomics; Bioinformatics Education.

Funding: This project is funded by the Chan Zuckerberg Initiative (CZI).

1129 - Disentangling the gene regulatory network controlling intrinsic antiviral immunity in stem cells

Mónica Padilla Gálvez

MÃ³nica Padilla GÃ¡lvez, Qing Yang, Daniel Blanco-Melo
Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, Postbaccalaureate program, Fred Hutchinson Cancer Center, Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center

Stem cells defend from viral infections by intrinsically expressing antiviral genes independent of viral recognition or signal transduction. As these cells differentiate, these intrinsically expressed antiviral genes (IAGs) become interferon signaling dependent. Previous work indicates an incompatibility of pluripotency with interferon response, highlighting both the need for IAGs expression and that these processes might have evolved from similar pathways. However, two main questions remain: (1) how do stem cells regulate IAGs? (2) What regulatory switch during differentiation makes IAGs expression interferon signaling dependent? Here, we aim to elucidate the gene regulatory network controlling IAGs expression dynamics during early differentiation using bulk RNA-seq and multiome analysis. We first identified putative IAGs comparing gene expression between human embryonic stem cells (hESCs) at baseline and infected differentiated cells. We further classified them by expression levels and characterized their behavior in hESCs upon influenza A virus infection. We found 20 transcription factors (TFs) regulating IAGs, including key pluripotency factors, such as KLF4 â€“ a Yamanaka factor shown to inhibit interferon signaling. Our analysis also highlights that some IAGs could be further induced upon infection, including IFITM1 and 28 genes known to have a strong antiviral effect in ESCs. Moreover, co-expression with IFN during infection suggests the underlying IAGs network converges with the canonical interferon pathway. In general, our results indicate that maintaining IAGs expression is a key property of pluripotency. This regulatory network permits limited IAG inductions via interferon signaling upon infection, highlighting a unique antiviral immune response mounted by stem cells

168 - Population Genetics and Forensic Analysis of Western African Populations

Moutanou Modeste Judes ZEYE

Moutanou Modeste Judes ZEYE, Serge Yannick Ouedraogo, Jacques SimporÃ©
Human Evolution, Department of Organismal Biology, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden, Laboratory of Molecular Biology and Genetics (Labiogene), University Joseph KI-ZERBO, 03 BP 364, Ouagadougou 01, Burkina Faso, Department of Oncology, School of Clinical Medicine, Shandong Cancer Hospital, Shandong First Medical University, Shandong Academy of Medical

Sciences, 6699 Qingdao Road, Huaiyin District, Jinan 250000, Shandong, China, Laboratory of Molecular Biology and Genetics (Labiogene), University Joseph KI-ZERBO, 03 BP 364, Ouagadougou 01, Burkina Faso

In West Africa, most population groups speak languages belonging to the Niger-Congo family, which are divided into five groups: Mande, Gur, Kru, West Atlantic, and Dogon2. In Burkina Faso (BF), a country located in the middle of West Africa, there are 63 ethnic groups speaking local languages belonging to these five groups. Despite the large population size, 20 million people reported in the last census, genetic data based on autosomal short tandem repeats (aSTR), X-chromosome STRs (X-STRs), and Y-chromosomes (Y-STRs) are still lacking in genetic and forensic research. This research gap also presents a security concern in the Sahel region, due to the lack of effective measures to identify individuals and the remains of deceased persons and to conduct family relation searches. To fill this information gap, 500 unrelated individuals from BF were genotyped using three microsatellite marker sets: MICROREADER 21 ID, 19X, and 29Y Direct ID System kit. None of the investigated individuals shared common microsatellite profiles. Based on our results, MICROREADER ID System kits are highly polymorphic and informative. Population genetic analysis with comparative Western African populations detected population structure and showed that geographical proximity is a reliable indicator of genetic affinities between populations. Our findings underscore also the necessity for expanded genetic research in West Africa, including genome-wide analyses, aiming to enhance both scientific understanding and forensic applications.

479 - A computational approach for positive genetic identification and relatedness detection from low-coverage shotgun sequencing data

Remy Nguyen

Remy Nguyen, Joshua D. Kapp, Samuel Sacco, Steven P. Myers, Richard E. Green
University of California, Santa Cruz, University of California, Santa Cruz, University of California, Santa Cruz, California Department of Justice Jan Bashinski DNA Laboratory, University of California, Santa Cruz

Recent developments in next-generation sequencing technologies have enabled the recovery of genetic data from minute, contaminated, or heavily degraded samples, removing roadblocks to many forensic investigations. Nevertheless, current identity analysis methods, often relying on hyper-variable short tandem repeat (STR) or single nucleotide polymorphic (SNP) markers, can fail when trace quantities of DNA produce inaccurate or
incomplete genotype calls at the sites used for comparison. This is particularly true for difficult but forensically relevant samples, such as single, rootless hairs. Here, we present IBDGem, a fast and robust computational procedure for detecting genomic regions of identity-by-descent from low-coverage shotgun sequence data. At less than one-fold genome coverage, this method reliably locates segments of relatedness and can make high-confidence identity detections with as little as 0.01-fold genome coverage.

545 - Enhancing Py_ped_sim Simulations to Create More Dynamic Pedigrees

Sthen A Campana

Sthen A Campana, Miguel Guardado, Elena Zavala, Rori Rohlfs
University of Oregon, University of California San Fransisco, University of California Berkeley, University of Oregon

Investigative genetic genealogy (IGG) relies heavily on the accuracy of pedigrees used for analysis. Following its pivotal role in the identification of the Golden State Killer in 2018, this method has gained significant traction in forensic identification. However, little work has been done to understand how the structure of a pedigree influences the accuracy of results. Py_ped_sim has been created to simulate pedigrees structures and the genomes of individuals in a pedigree. One limitation of py_ped_sim and other pedigree simulators

is that only descendants from one starting pair of root founders are simulated. This leads to an underestimation of cousins for individuals because we would have family data on one parent, but none on the other parent who is a non-root founder, that is, not a descendant of the root founder. I present two software programs that will aid in generating dynamic pedigrees via adding in parameters for misattributed paternity (MAP) and extending the family's breadth. Family broadening simulates families onto these non-root founders and gives us a more accurate number of cousins for each individual.

This software is made with a python interface and will be incorporated into the py_ped_sim software framework. We present validations of how extending the breadth of the family creates more distant cousin relationships providing a more accurate representation of genetic relatives. Overall, these suites of software will offer a valuable resource for advancing the field of IGG by providing dynamic pedigrees for comprehensive genetic investigation.

691 - Improved identification of victims from the Spanish Civil War with HTS technology

Fernando González-Candelas

Sandra CarbÃ³-RamÃ-rez, LlÃºcia MartÃ-nez-Priego, Alan CodoÃ±er-Alejos, Fernando GonzÃ¡lez-Candelas

FISABIO, FISABIO, FISABIO, University of Valencia, FISABIO

High-throughput sequencing (HTS) offers new opportunities in forensic genetics, providing several advantages over traditional methodologies. We have assessed the applicability and performance of HTS for the identification of victims from mass graves of the Spanish Civil War (1936â€“1939). Here, we present the genetic identification of victims buried in 2 mass graves located in Valencia which were analyzed previously by traditional technologies. A total of 102 post-mortem remains and 76 relatives (41 families) were sequenced using HTS with Verogen-Qiagen reagents and sequencer, which included 59 STRs (28 autosomal, 24 Y-chromosome, 7 X-chromosome) and 94 SNPs. We used the DVI module of Familias software for kinship analysis. Informative profiles, with >50% genotyped markers, were obtained in 70 remains. To the 30 identifications obtained with capillary electrophoresis sequencing of 32 STR markers, we have added the identification of 14 victims. These identifications were obtained after comparison with reference samples from first-, second- and third-degree relatives. As expected, parent-child relationships were more robustly identified than other relationships. The main limitations for the identification of victims from these mass graves were the low number of first-degree living relatives and the low DNA concentration and high levels of DNA degradation in the remains. Despite this, NGS-based genotyping provides numerous advantages that improve the interpretation of challenging or degraded DNA and have the potential to increase the power of discrimination for identification and kinship analysis, especially for low-quality samples and distant relatives.

935 - Quantifying the Accuracy of Investigative Genetic Genealogy (IGG) through Population Genetic Simulations via py_ped_sim

Miguel Alberto Guardado

Miguel A Guardado, Elena Zavala, Sthen Campana, Ryan D Hernandez, Rori Rohlfs

University of California, San Francisco, University of California, Berkeley, University of Oregon, San Francisco State University, University of California, San Francisco, University of Oregon

Investigative Genetic Genealogy (IGG) is the practice of identifying people of interest using distant relatives' genetic data from genetic genealogy datasets. This tool has grown in popularity for forensic identification after it was used to identify the Golden State Killer in 2018. Still, the accuracy and reliability of this technique are unknown, with existing evaluations of the limitations focused on individuals with European ancestry and simple two-children-per-family pedigree models. We aim to quantify the accuracy of IGG with imperfect SNP

panel data using population genetic simulations of realistic pedigrees with varying genetic ancestry. We present py_ped_sim, an open-source software with a dynamic pedigree structure and genome simulation framework. py_ped_sim is designed to support investigation of the accuracy of IGG, but can also be applied to genetic pedigree analyses in evolutionary and medical genetics. We validate our genomic pedigree simulation, showing a strong correlation between observed kinship and expected genetic kinship across six simulated families (average R2=0.90). We then utilize py_ped_sim to compare how kinship estimates were impacted by the genetic ancestry of founders across five super-continental population groups from the 1000 genomes consortium. We quantified differences in kinship estimates across the genetic ancestry of founders, considering the five super-continental groups. In addition, we use this simulation framework to test the accuracy of other popular kinship methods by generating pedigrees, including up to 6th-degree relationships. Overall, we provide a template for criminal justice practitioners to develop guidelines on the accuracy/limitations of IGG methodologies.

982 - Comparing accuracy of forensic DNA mixture analysis across groups with varying genetic diversity.

Kamillah T Felix

Maria Flores, Cara Ly, Evan Ho, Niquo Ceberio , Kamillah Felix, Hannah Mariko , Miguel Guardado, Matt Paunovich, Chris Godek , Carina Kalaydjian, Rori Rohlfs

San Francisco State University, University of California, Los Angeles, San Francisco State University , San Francisco State University, University of California, Santa Cruz, San Francisco State University , San Francisco State University, George Washington University , University of California, San Francisco , San Francisco State University , San Francisco State University , San Francisco State University , San Francisco State University , University of Oregon

Police are increasingly using trace amounts of DNA in investigations, often when samples contain DNA from multiple contributors. The reliability of interpreting these DNA mixtures is complex. A likelihood ratio (LR) is a value used to assess whether genetic evidence at the crime scene supports the defense hypothesis (the suspect did not contribute to the DNA mixture), or the prosecutor's hypothesis (the suspect did contribute to the mixture). The LR is calculated based on the allele frequency distribution of the assumed group that the person of interest (POI) belongs to, which will be referred to as the reference group. We hypothesized higher error rates when the reference group is incorrectly assumed (different from the POIâ€™s true genetic background), and even higher when the true group has low genetic diversity. We also predict that as the number of contributors in a mixture increases, the amount of false positives increases. To test our hypothesis, we used Forensim â€“ a free open source R package â€“ to simulate individual genotypes, forensic DNA mixtures, and calculate the LR. We observed that when there are more individuals in the DNA mixture, there is less reliability in assessing the evidence where the POI did not contribute. With the correct reference group, there was an increase in the false positive rate for groups with low genetic diversity compared to groups with high genetic diversity. Our results indicate that forensic DNA mixture analysis tools used today may falsely identify individuals with certain genetic backgrounds.

179 - Geographic and Environmental Impact on Genetic Variation in Khoe-San Descendant Communities in South Africa

Stacy L Edington

Stacy L Edington, Dana R Al-Hindi, Neus Font-Porterias, Justin W Myrick, Caitlin Uren, Paul J Norman, Marlo MÃ¶ller, Brenna M Henn, Austin W Reynolds

Baylor University , University of California Davis, University of California Davis, University of Colorado School of Medicine, University of

California Davis, Stellenbosch University, University of Colorado School of Medicine, Stellenbosch University, University of California Davis, Baylor University

While recent studies have focused on reconstructing the population history of Bantu-speaking populations of southern Africa, very little work has been done on the recent population history of Khoe-San communities. In particular, there has been limited research on the effects of European colonization and the Dutch-East India slave trade on the population structure of Khoe-San and Khoe-san descendant (K-SD) communities. We have sampled over 3,000 Khoe-San and K-SD individuals from South Africa, including 150 newly genotyped for this study to better understand how geographic and environmental factors may be impacting genetic variation in these communities across South Africa. We first used MAPS to infer effective migration rates across space using IBD tract information. Our MAPS results show relatively more migration along the southern coastal plain of South Africa and within the arid interior region, with barriers to migration between these regions. These observed decreases in effective migration rates are consistent with the location of the Cederberg and Roggeveld mountains that divide the coastal plain from the interior of the country. We next applied SPRUCE, a machine learning approach to integrate genetic and environmental data, to determine which factors may be influencing the migration patterns in our study region. Our results show that for the past 30 generations, landscape steepness and precipitation variation are the most important factors explaining effective migration patterns. Together, the environmental variables included in our model explain ~33% of the variation in effective migration in our dataset.

180 - Ancient mitogenomes highlight legacy of Transatlantic Slave Trade in Peru

Laura N Pott

Laura N. Pott, Jaime Zolik, JosÃ© Luis Santa Cruz AlcalÃ¡, Claire Maass, Maria A. Nieves-ColÃ³n
University of Minnesota, University of Minnesota, Universidad Nacional Mayor de San Marcos, APACHETA Consultores, Museo Virtual Afroperuano de San Luis de CaÃ±ete, SNA International, APACHETA Consultores, Museo Virtual Afroperuano de San Luis de CaÃ±ete, University of Minnesota

Contemporary Latin American populations have complex admixture patterns due to colonization, indentured servitude, and the Transatlantic Slave Trade. However, the population history of Afro-descendant communities remains understudied. As part of a community-engaged research project and the first ancient DNA (aDNA) study of an Afro-diasporic community in South America, we use aDNA to investigate patterns of mitogenomic diversity among individuals buried in the cemetery of the 18th century Hacienda La Quebrada sugar plantation in Peru. We isolated aDNA and prepared single-stranded and double-stranded libraries for 30 individuals, followed by shotgun sequencing and mitochondrial enrichment. We recovered 36 mitogenomes (average depth of 190.5X), retaining 32 mitogenomes (average depth of 214X) representing 28 individuals for further analysis after considering coverage and contamination. We successfully obtained 32 mitochondrial haplogroup assignments with an average quality of 93% using Haplogrep 3 and a hard-filtering variant coverage threshold of 5X. 75% of the individuals carry haplogroups L0, L1, L2, or L3 (associated with African populations); 21.4% carry A2, B2, C, or D1 (associated with Indigenous American populations); and 3.6% carry H1bw (associated with Iberian populations). Median joining networks were generated using popART to visualize phylogenetic relationships. These diverse results highlight several mitolineages that have not been commonly reported among enslaved populations in the Americas, including Eastern Africa and the Iberian Peninsula. They also suggest that contemporary admixture patterns in Afro-Peruvians are longstanding, providing insight into the demographic history and genomic impacts of slavery in Afro-diasporic communities in Latin America.

519 - Genetic signatures from the Dutch East India slave trade in southern Africa

Dana R. Al-Hindi

Dana R. Al-Hindi, Stacy Edington, Austin W. Reynolds, Neus P. Font, Brenna M. Henn
University of California, Davis, Baylor University, Baylor University, University of Colorado, University of California, Davis

People who self-identify as South African Coloured (SAC) are often described as five-way admixed, harboring majority of their genetics from indigenous Khoe-San (avg. 46%) populations, with other ancestries from equatorial African populations, Europeans, and South and East Asians. Our primary focus are the equatorial African, and South and East Asian ancestries that were introduced after 1652 through the Dutch East India slave trade. About half of the enslaved individuals brought to the Cape Colony between 1652 and 1808 were from South and East Asia, but nearly all of these were from the Indonesian archipelago (23%) or the Indian subcontinent (26%). Previous studies have frequently used the Han Chinese as a reference for the South and East Asian signals among the SAC people, but given the region's slave trade and historical record, we hypothesize populations from Indonesia and/or the Indian subcontinent would be more appropriate. Here, we fine tune which South and East Asian populations best serve as source population for those genetic components in the SAC. The SAC admixed individuals in our analysis stem the Northern and Western Cape (n=619) of South Africa. Reference populations were extracted from the 1000 Genomes project, alongside 162 whole genome Indonesian samples from the Indonesia Genetic Diversity Project. We combine the use of highly accurate local ancestry inferences and ancestry specific MDS to establish if populations other than the Han Chinese are more representative of the Asian components in the SAC.

654 - Ghost admixture in present-day African genomes

Aigerim Rymbekova

Aigerim Rymbekova, Gwenna Breton, Paul Verdu, Mattias Jakobsson, Carina Schlebusch, Martin Kuhlwilm
Department of Evolutionary Anthropology, University of Vienna, Human Evolution and Archaeological Sciences (HEAS), University of Vienna, Human Evolution, Department of Organismal Biology, Uppsala University, National Centre for Scientific Research (CNRS), National Museum of Natural History (MNHN), Paris CitÃ© University, Human Evolution, Department of Organismal Biology, Uppsala University, Human Evolution, Department of Organismal Biology, Uppsala University, Department of Evolutionary Anthropology, University of Vienna, Human Evolution and Archaeological Sciences (HEAS), University of Vienna

Admixture has played a critical role in shaping the genetic diversity among human populations. Exploring the genomes of diverse and deeply divergent present-day human populations such as the Khoe-San populations of Southern Africa and Central African rainforest hunter-gatherers (RHG) is key to understanding our evolutionary history as humans. Although several studies have previously suggested contributions from unknown deeply diverging lineages into the ancestors of different present-day African populations, there is no coherent picture due to differences in datasets, methods and frameworks considered. This problem is particularly difficult since no source genomes from archaic populations are available.
Using methods that do not require the reference archaic genomes, we analyzed 25 high-coverage genomes from five different Khoe-San populations and a set of genomes from different RHG populations, in the context of African genomic diversity. We explored two models of early human population structure: the reticulation of multiple stems migrating between early human populations, and archaic hominin admixture from deep ghost lineages. We found that a model with introgression from an unknown source population into the ancestors of Khoe-San populations better explains patterns of outliers in terms of private polymorphisms than the weakly structured stem model. We identified genomic segments that represent putative fragments of ghost introgression in the Khoe-San individuals.

1084 - Evolutionary History of Transient Receptor Proteins in Sub-Saharan Africa

Jasmin Rees

Jasmin Rees, Shaohua Fan, Jeffrey Spence, Yuanqing Feng, Matthew Hansen, Jonathon Terhorst, Alessia Ranciaro, Jibril Hirbo,

William Beggs, Neil Thomas, Thomas Nyambo, Sununguko Wata Mpoloka, Alfred Njamnshi, Charles Fokunang, Gurja Belay, Yun Song, Sarah Tishkoff, TOPMED CONSORTIUM
University of Pennsylvania, Fudan University, University of Pennsylvania, Stanford University, University of Pennsylvania, University of Pennsylvania, University of Michigan, University of Pennsylvania, University of Pennsylvania, University of Pennsylvania, University of California, Kampala International University in Tanzania, University of Botswana Gaborone, The University of Yaounde Ì•, The University of Yaounde Ì•, Addis Ababa University, University of California, University of Pennsylvania, University of Pennsylvania

Transient receptor potential (TRP) channels are expressed across a number of different skin cells, including sensory neurons, keratinocytes, melanocytes and skin-resident immune cells. As such,
skin expressing TRP channels not only act as chemical, thermal or mechanical sensors, but also play a key role in the growth and development of skin and hair, as well as in mediating immunological and inflammatory responses of the skin. Various TRP genes have hence been implicated across a wide range of skin disorders and diseases, including atopic and pruritic dermatitis, psoriasis, alopecia, rosacea, and skin cancer.
Using a large number of high-quality whole-genome data from diverse sub-Saharan Africans, together with publicly available non-African genomes, to trace the evolutionary history of TRP genes across sub-Saharan Africa. We demonstrate that TRP genes contain alleles that are highly differentiated amongst African populations, including within and between major languages groups, and have likely been targets of positive selection. We evaluate the role of abiotic factors in driving the genetic variation of TRP genes across sub-Saharan Africa, considering the potential thermal and UVB selective pressures acting on TRP genes with known thermoregulatory or melanocyte function. Finally, using large-scale BioBank data, we identify associations between common and selected variants within TRP genes and disease traits, specifically skin-based disorders.
With thanks to the TOPMED Consortium and funding sources ADA 1â€“19-VSN-02, NIH grants 1R35GM134957, R01DK104339, and R01AR076241, and 1X01HL139409-01.

625 - Small-scale genetic histories from Roman Britain

Marina Silva

Marina Silva, Thomas Booth, Joanna Moore, Kyriaki Anastasiadou, Christopher Barrington, Sharon Clough, Alexandre Gilardet, Paolo Guarino, Michael Henderson, Sarah Johnston, Monica Kelly, Jesse McCabe, Alex Smith, Leo Speidel, Pooja Swali, Frankie Tait, Don Walker, Mia Williams, Alistair Barclay, David Bowsher, Janet Montgomery, Pontus Skoglund
Ancient Genomics Laboratory, The Francis Crick Institute, Ancient Genomics Laboratory, The Francis Crick Institute, Department of Archaeology, University of Durham, Ancient Genomics Laboratory, The Francis Crick Institute, Bioinformatics and Biostatistics, The Francis Crick Institute, Cotswold Archaeology, Ancient Genomics Laboratory, The Francis Crick Institute, Cotswold Archaeology, Museum of London Archaeology (MOLA), Ancient Genomics Laboratory, The Francis Crick Institute, Ancient Genomics Laboratory, The Francis Crick Institute, Ancient Genomics Laboratory, The Francis Crick Institute, Headland Archaeology, Ancient Genomics Laboratory, The Francis Crick Institute, Genetics Institute, University College London, Ancient Genomics Laboratory, The Francis Crick Institute, Ancient Genomics Laboratory, The Francis Crick Institute, Museum of London Archaeology (MOLA), Ancient Genomics Laboratory, The Francis Crick Institute, Cotswold Archaeology, Museum of London Archaeology (MOLA), Department of Archaeology, University of Durham, Ancient Genomics Laboratory, The Francis Crick Institute

The Imperial Roman period witnessed significant political and societal changes, fueled by road and maritime network development associated with military campaigns, governance and trade, resulting in an intensification of human movement across Western Eurasia, including long-range mobility. We aim to study the Roman period in Britain, at the northwestern fringe of the Roman Empire, from a

genomics perspective by building a bottom-up sampling approach, focusing on intra-site questions arising primarily from the archaeological context and stemming from close partnerships with collaborators from academic, commercial and cultural heritage sectors.

Here we showcase two examples of how our multidisciplinary approach, combining multiple strands of evidence (e.g. aDNA, osteological assessment, radiocarbon dating, isotope analysis), has contributed to illuminate social practices within small and rural communities and uncover life histories of single individuals and family groups in Roman Britain: 1) the life history of long-distance mobility of a male individual in a isolated inhumation in Cambridgeshire, who did not carry traceable ancestry related to local populations in Britain and was instead genetically closer to Caucasus and Sarmatian groups; and 2) the reconstruction of small family trees spanning up to four generations in a small rural cemetery in Childrey Warren, Oxfordshire.

These two contrasting archaeological contexts â€“ isolated inhumation versus a formal cemetery â€“ highlight the diversity of burial traditions and provide insights on different aspects of life and death during this period, ultimately contributing for a more nuanced view of Roman Britain.

1124 - Development of a bioinformatic tool in Python for the Identification of vestigial octameric palindromic sequences in cyanobacterial genomes.

Alberto Balderas Martinez

Alberto Balderas Martinez
Universidad Autonoma del Estado de Mexico

This study introduces a cutting-edge bioinformatic tool developed in Python, specifically tailored to identify vestigial octameric palindromic sequences residing within cyanobacterial genomes. Through the utilization of binary encoding and Boolean algebra, our tool optimizes algorithms, significantly reducing computational overhead and memory usage. By implementing these innovative strategies, our tool ensures efficient and streamlined sequence analysis, thereby enhancing the speed and precision of palindromic sequence identification in cyanobacterial genomes. Additionally, our tool facilitates easy traversal of various combinations and computation of Hamming distance by employing a binary DNA sequence code compatible with a tesseract system. This unique combination of computational and mathematical approaches offers a valuable resource for researchers investigating genetic elements and evolutionary processes in cyanobacteria.

170 - Estimating relatedness in structured and admixed populations using the distributions of Identity By Descent tract lengths

Arun Sethuraman

Arun Sethuraman
San Diego State University

Genetic relatedness estimators often utilize Identity By State (IBS) as a proxy for Identity By Descent (IBD), thereby leveraging observed allele frequency spectra in their inference. However, in highly structured or admixed populations, conflating IBS and IBD results in erroneous inference with large variance in relatedness estimates. Here I derive and describe a new computational method that utilizes the distribution of IBD tract lengths under the admixture model to estimate relatedness in structured or admixed populations. I describe a maximum likelihood-based estimator of relatedness, solved using an Expectation Maximization (EM) algorithm. I utilize extensive simulations of genome-scale haplotypic data under a variety of demographic models using stdpopsim, to test the accuracy and precision of our estimator. I also compare its performance against other estimators, including those implemented in InRelate, KING and IBSrelate.

Finally, I test my method to infer close first and second degree relatives in the 1000 Genomes Project data, while accounting for the presence of population structure and admixture.

223 - Leveraging the Ancestral Recombination Graph to estimate the STR mutation model

Sebastián Iturbe

SebastiÃ¡n Iturbe, Alan Izarraras-GÃ³mez, Rori Rohlfs, Carolina De Lima-Adam, Luis FernÃ¡ndez-Luna, Diego Ortega-Del Vecchyo
Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico, Department of Data Science and Institute of Ecology and Evolution, University of Oregon, Department of Data Science and Institute of Ecology and Evolution, University of Oregon, Department of Data Science and Institute of Ecology and Evolution, University of Oregon, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico

Short tandem repeats (STRs), also known as microsatellites, are repetitive base pair sequences that constitute approximately 3% of the human genome. Over 10,000 STR variants contribute to gene expression, accounting for 10-15% of the heritability of gene expression highlighting their significance in various clinical conditions. The evolution and mutation dynamics of STRs are a crucial factor we need to understand in order to perform more comprehensive analysis of the phenotypic impact of STRs. To achieve this understanding, it is imperative to infer the mutation model determining patterns of genetic variation of STRs. The evolution of STRs relies on the overarching evolutionary history across the genome. This history is manifested in the genomic relationships among all samples represented in the Ancestral Recombination Graph (ARG). Recent advancements in methods for inferring the ARG, such as Relate and tsinfer, present opportunities to devise strategies for inferring the most plausible STR mutation model integrating the evolutionary history encoded in the ARG. Our proposal involves the development of a novel maximum-likelihood method for inferring the most probable mutation model by leveraging information encoded in the ARG. Through exhaustive simulations, we demonstrate the accuracy of our method in inferring the STR mutation model. Additionally, we show an application of our method using data from the 1000 Genomes Project.

363 - HaploSweep: detecting and distinguishing recent soft and hard selective sweeps through haplotype structure

Shilei Zhao

Shilei Zhao, Lianjiang Chi, Hua Chen
Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing Institute of Genomics, Chinese Academy of Sciences

Identifying soft selective sweeps using genomic data is a challenging yet crucial task in population genetics. In this study, we present HaploSweep, a novel method for detecting and categorizing soft and hard selective sweeps based on haplotype structure. Through simulations covering a broad range of selection intensities, softness levels, and demographic histories, we demonstrate that HaploSweep outperforms other methods in detecting soft sweeps. Specifically, HaploSweep achieves high classification accuracies of 0.9247, 0.9484, and 0.9829 for the CHB, CEU, and YRI populations respectively, when applied to simulations under the human Out-of-Africa demographic model. We further propose a method to accurately distinguish soft shoulders around hard sweeps from soft sweeps. Applying this method to genomic data from three populations (CHB, CEU, and YRI) of the one thousand genomes project, we identify multiple novel genes exhibiting strong population-specific soft sweep signals, including HRNR, AMBRA1, CBFA2T2, DYNC2H1, and RANBP2 etc, showing an enrichment of immunity and metabolism-related terms. By demonstrating the superior performance in both simulated and real data, HaploSweep underscores its potential as a valuable tool for investigating soft sweeps and providing insights into the mechanisms of

adaptive evolution.

400 - Assessing the impact of post-mortem damage and contamination on imputation performance in ancient DNA

Antonio Garrido Marques

Antonio Garrido Marques, Simone Rubinacci, Anna-Sapfo Malaspinas, Olivier Delaneau, BÃ¡rbara Sousa da Mota
Department of Computational Biology, University of Lausanne, Division of Genetics, Department of Medicine, Brigham and Womenâ€™s Hospital and Harvard Medical School, Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Department of Computational Biology, University of Lausanne, Swiss Institute of Bioinformatics, University of Lausanne, Regeneron Genetics Center, Department of Computational Biology, University of Lausanne, Swiss Institute of Bioinformatics, University of Lausanne

Low-coverage imputation is becoming ever more present in ancient DNA (aDNA) studies. Imputation pipelines commonly used for present-day genomes have been shown to yield accurate results when applied to ancient genomes. However, post-mortem damage (PMD), in the form of C-to-T substitutions at the reads termini, and contamination with DNA from closely related species can potentially affect imputation performance in aDNA. In this study, we evaluated imputation performance i) when using a genotype caller designed for aDNA, ATLAS, compared to bcftools, and ii) when contamination is present. We evaluated imputation performance with principal component analyses (PCA) and by calculating imputation error rates. With a particular focus on differently imputed sites, we found that using ATLAS prior to imputation substantially improved imputed genotypes for a very damaged ancient genome (42% PMD). Trimming the ends of the sequencing reads led to similar improvements in imputation accuracy. For the remaining genomes, ATLAS brought limited gains. Finally, to examine the effect of contamination on imputation, we added various amounts of reads from two present-day genomes to a previously downsampled high-coverage ancient genome. We observed that imputation accuracy drastically decreased for contamination rates above 5%. In conclusion, we recommend i) accounting for PMD by either trimming sequencing reads or using a genotype caller such as ATLAS before imputing highly damaged genomes and ii) only imputing genomes containing up to 5% of contamination.

588 - Distinguishing between 2nd-degree relationships in ancient pedigrees using IBD detection

Théo Cavinato

ThÃ©o Cavinato, Simone Rubinacci, Olivier Delaneau, Anna-Sapfo Malaspinas
Department of Computational Biology, University of Lausanne, Swiss Institute of Bioinformatics, University of Lausanne, Division of Genetics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Regeneron Genetics Center, Department of Computational Biology, University of Lausanne, Swiss Institute of Bioinformatics, University of Lausanne

Kinship inference between individuals from an archeological site allows pedigree reconstruction and improves our knowledge on past social structures. However, reconstructing a pedigree can be challenging: kinship inference methods often only return the degree of relatedness between two individuals, which in turn can be explained by multiple underlying genealogical structures. For instance, an uncle - niece pair or a half-sibling pair would have the same degree of relatedness, but distinct pedigrees. Recent methodological improvements in ancient DNA now allow the detection of Identical-By-Descent (IBD) segments between two ancient relatives. In this work, we developed a method that leverages the analysis of IBD segments to solve complex genealogical structures. Briefly, in cases where siblings show a 2nd-degree relatedness with a third individual, our method first infer shared IBD between the siblings and that individual and then determine whether this person is their uncle/aunt or not based on Mendelian principle. The method also takes into account the rest of the pedigree to distinguish between other types of 2nd-degree relatedness when possible. We benchmarked our approach on simulated

genomes and our results show a strong precision even for extremely low coverages. Finally, we applied it to the case study of the Koszyce archeological site to correctly place a young deceased child in a reconstructed genealogy.

833 - The effect of divergent and parallel selection on the genomic landscape of divergence

Hisham Ali

Hisham AA Ali, Tim Coulson, Sonya M Clegg, Claudio S Quilodran

Arizona State University, University of Oxford, University of Oxford, University of Oxford, -, University of Geneva, Geneva, Switzerland.

While the role of selection in divergence along the speciation continuum is theoretically well understood, defining specific signatures of selection in the genomic landscape of divergence is empirically challenging. Modelling approaches can provide insight into the potential role of selection on the emergence of a heterogenous genomic landscape of divergence. Here, we extend and apply an individual-based approach that simulates the phenotypic and genotypic distributions of two populations under a variety of selection regimes, genotype-phenotype maps, modes of migration, and genotype-environment interactions. We show that genomic islands of high differentiation and genomic valleys of similarity may respectively form under divergent and parallel selection between populations. For both types of between-population selection, negative and positive frequency-dependent selection within populations generated genomic islands of higher magnitude and genomic valleys of similarity, respectively. Divergence rates decreased under strong dominance with divergent selection, as well as in models including genotype-environment interactions under parallel selection. For both divergent and parallel selection models, divergence rate was higher under an intermittent migration regime between populations, in contrast to a constant level of migration across generations, despite an equal number of total migrants. We highlight that interpreting a particular evolutionary history from an observed genomic pattern must be done cautiously, as similar patterns may be obtained from different combinations of evolutionary processes. Modelling approaches such as ours provide an opportunity to narrow the potential routes that generate the genomic patterns of specific evolutionary histories.

955 - Single-Haplotype Genome Assemblies Reveal the Enigmatic Evolution of Felid Genome Architecture

Andrew J Harris

Andrew J Harris, William J Murphy

Texas A&M University, Texas A&M University

Over the past decade, the ability to assemble high-quality genome assemblies and parse haplotype-specific differences has enabled close evaluation of genome architecture evolution between individuals, populations, and species. Advancements in long-read sequencing accuracy and genome assembly algorithms revolutionized the ability to parse parental haplotypes within trios to directly interrogate haplotype-specific evolutionary questions. These questions often target key aspects of adaptation, selection, and speciation alike. Here, we report how using trio-binned F1-interspecific hybrids for single-haplotype genome assembly enabled near Telomere-to-Telomere genomes that provide novel insights into the evolution of genomic architecture spanning the family Felidae. We find an enrichment of inversions as well as large structural differences between species on the X chromosome, within a region previously associated with hybrid dysfunction that acts as a barrier to gene flow. For the first time, we evaluate the size and architecture of multi-megabase transcriptionally active FA-SAT macrosatellites in domestic cats, which are implicated in the formation of cancer and structure of centromeres. Finally, we provide further insight into the evolution of the large, copy-number variable macrosatellite DXZ4 which has been shown to have a role in speciation and hybrid sterility. Together, we present an intriguing view into the structurally dynamic evolution of felid genomes across roughly 15 million years. With continued additions of high-quality single haplotype genome assemblies, we aim to

fully characterize the repertoire of structural differences between and within all cat species to facilitate genotype to phenotype discovery.

961 - Improving Local Ancestry Inference through Neural Networks

Jazeps Medina Tretmanis

Jazeps Medina Tretmanis, Maria C Avila-Arcos, Flora Jay, Emilia Huerta-Sanchez
Center for Computational Molecular Biology, Brown University, International Laboratory for Human Genome Research, UNAM, CNRS, INRIA, Laboratoire Interdisciplinaire des Sciences du Numerique, Universite Paris-Saclay, Center for Computational Molecular Biology, Brown University, Department of Ecology, Evolution and Organismal Biology, Brown University

Local Ancestry Inference (LAI) is the process of classifying the genetic ancestry of chromosomal segments within a genome. LAI enables the study of ancestry-specific disease risk, as well as the evolution of admixed populations.

Historically, LAI methods have been implemented through Hidden Markov Models (HMMs) underpinned by Li & Stephenâ€™s model. However, recent methods have started to leverage Neural Networks (NNs) for the task of LAI. We present the first comprehensive benchmarking of both traditional and NN models for LAI using a combination of real and simulated data. We also benchmark the performance of NN architectures that have not previously been used for LAI. We include model performance for scenarios that are not usually benchmarked: short ancestry tracts (<1000 SNPs), and admixture between closely related populations.

We find that Convolutional NNs achieve the highest performance out of all methods for traditional LAI. While we can reliably infer inter-continental genetic ancestry within admixed individuals, current methods cannot reliably infer intra-continental ancestries. However, we find that this can be achieved by supplementing SNP data with other statistics such as minor allele frequencies. Finally, we find that short ancestry tracts still pose a challenge for LAI, but that this can be ameliorated by incorporating knowledge of recombination rates into model architectures.

The work presented here is useful both for its value in helping researchers make informed decisions regarding which methods fit their needs, as well as being an important initial data point for how NNs can improve the LAI method landscape.

85 - Simple, general tests for accelerated evolution and positive selection

Alexander L Starr

Alexander L Starr, Leslie Magtanong, Gabriella Cale, Alexander Hoefler, Hunter B Fraser
Stanford University, Stanford University, Stanford University, Stanford University, Stanford University

A large body of work has established that the genetic basis of human complex traits primarily lies in many non-coding variants with small phenotypic effects. However, widely used methods to detect ancient changes in the strength of natural selection are primarily applicable only to small genomic regions or protein-coding sequences, creating a mismatch between what is known about human genetics and the methods used to understand human evolution. To address this, we developed two simple, general methods to detect mono- and polygenic (1) accelerated evolution of and (2) positive selection on coding and non-coding sequence genome-wide. Using the first method, we identified hundreds of genes that have undergone accelerated evolution on coding and nearby non-coding sequences in the human lineage. We highlight the gene with the strongest evidence for cis-regulatory accelerated evolution, LMO4, which plays a key role in determining handedness and brain asymmetry. Although accelerated evolution can be the result of positive selection, it can also result from an increased local mutation rate or loss of evolutionary constraint. Therefore, we developed a new test for positive selection that is

applicable genome-wide. Using this test, we find evidence for positive selection at non-coding sites near dozens of genes, including at the EVC/EVC2 locus previously linked to human-specific craniofacial morphology and many transcriptional regulators. Overall, we anticipate that our broadly applicable frameworks will improve our understanding of how natural selection has shaped the evolution of humans and other species.

112 - Diverse patterns of transposable elements expressions across tissues exhibited by Rhesus macaque and possible regulation of gene expression of adjacent genes by tissue-specific TEs

Dae-Soo Kim

Kun-Hyang Park, Nakyoung Kim, Dae-Soo Kim

Korea Research Institute of Bioscience & Biotechnology, Department of Core Facility Management Center, Korea Research Institute of Bioscience & Biotechnology, Korea Research Institute of Bioscience & Biotechnology

Transposable elements (TEs) constitute approximately half of primate genomes and play a pivotal role in gene regulation, potentially contributing to the evolution of phenotypic diversity. In this study, we examined TE expression patterns across 13 distinct tissue types in the rhesus macaque, a widely used non-human primate model, using publicly available RNA-seq and Cap analysis gene expression datasets. Our analysis revealed non-random distribution and differential enrichment of TEs across genomic regions, with intergenic TEs being predominantly silenced and highly expressed TEs notably concentrated in the thymus.

We identified four distinct clusters of TEs exhibiting tissue-specific enrichment patterns, with enrichment levels varying by tissue type and TE age. Particularly noteworthy were the 365 short interspersed nuclear elements (SINEs) and 269 long interspersed nuclear elements (LINEs) highly enriched near transcription start sites (TSSs), suggesting potential regulatory roles. Furthermore, we uncovered 37 tissue-specific TEs highly expressed within coding DNA sequences (CDS) and 1226 upstream of genes, potentially driving tissue-specific gene overexpression.

By examining pairwise comparisons and associations between TEs and genes, we elucidated the role of tissue-specific TEs in regulating nearby gene expression in primates. Our findings contribute to a deeper understanding of TEs as regulators of gene expression in a tissue-specific context, shedding light on the intricate mechanisms underlying phenotypic diversity in primates. This study enhances our comprehension of the complex interplay between TEs and gene regulation, offering valuable insights into primate genome evolution and functional genomics.

122 - Uncovering the regulatory function of variants that shaped human evolution

David Gokhman

Carly Weiss, Lana Harshman, Katharina Lange, Ryder Easterlin, Yizhi Yan, Hunter Fraser, Dmitri Petrov, Simon Fishilevich, Fumitaka Inoue, Nadav Ahituv, David Gokhman

Stanford University, UCSF, Weizmann Institute of Science, UCSF, ASHBi, Stanford University, Stanford University, Weizmann Institute of Science, ASHBi, UCSF, Weizmann Institute of Science

Changes to gene regulation are key drivers of human evolution. However, our ability to predict the effects of sequence variants on expression is limited. Consequently, the effects of most human variants remain unknown. Here, we used massively parallel reporter assays to experimentally determine the regulatory effect of each of the hundreds of thousands of single-nucleotide variants that emerged and spread to fixation in human evolution. Using a range of skeletal, adipose, skin, and neural cells, we compared the functional output of the ancestral and the derived allele for each of the 71,443 variants distinguishing modern humans from Neanderthals and Denisovans, as well as the 541,851 variants distinguishing all human lineages from other great apes. We found that 15% of the variants arose in active

regulatory elements, and that 20% of these altered human expression levels. To identify genes affected by expression-altering variants, we generated human-chimpanzee hybrid cells, where expression differences between the species are driven solely in cis (i.e., by nearby variants). We found that genes controlling the vocal tract, face, and cerebellum were particularly enriched in regulatory changes, suggesting that these systems might have experienced unique selective pressures. Finally, we identified specific variants that likely contributed to human-specific traits, including in SATB2, a key skeletal and neural regulator, and in IRF4, a regulator of skin pigmentation. Overall, this work uncovers functionally important variants in deep and recent human evolution, and elucidates their role in shaping human adaptations.

131 - Unraveling the Genomic Influence of Archaic Hominins on 3D Genome Interactions, Immune Pathways, and Gene Expression in Modern Humans

Isabela Alvim

Isabela Alvim, Matthew Leavesley, FranÃ§ois-Xavier Ricaut, Nicolas Brucato, Murray Cox, Irene Gallego Romero

Massey University, St. Vincentâ€™s Institute, School of Humanities and Social Sciences, , University of Papua New Guinea, Centre de Recherche sur la Biodiversi et lâ€™Environnement, Universite de Toulouse, Centre de Recherche sur la Biodiversi et lâ€™Environnement, Universite de Toulouse, University od Aucklans, St. Vincentâ€™s Institute

Modern humans have inherited as much as 8% of their genome from extinct hominins, with the highest frequencies of Neanderthal and Denisovan DNA today found in the Pacific region. This â€˜archaicâ€™ DNA is enriched for immune pathway genes and regulatory elements, and clusters non-randomly across the genome in a pattern characteristic of the co-associating domains that form during 3D folding of DNA in the nucleus. Yet how archaic DNA interacts in 3D space and thus influences gene expression remains poorly understood.

To address this knowledge gap, we employ a multi-faceted approach: (i) we identify archaic DNA regions in the genomes of 12 individuals from Papuan New Guinea; (ii) subsequently, we empirically determined how archaic DNA aligns in the three-dimensional space of living immune cells from these individuals through Hi-C experiments; (iii) we then explore how this 3D structure influences DNA accessibility and gene expression by integrating ATAC-seq and RNAseq data. Finally, (iv) we compare the results to data from the same experiments performed with 6 African samples, which have no archaic introgression, and 6 European samples, which show lower levels of archaic introgression coming mostly from Neanderthals. This study contributes to a deeper understanding of our genetic inheritance from archaic hominins by shedding light on the complex interplay between archaic DNA, three-dimensional genomic organization, and gene regulation in immune cells

165 - Inference of the distribution of fitness effects using local genealogies

Alan Izarraras-Gomez

Alan Izarraras-Gomez, Christian Huber, Charleston W.K. Chiang, Diego Ortega-Del Vecchyo

International Laboratory for Human Genome Research, National Autonomous University of Mexico, Department of Biology, Pennsylvania State University, State College., Center for Genetic Epidemiology, Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Department of Quantitative and Computational Biology, University of Southern California, International Laboratory for Human Genome Research, National Autonomous University of Mexico

Mutations in the human genome have deleterious, neutral or advantageous effects, which are quantified with a selection coefficient (s). The distribution of fitness effects (DFE) measures the probability of new mutations having a certain selection coefficient and, as such, is central to fundamental questions in population genetics such as how natural selection influences genetic diversity. The DFE has

previously been inferred for a variety of species using the site frequency spectrum (SFS), measures of haplotype lengths as well as different groupings of summary statistics. A drawback of these approaches is that they focus on specific summary statistics of the genetic data which are not guaranteed to be sufficient to estimate the impact of natural selection. Notably, all patterns of genetic variation explained by summary statistics are encoded on the Ancestral Recombination Graph (ARG), which is a data structure that shows the structure of local genealogies across the genome. Natural selection changes coalescent patterns from local genealogies making the ARG a rich data source to infer the DFE. In this study we leverage information from the ARG to build a maximum likelihood approach to infer the DFE using information from the number of lineages with a derived allele at specific time points. We show through simulation that this approach can accurately infer a wide range of selection coefficients (0.1>2Ns<1000) and show an application of the method to infer the DFE of non-synonymous mutations of the Yoruba population from the 1,000 genomes project.

181 - Curbing the accumulation of deleterious mutations: the roles of weak epistasis and compensatory beneficial mutations

Ulises Hernández Martín del Campo

Ulises HernÃ¡ndez, Joseph Matheson, Walid Mawass, Joanna Masel
University of Arizona, University of Arizona, University of California San Diego, University of Arizona, University of Arizona

Under realistic human genome-wide deleterious mutation rates Ud>1, new deleterious mutations appear faster than selection can purge them. This can cause fitness decline and eventual population extinction. Using a novel simulation framework that efficiently handles genome-wide linkage disequilibria across many segregating sites, we found that large effect beneficial mutations can compensate for the accumulation of many smaller-effect deleterious mutations. Indeed, populations might even adapt faster following increased mutation rate. However, in the absence of epistasis, the number of beneficial mutations required can exceed Haldaneâ€™s â€œcost of selectionâ€• limit created by finite reproductive excess. When mutations have larger fitness effects in individuals of lower fitness (global negative epistasis), the variance in fitness increases as the population degrades, increasing the rate at which selection purges mutations, sometimes making the rate of further deleterious fixations negligible. We combined expressions from the Price equation with expressions from a one locus two allele finite Markov chain model to derive a novel solution for mutation-selection-drift metastable balance under high Ud and global epistasis. Very weak negative epistasis is sufficient for population metastability in large but not in small populations. Indeed, weak epistasis combined with rare large effect beneficial mutations allow population persistence even at Ud=10. Given humansâ€™ high Ud>2, conventional models suggest that we are on the edge of continual genome degradation - our results suggest that under weak global epistasis and beneficial substitutions, human populations might even tolerate further increases in mutation rate.

245 - Estimating gene conversion tract length and rate from PacBio HiFi data

Anders Poulsen Charmouh

Anders Poulsen Charmouh, Peter Porsborg SÃ¸rud, Thomas Bataillon, Asger Hobolth, Lasse Thorup Hansen, SÃ¸ren Besenbacher, Mikkel Heide Schierup
Bioinformatics Research Centre, Aarhus University., Bioinformatics Research Centre, Aarhus University., Bioinformatics Research Centre, Aarhus University., Department of Mathematics, Aarhus University., Department of Mathematics, Aarhus University., Department of Mathematics, Aarhus University., Department of Molecular Medicine (MOMA), Aarhus University, Bioinformatics Research Centre, Aarhus University.

A non-crossover event (NCO) occurs when the DNA double-strand break (DSB) is repaired using the homologous chromosome as template through the synthesis-dependent strand annealing pathway, or double Holliday junction pathway (Chen et al. 2007; McMahill et

al. 2007). Some NCOs are observable as gene conversions which occur when one SNP is transferred unidirectionally from one sequence to another (Mansai et al. 2011; Lorenz and Mpaulo 2022). NCO is an understudied recombination mechanism despite most DSBs yielding NCO rather than a crossover (Cole et al. 2012; Halldorsson et al. 2019). Further, gene conversions have several effects on genomic diversity and hence many evolutionary implications as they can generate new combinations of alleles, cause de novo mutations due to the potentially mutagenic effect of recombination and even counteract the mutation load by reverting germline mutations through GC-biased gene conversion (Arbeithuber et al. 2015; Halldorsson et al. 2019). Here, we show that gene conversion events can be called directly through HiFi PacBio sequencing data, without using trio data or breeding experiment. We develop a new model of gene conversion rate and tract length and show that very few percent of all NCOs are observable as gene conversions. We estimate the rate of NCO for the genome of humans and other great apes. The results are relevant with respect to understanding NCOs and the extent to which they affect the genome.

253 - Setting up an integrated pipeline of analyses to explore the genetic architecture of complex adaptive traits evolved by modern human populations

Giulia Ferraretti

Giulia Ferraretti, Paolo Abondio, Marta Alberti, Agnese Dezi, Phurba T. Sherpa, Paolo Cocco, Massimiliano Tiriticco, Marco di Marcello, Guido Alberto Gnecchi-Ruscone, Luca Natali, Angela Corcelli, Giorgio Mrinelli, Davide Peluzzi, Stefania Sarno, Marco Sazzini
Laboratory of Molecular Anthropology & Centre for Genome Biology, Dept. of Biological, Geological and Environmental Sciences, University of Bologna, Italy, Dept. of Cultural Heritage, Ravenna Campus, University of Bologna, Italy, Laboratory of Molecular Anthropology & Centre for Genome Biology, Dept. of Biological, Geological and Environmental Sciences, University of Bologna, Italy, Dept. of Emergency and Organ Transplantation, University of Bari Aldo Moro, Italy, Mount Everest Summiters Club, Rolwaling, Dolakha, Nepal, Explora Nunaat International, Montorio al Vomano, Teramo, Italy, Explora Nunaat International, Montorio al Vomano, Teramo, Italy, Explora Nunaat International, Montorio al Vomano, Teramo, Italy, Dept. of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, Italian Institute of Human Paleontology, Rome, Italy, Dept. of Basic Medical Science, Neuroscience and Sense Organs, University of Bari Aldo Moro, Italy, Explora Nunaat International, Montorio al Vomano, Teramo, Italy, Explora Nunaat International, Montorio al Vomano, Teramo, Italy, Laboratory of Molecular Anthropology & Centre for Genome Biology, Dept. of Biological, Geological and Environmental Sciences, University of Bologna, Italy, Laboratory of Molecular Anthropology & Centre for Genome Biology, Dept. of Biological, Geological and Environmental Sciences, University of Bologna, Italy, Interdepartmental Centre Alma Mater Research Institute on Global Changes and Climate Change, University of Bologna, Italy

Biological adaptations mediated by complex/polygenic traits are increasingly supposed to represent most of the adaptive responses evolved by human populations to cope with a range of selective pressures. However, studies in such a research field still lack specifically defined methodologies able to formally test a realistic approximation of the polygenic model of adaptation, with many approaches relying on results from Genome-Wide Association Studies, which are biased by well-known conceptual limitations.
Here, we combined three different methods to set up an integrated pipeline of analyses capable to detect functionally related loci putatively implicated in the modulation of complex/polygenic adaptive traits. Particularly, we tested the reliability of such an approach by investigating the genetic bases of high-altitude adaptations evolved by Himalayan populations, which provide a case-study characterized by a well-understood selective pressure. For this purpose, whole genome data from individuals of Tibetan ancestry were first scanned with a likelihood haplotype-based method that discriminates between strong and weak selective signals. Distributions of the latter were then used to identify networks of genes significantly enriched in such signatures and strictly related from a functional perspective. Finally, we shortlisted the obtained results by focusing on those putative adaptive genes confirmed also by a supervised-machine learning algorithm trained on simulated data build on demographic parameters inferred for Tibetan populations. Accordingly, we pinpointed

angiogenesis-related pathways as processes pervasively shaped by the action of natural selection in Tibetans and potentially related to their improved tissue blood perfusion, as previously proposed.

351 - Inference of selection acting on coding sequences in the human lineage.

Hossameldin Loay

Hossameldin Loay, Donate Weghorn

Center of Genomic regulation CRG, Universitat Pompeu Fabra UPF, Center of Genomic regulation CRG

Signs of selection within DNA sequences are usually detected using molecular clock models. However, these methods depend on certain intrinsic assumptions. First, they assume the same mutation rate across different genomic compartments and for different mutation types. Second, they often assume infinite sites: identical by state mutations are assumed to be identical by descent (no parallel mutations), and a mutated site is not allowed to re-mutate to a different state (no multiple hits). Sequencing more extant individuals of the human population shows that the infinite sites assumptions are violated, where more individuals harbor identical by state mutations from independent mutation events.

Under the infinite sites assumptions, substitution rates (fixation rates) are equal to the underlying mutation rates.A difference between the two values is dependent on the rate of acquiring new mutations: population size multiplied by mutation rate. Here, we show that we can leverage the mutation rate variability across different genomic regions as well as different mutation types to have an estimate for an effective population size that can be used to correct for the bias induced by parallel mutations. Moreover, we correct for multiple hits without assuming detailed balance often assumed by some models like the HKY model. We use the corrected mutation rate values and the estimated effective population size to calculate gene-specific selection coefficients. We are also able to retrieve this information for specific historic evolutionary episodes that are not accessible by traditional methods.

365 - Global patterns of Holocene natural selection

Laura L Colbran

Laura L Colbran, Iain Mathieson

University of Pennsylvania, University of Pennsylvania

Ancient DNA (aDNA) provides direct access to the evolutionary history of individual variants and has been used extensively to detect and classify occurrences of natural selection. These efforts have largely focused on European populations. Populations in other parts of the world share many selective pressures with Europeans, but have also been exposed to independent environments. The extent to which the response to Holocene selection is shared across populations is unknown. New aDNA from different parts of the world provides an opportunity to study the common effects of shared selective pressures and to identify novel selected loci and traits.

In this study, we tested for selection in populations in Europe, East Asia, Africa and Central/South America. We modelled admixture proportions in 6,567 individuals from 12 ancient populations, and 1,228 individuals from 14 modern populations and tested for significant deviations using a maximum likelihood approach. In addition to known loci such as LCT and ADH1B, we identified novel loci such as DDB2 in Europe, Asia and the Americas and PDK4 in the Americas. We also adapted this approach to test for polygenic selection on complex traits and found that balancing selection on waist-hip ratio and directional selection on skin pigmentation are found on every continent. However other signals of directional selection on viral immunity in Eurasia, white blood cell counts in Africa and bone mineral density in Europe are more region-specific. This study represents the first comprehensive aDNA-based comparison of selective pressures

across much of the inhabited world.

407 - 50,000 years of Evolutionary History of India: Insights from ~2,700 Whole Genome Sequences

Elise Kerdoncuff

Elise Kerdoncuff, Laurits Skov, Nick Patterson, Wei Zhao, Yuk Yee Lueng, Gerard D Schellenberg, Jennifer A Smith, Sharmistha Dey, Andrea Ganna, AB Dey, Jinkook Lee, Priya Moorjani

Department of Molecular and Cell Biology, University of California, Department of Molecular and Cell Biology, University of California, Department of Human Evolutionary Biology, Harvard University, Survey Research Center, Institute for Social Research, University of Michigan, Department of Epidemiology, School of Public Health, University of Michigan, Neurodegeneration Genomics Center, Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Neurodegeneration Genomics Center, Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Department of Epidemiology, School of Public Health, University of Michigan, Survey Research Center, Institute for Social Research, University of Michigan, Department of Biophysics, All India Institute of Medical Sciences, Institute for Molecular Medicine Finland, Department of Geriatric Medicine, All India Institute of Medical Sciences, Center for Economic & Social Research, University of Southern California, Department of Molecular and Cell Biology, University of California, Center for Computational Biology, University of California

Indians have been underrepresented in human genomic studies. We generated ~2,700 high coverage genomes from Indiaâ€“â€“including individuals from most geographic regions, speakers of all major languages, from tribal and caste groupsâ€“â€“providing a comprehensive survey of genetic variation in India. We show that most Indians derive ancestry from ancient Iranian farmers, Eurasian Steppe pastoralists and South Asian hunter-gatherers. We uncover a common source of Iranian-related ancestry from early Neolithic cultures of Central Asia across diverse Indian groups, including Ancestral South Indians, Ancestral North Indians, Austroasiatic-related and East Asian-related individuals, suggesting that the Iranian-related gene flow occurred well before the Steppe pastoralist-related gene flow. Following these admixtures, India experienced a major demographic shift towards endogamy, resulting in extensive homozygosity and identity-by-descent sharing among individuals. At deep timescales, Indians derive ~1-3% ancestry from Neanderthals and Denisovans. By assembling the surviving fragments of archaic ancestry in Indians, we recover ~1.5Gb of the introgressing Neanderthal and ~0.6Gb of the introgressing Denisovan genomes, more than any previous archaic ancestry study. Indians have the largest variation in Neanderthal ancestry and the highest amount of population-specific Neanderthal segments among worldwide groups. Finally, we demonstrate that most of the genetic variation in Indians stems from a single major migration out of Africa that occurred around 50,000 years ago, with minimal contribution from earlier migration waves. Together, these analyses provide a detailed view of Indian evolutionary history and underscore the value of expanding genomic surveys to diverse groups outside Europe.

441 - Ancient AMY1 gene duplications primed the amylase locus for adaptive evolution upon the onset of agriculture

Charikleia Karageorgiou

Charikleia Karageorgiou, Feyza YÄ±lmaz, Kwondo Kim, Petar Pajic, Kendra Scheer, Christine Beck, Ann-Marie Torregrossa, Charles Lee, Omer Gokcumen

University at Buffalo, The Jackson Laboratory for Genomic Medicine, The Jackson Laboratory for Genomic Medicine, University at Buffalo, University at Buffalo, The Jackson Laboratory for Genomic Medicine, University of Connecticut, The University of Connecticut Health Center, University at Buffalo, The Jackson Laboratory for Genomic Medicine, University at Buffalo

Starch digestion is a cornerstone of nutrition, with the amylase enzyme playing a crucial role in starch metabolism. The copy number of

the human amylase gene has been linked to both metabolic diseases and the adaptation to agricultural diets. Prior studies have suggested recent origins of salivary amylase gene duplications. Here, we characterized 51 distinct amylase haplotypes across 98 individuals employing long-read DNA sequencing and optical mapping methods. Additionally, analysis of 31-mers linked to individual haplotypic blocks in the locus suggests that the first duplication of the amylase locus occurred more than 700,000 years ago, predating the human-Neanderthal split. Further, our analysis of ancient genomes revealed a shift towards higher amylase copy numbers in Neolithic farmers than in hunter-gatherer groups. The comparison of amylase haplotypes allowed us to characterize multiple nonallelic homologous recombination (NAHR) and microhomology-mediated break-induced replication (MMBIR) events, underlying different types of structural variants in the locus. Overall, our study reveals the mutational mechanisms and evolutionary forces that shape the remarkable structural complexity of the amylase locus.

442 - Genetic adaptations and demographic history in East Asia during the period of rice domestication

Mariko Isshiki

Mariko Isshiki, Srilakshmi M Raj

Albert Einstein College of Medicine, The University of Tokyo, Albert Einstein College of Medicine

Rice is a staple crop in East Asia, and its domestication created one of the largest dietary and environmental changes in East Asia in the last 10,000 years. However, except for a few examples ((e.g. ADH1B Arg47His), our knowledge about human adaptation during rice domestication is limited. We aimed to reveal detailed human demographic history and adaptation in East Asia during the last 10,000 years using whole genome sequence data of ~7000 individuals of East Asian ancestry. We found substantial effective migration in East Asia mirroring major geographic barriers, using FEEMS. A population size increase after rice domestication was observed in some populations through IBDNe analyses. We also identified variants that may have undergone selection in East Asia during different time periods meaningful to rice domestication, using the methods Relate, ASMC and SDS. The variants selected during the last ~10,000 years and ~5,000 years were enriched in genes associated with alcohol-related traits and plasma polyunsaturated fatty acid (PUFA) levels, respectively. PUFA levels are known to lower cardiovascular disease (CVD) risk, and higher rice consumption is known to be associated with CVD risk factors such as Type 2 diabetes (T2D). We conducted a PheWAS in All of Us Research Program that demonstrated lower risk of CVD in East Asians compared to Europeans, despite their higher risk of T2D. We identify new relationships between genetic adaptations and modern disease risk in East Asians that may be influenced by rice domestication.

445 - From Foraging to Farming: Tracing the Impact of Agricultural Adoption on Adaptation and Selection using whole-genome sequencing

Bridget Chak

Bridget Chak, Elise Lauterbur, Evan Wu, Maanasa Raghavan, David Enard, Luis Barreiro

University of Chicago, University of Arizona, University of Chicago, University of Chicago, University of Arizona, University of Chicago

The Neolithic revolution marked a pivotal transition in human history; from altering infectious disease burdens to narrowing dietary composition and diversity. Despite its significance, we lack a comprehensive understanding of how the emergence of agriculture has shaped the genetic histories of and adaptation among different human populations. Leveraging cross-continental comparisons, we investigated the extent to which convergent evolution of phenotypes and biological functions related to immunity and metabolism are observed among hunter-gatherer and agriculturalist populations facing similar ecological environments or evolutionary pressures.

Here, we collated a cross-continental dataset of newly generated and published high-coverage (~30x) whole-genome sequencing data from 10 hunter-gatherer, 2 pastoralist, and 21 agriculturalist populations from 6 geographic regions: Southeast Asia, South Asia, West Central Africa, East Central Africa, East Africa, and Southern Africa. This dataset includes a total of 974 individuals. We characterized a broad spectrum of genetic diversity and variation, including 62,908,806 SNPs and 15,319,830 indels. Using Flex-sweep, a machine learning-based method designed to detect diverse selective sweeps, we present evidence of population-private and -shared signals of positive selection among hunter-gatherer and agriculturalist populations, including significant adaptation in virus-interacting proteins. Overall, our study sheds light on how the process of agricultural adoption played out in varied global contexts with respect to the impact on human demographic and selection histories.

506 - Spatial Genomic Scale and Determinants of Human Germline Mutation Landscape

Nathan Cramer

Nathan Cramer, Priya Moorjani

University of California, Berkeley, University of California, Berkeley

Germline mutations are the ultimate source of genetic diversity within and across species. Understanding the rate and mechanisms by which mutations occur is of paramount importance for studies of medical genetics (to interpret the incidence of de novo and heritable diseases) and evolutionary biology (to date demographic and adaptive events). It is well established that mutation rates vary across the genome, from adjacent base pairs to whole chromosomes. However, the mechanisms underlying these patterns remain elusive. To understand the determinants of local mutation rate, we use low frequency single nucleotide polymorphisms from 76,156 whole genomes in the Genome Aggregation Database (gnomAD.) We focus on putatively neutral mutations by removing conserved regions and exons. We apply wavelet transform to characterize the spatial scale of variation in mutation patterns across the genome. Many factors correlate with local mutation rate variation, including replication timing, recombination rate, germline methylation, and base composition. We quantify the impact of each of these factors at various genomic size scales. By considering different types of mutations â€“ accounting for 6 strand-independent base substitution types and CpG transitions â€“ we explore the contribution of distinct mechanisms impacting the mutation spectrum. We then apply this approach to study mutation rate differences across sexes and populations. Together, this analysis helps quantify the contribution of various genomic features and biochemical processes impacting the spatial landscape of mutation rate in humans.

532 - A large pool of novel translated open reading frames is neutrally evolving in the human genome

Jiwon Lee

Jiwon Lee, Aaron Wacholder, Maria Chikina, Martin Jinye Zhang, Anne-Ruxandra Carvunis

University of Pittsburgh School of Medicine, Center for Evolutionary Biology and Medicine, Joint CMU-Pitt Ph.D. Program in Computational Biology, University of Pittsburgh School of Medicine, Center for Evolutionary Biology and Medicine, University of Pittsburgh School of Medicine, Carnegie Mellon University, University of Pittsburgh School of Medicine, Center for Evolutionary Biology and Medicine

Human annotated proteins are translated from sequences spanned by open reading frames (ORFs) that are largely conserved. In contrast, recent advances in ribosome profiling have uncovered a vast number of unannotated human noncanonical ORFs (nORFs) that are evolutionarily novel, existing mostly in the primate lineage. Because they are generally short (<110 amino acids) and taxonomically restricted, nORFs have been missed by current functional annotations. Thus, they are largely uncharacterized, and their prevalence in the human genome is an open question. In this work, we expand the known set of nORFs by integrating 1,067 human ribosome profiling experiments, and investigate their evolutionary roles and contribution to phenotypes. Our study utilizes SNP information from TOPMed

whole genome sequencing (WGS) data (N=132,345), and GWAS summary association statistics across 59 diseases and complex traits (avg. N=351,446). We identify 66,000 nORFs, significantly expanding upon nORFs reported in recent studies. We show that these nORFs have a nonsynonymous-to-synonymous SNP ratio close to 1 (adjusted pN/pS=0.976), suggesting they are evolving close to neutrally. By meta-analyzing LD score regression (LDSC) results across 59 nonredundant GWAS phenotypes with genetic correlation <0.1, we further find that nORFs are deficient in heritability, but provide independent information relative to the 52 functional annotations in the baseline-LD model (LDSC Ï„=-6.45Â±1.15Ã—10-7, p=1.91Ã—10âˆ’8). Though there is a growing number of case studies reporting nORFs with biologically important function, our results suggest that the bulk of human nORFs are evolving neutrally and have low contributions to complex traits and disease.

538 - Functional characterisation of the regulatory activity of archaic DNA in Island Southeast Asia

Maddy Comerford

Maddy Comerford, Davide Vespasiani, Navya Shukla, Matthew Leavesley, Christopher Kinipi, FranÃ§ois-Xavier Ricaut, Nicolas Brucato, Irene Gallego Romero

School of Biosciences, University of Melbourne, Melbourne Integrative Genomics, University of Melbourne, Human Genomics and Evolution, St. Vincentâ€™s Institute of Medical Research, School of Biosciences, University of Melbourne, Melbourne Integrative Genomics, University of Melbourne, Walter and Eliza Hall Institute of Medical Research, School of Biosciences, University of Melbourne, Melbourne Integrative Genomics, University of Melbourne, Human Genomics and Evolution, St. Vincentâ€™s Institute of Medical Research, School of Humanities and Social Sciences, University of Papua New Guinea, Health Services, University of Papua New Guinea, CNRS, UniversitÃ© de Toulouse, CNRS, UniversitÃ© de Toulouse, Human Genomics and Evolution, St. Vincentâ€™s Institute of Medical Research, Center for Genomics, Evolution and Medicine, University of Tartu, School of Biosciences, University of Melbourne, Melbourne Integrative Genomics, University of Melbourne

Interbreeding between modern humans and archaic hominins has contributed to the genomes of present-day human populations. Introgressed archaic variants are not uniformly distributed throughout the genome, but are enriched in regulatory regions, particularly in enhancers active in immune cell types, suggesting they have the potential to contribute to gene regulation and organism-level traits. However, due to the lack of diversity in genomic databases, research on the impacts of archaic introgression has focused mainly on the Neanderthal variants present in European populations, to the detriment of our understanding of the Neanderthal and Denisovan variants present in the rest of the world.

Using a Massively Parallel Reporter Assay (MPRA), we functionally characterise the regulatory activity of over 26,000 Neanderthal and Denisovan variants present in modern-day Papuan populations. To prioritise evolutionary and phenotypically relevant variants, we focused primarily on variants located within accessible, active chromatin in immune cell types and segregating at allele frequencies > 0.15, which together suggest they might have been targeted by positive selection. By comparing the regulatory activity of these archaic variants and their non-archaic counterparts in three cell lines of different genetic background, we assess the contribution of archaic introgression in shaping modern human genetic diversity and demonstrate the value of MPRAs as tools to study how gene regulatory changes have shaped phenotypes across evolution. Overall, this project will give insight into the regulatory activity of archaic variants in Island Southeast Asia and contribute to ongoing efforts to increase human diversity in genomics studies.

561 - Worldwide patterns of diversity at the 17q21.31 locus in modern human genomes

Samvardhini Sridharan

Samvardhini Sridharan, Peter James Sudmant
University of California, Berkeley, University of California, Berkeley

Among the known 150 large inversions segregating in human populations, one extraordinary example is at the 17q21.31 locus which spans >900 kilobases and comprises ten genes. These genes include the medically relevant MAPT and KANSL1. Several major inversion haplotypes of this locus segregate in modern human populations, each exhibiting distinct frequencies across geographic regions. Some of these haplotypes have been identified as being under selection and linked to fecundity. All of these haplotypes maintain a complex duplication architecture in addition to its inversion status. Consequently, this locus provides a unique system to study the impact of copy number variation on recombination suppression and extended linkage disequilibrium in human populations.

Here, we report extensive patterns of temporal and geographic variation at the 17q21.31 locus in several primates and in different modern human populations. We find that while the inverted haplotype (H2) is the ancestral haplotype in humans, it is the less common haplotype in modern-day human samples. We observe an exceptional amount of copy number diversity in direct orientation (H1) haplotypes globally, particularly in understudied modern populations including South Asians. We hypothesize that recombination suppression in large and structurally complex regions of the genome may contribute to the differential accumulation of deleterious mutations. Finally, we observe several unique instances of ancient double recombination events in African populations and date the timing of these events. Together, these results provide valuable insights into the role of chromosomal rearrangements in shaping evolution, diversity, disease susceptibility, and fitness.

577 - Unraveling Vadnagar's Genetic Heritage: Paleogenomic Insights

Bhavna Ahlawat

Bhavna Ahlawat, Lomous Kumar, Abhijit Ambekar, Jagmahender Singh Sehrawat, Niraj Rai
Birbal Sahni Institute of Paleosciences, Panjab University, Birbal Sahni Institute of Paleosciences, Achaeological Survey of India, Panjab University, Birbal Sahni Institute of Paleosciences, AcSIR

Vadnagar, located in Gujarat, India, is an important archaeological site with a rich history spanning five major dynasties. Positioned strategically at the junction of ancient trade routes, it holds cultural significance. The Vadnagar archaeological complex comprises Vadnagar, Taranga Hill, Anaj Godown, Kota Ambaji Lake, and Sharmistha Lake. This study delves into the genetic ancestry of Vadnagar, analyzing skeletal remains and archaeological evidence to discern migration patterns. Using advanced sequencing technology, the research aims to unveil insights into the genetic history of the site, shedding light on human history in the region. Ancient DNA analysis of skeletal remains (N=4) through PCA and ADMIXTURE reveals primarily local ancestry with possible population amalgamation. Samples from Kota Ambaji Lake and Sharmistha Lake suggest a local genetic heritage with signs of medieval migration. Relics from Taranga monastery suggest the emergence of Hinayana Buddhism around the 4th century CE, highlighting Vadnagar's cosmopolitan nature and its role as a cultural exchange hub. This study's findings contribute to understanding the complex history of Vadnagar, showcasing its dynamic past and diverse influences.

586 - Is there any population differentiation at particular SNPs in the human CYP1A2 gene?

Sayaka Chiku

Sayaka Chiku, Jun Gojobori, Yoko Satta
SOKENDAI, SOKENDAI, SOKENDAI

Human CYP1A2 is a detoxification enzyme that metabolizes xenobiotics such as drugs (caffein is a well-known example) and is expressed only in human liver. Although CYP1A2 expression and activity levels vary among individuals and populations, the evolutionary background of CYP1A2 polymorphism has not been explored. Therefore, the purpose of this study is to investigate whether certain CYP1A2 SNPs are positively selected. In this study, after LD pruning for quality control, the remaining 215 SNPs from the vcf of 1KGP GRCh37 were used for estimating their frequency and FST among 1KGP populations. The derived allele frequencies at these SNPs showed that they were divided into three characteristics: 136 SNPs occur in â€œone populationâ€•, 73 SNPs in â€œsome populationsâ€• include being present only in a particular superpopulation, and 6 in â€œall populationsâ€•. The estimated frequencies show the presence of population- or regional-specific SNPs, some of which are in LD with specific SNPs. Based on the estimated FST values, a total of 7 SNPs were in the top 2.5%. rs762551 and rs57295890 have high FST values between Peru and 9 and 24 populations, respectively. Although both SNPs exist in all populations, the observed high FST values due to unique SNP frequencies in Peru: rs762551 is 36.6% on average but 13.5% in Peru, whereas rs57295890 is 16.4% on average but 68% in Peru. We will examine what makes unique SNP frequencies in Peru by using statistical methods such as Tajimaâ€™s D, iHS and 2DSFS.

610 - Genetic adaptation of Asian human populations to their environment

Johanne Adam

Johanne Adam, Romain Laurent, Nina Marchi, RaphaÃ«lle Chaix, Evelyne Heyer, Laure SÃ©gurel
MusÃ©um National d'Histoire Naturelle - CNRS - UniversitÃ© Paris CitÃ©, MusÃ©um National d'Histoire Naturelle - CNRS - UniversitÃ© Paris CitÃ©, MusÃ©um National d'Histoire Naturelle - CNRS - UniversitÃ© Paris CitÃ©, MusÃ©um National d'Histoire Naturelle - CNRS - UniversitÃ© Paris CitÃ©, MusÃ©um National d'Histoire Naturelle - CNRS - UniversitÃ© Paris CitÃ©, Biometry and Evolutionary Biology Laboratory Intitute

Positive selection drives the spread of advantageous mutations in populations, contributing to phenotypic diversity across species. While many studies uncovered examples of human local adaptation, they mostly focused on Western European and Eastern Asian populations. This has left little knowledge about human genetic adaptation in other areas of Asia, such as Northern, Central and South-East Asia. Yet, there is an important cultural and environmental variability across these populations, and in humans, cultural practices and climatic constraints are known to result in local adaptation events.
To address this gap, we studied a large genomic dataset of 894 individuals from 27 diverse Asian populations. Our goal is to provide a comprehensive analysis of adaptive pressures across the entire Asian continent, further incorporating publicly available genomes from East Asia and South Asia. Using iHS statistics, we conducted genome-wide scans for selection and investigated the most prominent signals shared across the Asian continent or across sub-regions. We further specifically quantified the relative importance of climatic and dietary variability in driving local adaptation. Our analysis led to the identification of novel genomic regions under selection, associated with farming, herding, or hunting-gathering practices. Additionally, we explored the interaction between climate and subsistence strategies in influencing selective events. This approach allowed us to pinpoint specific regions that were potentially involved in diet adaptation across diverse climates. We assembled gene sets based on the functions inferred from adaptive events identified in our study and applied selection enrichment analyses to uncover new cases of polygenic selection.

659 - Exploring the Landscape of Gene Family Size and their Associated Pseudogenes in the Human Genome

Gabriela Procopio Leite

Gabriela Procopio, Beatriz Mello, Carlos Schrago

Federal University of Rio de Janeiro, Federal University of Rio de Janeiro, Federal University of Rio de Janeiro

Gene family dynamics play a crucial role in shaping the evolutionary history of species. Segmental duplication, a key driver of expansion, provides the raw material for the evolutionary process. Although often deemed as adaptative, these expansions should be carefully evaluated against the neutral duplication rate shaped by factors like recombination and adjacent mobile/repetitive elements composition. This implies certain genomic regions may be predisposed to duplication, independent of natural selection pressures. For instance, Nozawa et al. (2007) showed the number of pseudogene matches that of functional counterparts within the human olfactory receptor (OR) gene family. This equivalence suggests a similar probability of persistence and degeneration for novel copies in populations, casting doubt on the traditionally attributed selective advantage. Therefore, expansions may result from the interplay between duplication, purifying selection, and demography. Our study evaluates whether family sizes are positively correlated with the corresponding pseudogene number in the human genome, indicating their dynamics are mainly driven by fundamental molecular mechanisms associated with errors leading to segmental duplications. We investigate whether gene family expansions predominantly correlate with differential recombination rates and mobile elements composition. We found the number of pseudogenes is significantly positively correlated with paralog count. Moreover, this correlation is stronger for larger gene families. We map gene family size distribution and their associated pseudogenes across human chromosomes, juxtaposing this pattern against recombination rates and mobile element compositions. This comprehensive analysis offers an understanding of the processes shaping the genomic architecture of gene family size in humans.

716 - Gene surfing with background selection during human range expansions

Flávia Schlichta

FlÃ¡via Schlichta, Adamandia Kapopoulou, Stephan Peischl, Laurent Excoffier
Computational and Molecular Population Genetics lab, Institute of Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics, Computational and Molecular Population Genetics lab, Institute of Ecology and Evolution, University of Bern, Computational and Molecular Population Genetics lab, Institute of Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics, Computational and Molecular Population Genetics lab, Institute of Ecology and Evolution, University of Bern, Interfaculty Bioinformatics Unit, University of Bern, Computational and Molecular Population Genetics lab, Institute of Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics

Range expansions have been ubiquitous in human history, and they have played a significant role in shaping human genomic diversity. We use genomes from eight different human populations located at different distances from Africa to identify gene surfing events as local dips of diversity (troughs), and we study their distribution across the genome and their impact on functional regions. Populations closer to Africa have relatively more troughs in functional than in neutral regions, while those further from Africa (e.g. Karitiana) show similar levels of trough density in both regions. While an excess of such diversity dips in functional regions could be interpreted as signatures of positive selection and selective sweeps, we use forward simulations to show that background selection (BGS) alone can explain these same patterns. Moreover, we find that when negatively selected variants are partially recessive, effects of associative overdominance can be observed in regions of low recombination during expansions, despite very small population size on the wave front, where genetic drift usually overtakes selection. Finally, we find an excess of relative diversity loss in functional regions of the human genome that cannot be explained by BGS, suggesting that complex processes occurred after the expansion of human populations out of Africa. Our framework based on comparison of various populations along an expansion axis can be the foundation of a novel approach to quantify the impact of positive or purifying selection in human genomes, while controlling for demographic history.

718 - Investigation of features of recent positive selection at SLC8A1 region acting in the Japanese population

Risa L. Iwasaki

Risa L. Iwasaki, Yoko Satta

The Graduate University for Advanced Studies, SOKENDAI, The Graduate University for Advanced Studies, SOKENDAI

Mainland Japanese population has distinct three ancestral components from the Jomon hunter-gatherers, the Yayoi farmers, and the latter migrants from the East Asian continent. Population differentiation in the Japanese population has occurred through demographic changes and various environmental changes, including both environmental and cultural transitions resulting from admixture. Focusing on the effect of positive selection, we identified positively selected haplotypes, and then estimated the onset time and the intensity of selection coefficient, in order to assess how selection has shaped the genetic characteristics of the modern Japanese. We detected genome-wide candidate SNPs of positive selection, based on the difference in allele frequencies between Japanese and a genetically close East Asian population. Of them, we examined signatures of ongoing/completed positive selection against each haplotype with >10% frequency by using 2D-SFS, which evaluated the reduction of IAV in the focal haplotype. Here, we present a signature of ongoing positive selection on a haplotype nearby SLC8A1 in the Japanese with a frequency 20% higher than those in other East Asian populations. This haplotype has been selected since 18 kya with selection coefficient of 3%, whereas we could not detect any signatures of selection on any haplotypes in other East Asian populations. This result suggests that selected haplotypes are associated with recent population differentiation within East Asia. In the next step, we plan to classify genomic region into three ancestry-specific derived segments and discuss origins of selected haplotypes, in order to reconstruct the transition of adaptive phenotype in the Japanese.

723 - A compartmentalised model of sequence composition identifies underrepresented motifs exhibiting signs of purifying selection in Homo sapiens across UK Biobank whole genome sequencing data

Patrick Staunton

Patrick Staunton, Cathal Seoighe

University of Galway, University of Galway

Nullomers are DNA subsequences that do not appear in a given genome. Given that the number of possible DNA subsequences increases exponentially as a function of the subsequence length, for human studies, attention is typically focussed on the potential absence of shorter subsequences (i.e. fewer than 16 nucleotides in length).

Previous work on the human genome has identified nullomers expected to occur given the modelled probability of occurrence and the size of the genome. In addition, further studies have attempted to attribute function to identified nullomers.

Elaborating on previous models of sequence composition, we seek to address artefacts that influence the probability of occurrence. We achieve this by introducing a compartmentalisation framework that partitions the genome into disjoint regions to more accurately model heterogenous sequence composition effects. Moreover, we expand on the concept of missingness by considering, not only subsequences that are entirely absent, but, also, subsequences that are underrepresented relative to their expectation under the statistical model of sequence composition. In clustering underrepresented subsequences into cohesive classes, we form collections of significantly underrepresented motifs referred to as SUMs.

In many cases, we find that our compartmentalised approach improves model accuracy and alters underrepresentation status. Furthermore, in using these compartment-aware model results and a representative ancestral genome for Homo sapiens, we find evidence of purifying selection acting on classes of SUMs in UK Biobank whole genome sequencing data.

Given the evidence of purifying selection, we hypothesise that many of these SUMs may be involved in phenotypic variation.

734 - Stabilizing selection on human complex traits: the evidence from SNP-pair effect correlations

Evan Koch

Evan Koch, Mikhail Moldovan, Shamil Sunyaev

Harvard Medical School, Harvard Medical School, Harvard Medical School, Brigham and Women's Hospital

Analyses of human GWAS have the potential to inform us about selection acting on complex traits. Stabilizing selection has been proposed to broadly shape the genetic architecture of complex traits, including disease liabilities. Empirically, large-effect alleles have lower frequencies in a manner consistent with stabilizing selection, but such evidence is far from direct, and orthogonal approaches are needed. One strategy is to study patterns of LD which could be more specific to stabilizing selection. The Bulmer effect predicts that haplotypes carrying alleles with opposite effect directions will be overrepresented in the population as cancellation allows them to escape negative selection. A recent method, LDSPEC, estimates SNP-pair effect correlations (SPECs) between derived alleles. Correlations were consistently negative for nearby sites and within functional annotations, and were proposed to reflect stabilizing selection. We developed a population genetics model for SPECs and asked whether realistic scenarios of stabilizing selection generate strong correlations, and if other models of mutation and selection are similarly capable. Stabilizing selection, as well as Hill-Robertson interference under asymmetric mutation or selection, produce negative SPECs. However, the parameter range where stabilizing selection gives strong correlations is much wider, with Hill-Robertson models particularly sensitive to zero-effect mutations. We suggest that estimated SPECs can reflect stabilizing selection if causative mutations are clustered, capable of discordant directions, and have intermediate pleiotropic fitness effects. Our theoretical framework can be applied in studies linking selection, biological function, and complex trait genetics by leveraging information beyond allele frequencies and heritability enrichment.

743 - Estimating the Strength of Genome-wide Purifying Selection in Humans

Vince Buffalo

Rasmus Nielsen

UC Berkeley

Purifying selection at the top 10% most conserved sites in the human genome shapes patterns of genetic diversity across the genome. In turn, these genome-wide patterns of diversity can be used to infer the strength of selection, deleterious mutation rate, and the effective population size under drift. Here, we introduce a new quantitative genetic method to infer the strength of selection in humans from this linked selection signal. Unlike previous background selection models, our method handles weak selection and two-way mutations. This method uses highly accurate fine-scaled ARG-based estimates of pairwise coalescence times to estimate the distribution of selection coefficients for every base pair in the human genome using this linked selection signal. We show our model accurately predicts out-of-sample genetic diversity at different genomic scales, and is validated by realistic genome-scale simulations of background selection in humans. Our estimates of the strength of purifying selection have implications for understanding human health and the impact of selection on the genome.

795 - Genomic histories of the Adivasi and Sinhalese populations of Sri Lanka

Jose A Urban Aragon

Jose A Urban Aragon, Esha Bandyopadhyay, Constanza de la Fuente Castro, David Witonsky, Matthias SteinrÃ¼cken, Amali S Fernando, Anjana H Welikala, Joanne T Kotelawala, Kamani H Tennekoon, Ruwandi Ranasinghe, Niraj Rai, Maanasa Raghavan
University of Chicago, University of Chicago, University of Chicago, University of Chicago, University of Chicago, University of Colombo, University of Colombo, University of Colombo, University of Colombo, University of Colombo, Birbal Sahni Institute of Palaeosciences, University of Chicago

Sri Lanka, yielding the oldest-yet evidence of Homo sapiens in South Asia (~36,000 years ago), is central to our understanding of the early peopling of South Asia. Historically, Sri Lanka has been a vital node in the Indian Ocean trade network. Extensive migrations have, consequently, shaped the rich socio-cultural tapestry of this island nation. To investigate if the population history of Sri Lanka mirrors this socio-cultural diversity, we analyzed whole-genome sequence data from two present-day Sri Lankan populations: Adivasi (N = 19) and Sinhalese (N = 35). Based on some anthropological records, Adivasi are believed to be descendants of early inhabitants of Sri Lanka. Sinhalese are the most prevalent contemporary group in Sri Lanka, with oral traditions and chronicles suggesting that the Indo-European-speaking Sinhalese migrated from India in the first millennium BCE.

We evaluated the genome-wide affinities of the Sinhalese and Adivasi with global populations, including the Sri Lankan Tamils (STU, 1000G Project), finding that all three Sri Lankan populations fall on the modern South Asian genetic cline and share high levels of genetic drift each other and with South Indian populations. Moreover, while the three Sri Lankan populations were modeled with genetic sources used to model other South Asians, our analyses suggest that the Adivasi have maintained a small long-term effective population size, reflecting a unique demographic history compared to the other two Sri Lankan populations. Overall, our study provides novel genomic insights on the population structure of Sri Lanka.

807 - Uncovering the genetic architecture and evolutionary roots of androgenetic alopecia in African men

Ujani Hazra

Ujani Hazra, Rohini Janivara, Aaron Pfennig, Timothy Rebbeck, Akindele O Adebiyi, Ilir Agalliu, Joseph Lachance
Georgia Institute of Technology, Georgia Institute of Technology, Georgia Institute of Technology, Dana-Farber Cancer Institute, University of Ibadan, Albert Einstein College of Medicine, Georgia Institute of Technology

Androgenetic alopecia or male pattern baldness (MPB) is a highly heritable trait. However, much of our understanding of the genetics of MPB comes from individuals of European descent. Examining a novel dataset comprising 2,136 men from Ghana, Nigeria, Senegal, and South Africa that were genotyped using a custom genotyping array, we conducted the first African GWAS of MPB and inferred the evolutionary history of genetic variants associated with this complex trait. We found evidence that the genetic architecture of MPB varies across continental populations and identified 51 independent associations in Africa (p-value < 10^{-5}, r^2 < 0.2). Although Neanderthal alleles have previously been associated with skin and hair phenotypes, we did not find enrichment for signatures of ancient introgression in European-ascertained baldness hits. Next, we tested for signatures of polygenic selection on MPB using integrated Haplotype Scores. Regardless of the ascertainment scheme, autosomal SNPs were not enriched for outlier iHS statistics. However, X-chromosomal SNPs (rs12558842, rs2497911, rs1204041) near the EDA2R and AR genes were found to have large allele frequency differences between continents. These variants were in linkage disequilibrium and multiple divergent haplotypes were found in high frequency in Africa and Europe. Although these patterns are consistent with the Xq12 region having undergone multiple independent selection events we note that Xq12 is located near the centromere in a genomic region with low recombination rates. Because of genetic hitchhiking, divergent haplotype frequencies at Xq12 need not have been caused by selection acting directly on male pattern baldness.

815 - The genetic population history of the last Himalayan hunter-gatherers

Inez Derkx

Inez Derkx, Simone A Biagini, Sudarshan Subedi, Prajwal Rajbhandari, Elena Bosch, Jaume Bertranpetit, Lucio Vinicius, Andrea Migliano

University of Zurich, Universitat Pompeu Fabra, Tribhuvan University, Research Institute for Bioscience and Biotechnology (RIBB), Universitat Pompeu Fabra, Universitat Pompeu Fabra, University of Zurich, University of Zurich

The Himalayas, with their complex terrain and diverse populations, have played a pivotal role in shaping human migration and adaptation in Eurasia. While extensive research has focused on various Nepali populations like the Sherpas, the Raute of western Nepal, a nomadic hunter-gatherer group of roughly 150 individuals, remain relatively unexplored. With their foraging lifestyle under increasing pressure, it is crucial to investigate their genetic structure and potential population history and origins. This study aims to shed light on the Raute's genetic structure and their place within the broader Himalayan genetic and historical context. We analysed newly genotyped data from the Raute and individuals from neighbouring settlements to examine patterns of consanguinity, population history, and recent genetic connectivity. Our results not only augment the existing linguistic and cultural narratives of the Raute from a genetic standpoint but also highlight the methodological challenges in researching small, consanguineous groups. Through this research, we contribute to the broader comprehension of the Himalayan region's intricate human history and the existence of hunter-gatherer societies within it.

821 - Dynamic Rates and Patterns of Nucleotide Substitutions in Ape Telomere-to-Telomere Genomes: Substantial Effects of Sex Chromosomes

Xinru Zhang

Xinru Zhang, Saswat K Mohanty, Robert S Harris, Francesca Chiaromonte, Kateryna D Makova, The T2T Primates Consortium

Department of Biology, The Pennsylvania State University, Huck Institutes of the Life Sciences, The Pennsylvania State University, Department of Statistics, The Pennsylvania State University, Department of Biology, The Pennsylvania State University, Huck Institutes of the Life Sciences, The Pennsylvania State University, Department of Biology, The Pennsylvania State University, Department of Statistics, The Pennsylvania State University, Huck Institutes of the Life Sciences, The Pennsylvania State University, Center for Medical Genomics, The Pennsylvania State University, Department of Biology, The Pennsylvania State University, Huck Institutes of the Life Sciences, The Pennsylvania State University, Center for Medical Genomics, The Pennsylvania State University, Department of Biology, The Pennsylvania State University, National Human Genome Research Institute, National Institutes of Health, University of Washington School of Medicine

Mutation is a major evolutionary force that generates genetic diversity for the other evolutionary forces to operate on. Mutation rates vary among species, over time, and across different genomic regions. DNA sequence contexts also substantially affect mutation rates. Yet, the extent of how sequence context influences nucleotide substitutions remains understudied.

Leveraging the recently released Telomere-to-Telomere (T2T) whole-genome assemblies for most extant great ape species, we were able to construct the most comprehensive tri-nucleotide substitution spectrum in apes. This has allowed us to compare nucleotide substitution spectra across seven ape species and examine its variation among different phylogenetic branches and genomic locations.

Our analyses revealed chromosomal heterogeneity in the distribution of substitutions. Specifically, chromosomes Y and X demonstrated unique substitution patterns, as compared to those for the autosomes. Comparing nucleotide substitution spectra between the two sex chromosomes, we found transversions to be significantly more abundant on the Y than on the X, and transitions vice versa. These findings are broadly consistent with sex-specific signatures of de novo mutations from other studies; C>A, C>G and T>G were shown to be

enriched in paternal de novo mutations, whereas C>T mutationsâ€”in maternal de novo mutations. C>G might be related to meiotic double-strand breaks in the male germline. Additionally, we compared substitution rates and spectra in the X and Y pseudoautosomal regions (PARs) and on the autosomes. Our results allow us to formulate hypotheses about molecular mechanisms of mutagenesis that should be tested in subsequent wet-lab experiments.

938 - Selective dynamics of interruptions at short tandem repeats

Michael E Goldberg

Michael E Goldberg, Harriet Dashnow, Kelley Harris, Aaron R Quinlan
University of Utah, University of Utah, University of Washington, Fred Hutchinson Cancer Research Center, University of Utah

Short tandem repeats (STRs) are hotspots of genomic instability that mutate at rates orders of magnitude greater than non-repetitive loci due to frequent replication slippage. Expansions at some STR loci are linked to Mendelian diseases, while variation at other noncoding loci may affect complex traits, possibly by altering transcription factor occupancy of nearby binding sites. Accordingly, some STRs are inferred to be under purifying selection, regardless of their instability.

One or more â€˜interruptionsâ€™, or bases that disrupt the locusâ€™s canonical repeat, significantly decrease an STRâ€™s mutation rate. However, interruptions may themselves be deleterious at constrained loci, particularly at noncoding loci in gene regulatory elements, possibly disrupting the formation of secondary structures key to their function. We therefore hypothesized that the frequency of interruptions could depend on a locusâ€™s purifying selection, where the fitness effects of both expansions but also interruptions could be more deleterious than at neutral loci.

To test this hypothesis, we examined the distribution of interruptions at ~650,000 autosomal 2-6 bp motif STRs. We find that noncoding STRs under purifying selection harbor fewer interruptions than those evolving neutrally. In contrast, purifying selection is positively associated with interruptions in coding STRs. Our findings indicate that the abundance of interruptions may be partially explained at coding STRs by the benefit of a lower mutational burden at their linked loci. In contrast, maintaining a minimum core stretch of uninterrupted repeat may be critical for the function of noncoding STRs that fall within regulatory elements, outweighing the benefits of lowering mutation rate.

949 - Using genomic data to unravel the complexity of selection on HLA genes

Diogo Meyer

Diogo Meyer, AndrÃ© S MarÃ³stica, Montgomery Slatkin, Erick C Castelli
University of SÃ£o Paulo, University of SÃ£o Paulo, University of SÃ£o Paulo, University of California, Berkeley, Universidade Estadual Paulista

The HLA loci, involved in the innate immune response, are known to be under balancing selection. However, questions about the intensity, target (amino acids, alleles, or haplotypes) and timescale of selection remain open. These can be approached using NGS to integrate HLA allele calls with the SNP data for the intergenic regions of the MHC, where HLA genes are located. Haplotypes encompassing SNPs and HLA alleles can identify haplotypes under recent selection, and provide estimates of allele and haplotype age. Until recently, this task was hampered by the difficulty in making reliable calls for HLA alleles using NGS data, due to the mapping bias that affects extremely polymorphic regions.
We developed a computational pipeline that uses multiple reference sequences to align short reads, thus avoiding mapping bias. This

pipeline provides HLA allele calls that can be integrated with phased SNPs for the entire MHC region. We applied this pipeline to over 3,000 individuals from various regions of the world and found haplotypes of over 3Mb in several populations at frequencies of over 1%. These haplotypes are geographically, restricted, and display low nucleotide diversity (Ï€~0.0004), indicating they are young. Using intron variation we found that most alleles, when defined at the protein level, have low diversity (Ï€~0.0002) .
These results indicate that while long-term balancing selection maintains individual SNP positions segregating for spans of time greater than the human-chimp divergence, HLA haplotypes and alleles experience turnover at much shorter timescales, and are relatively young.

962 - Comparing Neanderthal introgression maps reveals substantial heterogeneity across algorithms, populations, and assumptions

Yaen Chen

Yaen Chen, John A. Capra
University of California, San Francisco, University of California, San Francisco

Motivation: Statistical methods to identify Neanderthal ancestry rest on different assumptions and inputs. Nonetheless, most studies of Neanderthal introgression use only a single method. Due to a lack of â€œground truth,â€• we have a limited understanding of comparative strengths, weaknesses, and accuracy of these methods.

Results: We performed large-scale comparisons of introgressed variant maps from 12 representative Neanderthal introgression detection algorithms. These span methods using archaic and African human reference genomes (ArchaicSeeker2, CRF, DICAL-ADMIX), only archaic genomes (S*, Sprime, HMM, SARGE, ARGWeaver-D), only African human reference genomes (IBDmix), or neither (ArchIE).

Overall, ~66% of the genome is called as introgressed by at least one method, and 15.7% of introgressed variants are unique to a method. We identify high pairwise overlap between most methods, excluding ArchIE and Argweaver-D, which identified substantially less introgression. Hierarchical clustering reveals that algorithms built with similar motivations do not always have the most similar predictions. Regions unique to a single method often show enrichment for known Neanderthal and modern human differences: for example, loci unique to ArchIE are significantly enriched for long clavicle formation. Thus, unique predictions may not reflect errors, but rather different strengths of the methods. Loci identified by all 12 algorithms are significantly enriched for iron homeostasis, motor coordination, and skeletal phenotypes.

Conclusion: Our results highlight substantial heterogeneity in commonly used Neanderthal introgression maps. We make integrated prediction sets available to enable further understanding of the legacy of Neanderthal introgression in modern humans and ensure robustness of future results.

1006 - Modeling gene regulatory mechanisms contributing to the evolution of the human cerebral cortex

Marybeth Baumgartner

Marybeth Baumgartner, Yu Ji, Mark A Noble, Atreyo Pal, James P Noonan
Yale School of Medicine, Yale School of Medicine, Yale School of Medicine, Yale School of Medicine, Yale School of Medicine

The vast expansion of the human cerebral cortex distinguishes us from our primate relatives, and this cortical expansion is the foundation of uniquely human higher-order cognition. Numerous developmental innovations, such as increased proliferation of cortical progenitor

cells, contributed to this cortical growth. Ultimately, these developmental innovations arose from genetic changes in the human lineage, which altered the molecular and cellular programs underpinning development. Efforts to illuminate these human-specific genetic changes have identified thousands of Human Accelerated Regions (HARs), highly conserved regulatory elements that exhibit a high rate of human lineage-specific sequence change, and Human Gain Enhancers (HGEs), regulatory elements with increased enhancer activity in the fetal human brain compared to that of non-human primates and mice. To understand how HARs and HGEs impact the events of corticogenesis, it is essential to model the effect of these elements on cortical development in vivo. While a growing body of evidence from cross-species regulatory genomics datasets and in vitro assays have implicated a subset of HARs/HGEs in human cortical development, this shortlist of candidate elements still numbers in the hundreds. Here, we present approaches to (i) identify HAR and HGE candidates with high predicted effect sizes, based on the increasing number of high-quality primate genomes, chromatin contact data, and functional genomics datasets available; and (ii) broaden the scope of in vivo model design beyond single HARs/HGEs, to accurately replicate the human regulatory landscape and thereby maximize the likelihood of recapitulating uniquely human molecular and developmental phenotypes.

1046 - Investigating the relationship between runs of homozygosity and changes in human height over 35,000 years

Ana Victoria Leon Apodaca

Ana V. Leon-Apodaca, Zachary A. Szpiech

Pennsylvania State University, Pennsylvania State University

Demographic processes such as consanguinity, bottlenecks, and natural selection increase the likelihood of haplotypes being inherited identical-by-descent (IBD) from a recent common ancestor. When these identical haplotypes are inherited through both the maternal and paternal lineages, they manifest as runs of homozygosity (ROH). In humans, ROH is commonly associated with recessive Mendelian disorders, however, increased homozygosity has also been associated with complex traits such as height. Height is a highly heritable and polygenic trait influenced by both genetic and environmental factors. Research conducted on present-day human populations of European descent have found an association between genome-wide homozygosity and height, with more homozygosity resulting in lower height. Archaeological evidence and genomic data obtained from ancient human remains spanning the Upper Paleolithic (38,000 yBP) to the Iron Age (2,400 yBP) have shown there was a marked reduction in human height that occurred during the transition to agriculture that took place during the Neolithic, followed by an increase during the subsequent post-Neolithic periods of agricultural intensification. In this work, we investigate the relationship between historical demographic changes, ROH, and height; we assess whether changes in ROH patterns observed across ~35,000 years across ancient European populations are associated with human height, and whether genetic variants known to influence height are more often observed within ROH regions.

1065 - Exploring selective scanning by linkage disequilibrium statistic Dz: simulations and empirical studies

Alouette Zhang

Alouette Zhang, Kevin R Thornton, Aaron P Ragsdale, Simon Gravel

McGill University, Kyoto Univeristy, University of California, Irvine, University of Wisconsinâ€“Madison, McGill University

Selective sweeps are significant evolutionary events where beneficial variants increase rapidly in frequency and reach fixation in the population. Sweeps affect diversity at surrounding variants via hitchhiking. Several methods exploit the excess linkage disequilibrium (LD) pattern caused by hitchhiking for sweep detection.

In this project, the LD statistic Dz is utilized to detect selective sweeps in simulations and in 1000 Genomes Project data. Dz was introduced by Hill and Robertson as a mathematical stepping stone to compute D^2, which is the usual measure of covariance between allele frequencies. Dz computes the degree of LD between low frequency variants. It has recently emerged as a useful empirical measure of diversity. We show that the expectation of Dz is more informative about recent and ancient sweeps compared to other commonly used LD statistics, with adequate power for up to 100 thousand years ago in simulation with human-like parameters. Incorporating expected Dz in a scanning window in different populations from 1000 Genomes Project data recaptures sweep signals from widely studied sweeps. This project demonstrates Dzâ€™s potential for incorporation into the current line of sweep detection techniques, taking advantage of genetic patterns not considered by most current methods.

1082 - Fine-scale genetic structure and natural selection within and across ethnic groups of Pakistan

Rodelmar Ocampo

Rodelmar Ocampo, MarÃ-a JosÃ© Palma, Luis Enrique RamÃ-rez-Serrano, Zara P MartÃ-nez-SÃ¡nchez, Mashaal Sohail
Center for Genomic Sciences, Center for Genomic Sciences, Center for Genomic Sciences, Center for Genomic Sciences, Center for Genomic Sciences

Pakistan is the fifth most populous country in the world, home to a vast cultural diversity represented in multiple languages and cultural practices embedded in ethnic identities, yet remains under-studied genetically. We analyzed 9 Pakistani ethnic groups from the newly released joint call of the Human Genome Diversity Project and 1000 Genomes Project, a total 279 individuals from Brahui, Balochi, Burusho, Hazara, Kalash, Makrani, Pathan, Punjabi and Sindhi groups. We infer fine-scale genetic structure across Pakistan and in relation to global groups. Amidst the results, Makrani showed an enrichment of African ancestries. To further explore this finding we infer the timing of arrival of African groups and their specific origins using Gnomix, tracts and MAAS-MDS with an African reference panel. Further, we inferred natural selection per SNP and gene using the Population Branch Statistic (PBS). Focusing on the top 99th percentile of PBS values, we found unique as well as shared genes under natural selection across ethnic groups. For example, we found genes like CTXN2 and SLC24A5 (associated with skin pigmentation) highly selected in Pakistan as a whole compared to Bengali and Yorubans, as well as the APPAT lncRNA, associated with atherosclerosis, highly prevalent in Pakistan. We observe the highest PBS value for LINC00319 in Makrani, Kalash and Brahui groups (also in top 99th-percentile of all groups except Punjabi) that is related to 9 cancer types. Another interesting lncRNA found in Balochi, Brahui and Sindhi, LINC00313, is related to cancer but also to osteoarthritis development inhibition.

1090 - Investigating the role of early neurodevelopment in human brain evolution

Breanna Takacs

Breanna J Takacs, Soojin Yi, Sebastian Streichan, Susan Wopat, Matthew Lefebvre
University of California, Santa Barbara, University of California, Santa Barbara, University of California, Santa Barbara, University of California, Santa Barbara, University of California, Santa Barbara, University of California, Santa Barbara

The human brain has evolved distinct traits compared to nonhuman primates, including differences in absolute size, relative size of prefrontal cortical regions, distribution of cell populations, and changes in gene regulation and expression. Literature indicates that these variations stem from a longer duration of neurodevelopment during late fetal and postnatal stages in humans compared to nonhuman primates. Early neurodevelopment is generally conserved across many vertebrate species, involving the transformation of a flat neural plate into a tube shape through extension, bending, and folding of neural tissue. This process has been observed through in-vivo morphological and molecular comparisons across zebrafish, frogs, chicks, mice, and humans. However, these stages differ in detail and give rise to the central and peripheral nervous systems and thus may be the origin of fine grain, human-specific traits in the brain.

Advancements in in-vitro organoid models have expanded the scope of comparative studies on early ape brain development, which were previously limited by the availability of subjects. These aforementioned factors motivate the investigation of the the role of early neurodevelopment in the evolution of the human brain through the generation of novel 3D organoids representing early neurodevelopment from stem cells of various primate species, as well as a comparison of early and late neurodevelopment using previously published RNA-seq data.

1118 - Reconstructing the Demographic History of South Asia: From Antiquity to Modernity

Niraj Rai

Niraj Rai
Birbal Sahni Institute of Palaeosciences

Reconstructions of historical effective population sizes strongly indicate that South Asia has been inhabited since the Early Upper Palaeolithic era. This proposition stems from the fact that South Asia was one of the earliest regions settled by modern humans following the Out of Africa migration. South Asia is characterized by a diverse array of tribes, castes, and religiously identified populations. Many of these groups exhibit high levels of endogamy, revealing intricate genetic distinctions within the region. The complexity of South Asian population dynamics has spurred numerous inquiries from our research group and others. Until recently, these questions remained largely unanswered, but advancements in technology and resolution have facilitated analysis. New methodologies leverage comprehensive sequencing of modern and ancient samples, autosomal SNP genotyping of specific populations, extensive resequencing of the Y chromosome in selected groups (including tribal and caste populations from various geographic regions), and statistical analysis of results in the broader context of global genomic variation. In my presentation, I will discuss and present data concerning the intricate population structure of the Indian Subcontinent, as well as future research avenues aimed at comprehending the deep ancestral components of South Asians on Eurasian and global scales. These investigations will also shed light on patterns of admixture and migrations by studying both ancient and modern genomes.

1131 - Polygenic adaptation leads to a higher reproductive fitness of native Tibetans at high altitude

Yaoxi He

Yaoxi He, Bing Su
Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming Institute of Zoology, Chinese Academy of Sciences

The adaptation of Tibetans to high-altitude environments has been studied extensively. However, the direct assessment of evolutionary adaptation, i.e. the reproductive fitness of Tibetans and its genetic basis remain elusive. Here, we conduct systematic phenotyping and genome-wide association analysis of 2,252 mother-newborn pairs of indigenous Tibetans, covering 12 reproductive traits and 76 maternal physiological traits. Compared to the lowland immigrants living at high altitudes, indigenous Tibetans show better reproductive outcomes, reflected by their lower abortion rate, higher birth weight, and better fetal development. The results of genome-wide association analyses indicate a polygenic adaptation of reproduction in Tibetans, attributed to the genomic backgrounds of both the mothers and the newborns. Furthermore, the EPAS1-edited mice display higher reproductive fitness under chronic hypoxia, mirroring the situation in Tibetans. Collectively, these results shed new light on the phenotypic pattern and the genetic mechanism of human reproductive fitness under extreme environments.

174 - 300 Whole Genomes of People from the border of Ukraine and Romania

Taras K Oleksyk

Taras K Oleksyk, Walter W Wolfsberger, Khrystyna K Shchubelka
Oakland University - Biology, Uzhhorod National University, Oakland University - Biology, Oakland University - Biology, Uzhhorod National University

We present the results of the Whole Genome Sequencing of two human populations in the Carpathian Mountains region in Eastern Europe, specifically Ukraine's Transcarpathia and Romania's Satu Mare and Baia Mare provinces, areas previously underexplored in population genomics. The database contains the raw and annotated files of the whole genome sequences from 300 individuals from these regions, including annotations of common and unique genetic variants following a sampling protocol designed to capture the genetic diversity of Ukrainians and Romanians, including minority groups like Wallachians and Roma. The data is hosted on a dedicated web resource. We provide information on how to access to results of primary and secondary analysis of the data, including comparative analysis with previously published populations from Ukraine, and populations from the International Genome Sample Resource and Human Genome Diversity Project. The free research access to this database is contributing to a growing understanding of human genetic diversity in Central Europe. This effort emphasizes the potential for reuse of the generated data, advocating for open access to support future research in genomics, bioinformatics, and personalized medicine.

757 - Tandem tales: comparative analysis of tandem repeats in Ape genomes

Carolina de Lima Adam

Carolina de Lima Adam, Joana L Rocha, Peter Sudmant, Rori Rohlfs
University of Oregon, University of California Berkeley, University of California Berkeley, University of Oregon

Tandem repeats (TRs) are important contributors to genetic variation in eukaryotes. Their high mutation rate makes them strong candidates for surveys of recent evolution across species. Also, recent studies highlight the association of TRs with gene expression and complex traits. However, commonly used short-read sequencing technologies have failed to sequence long, repetitive TRs and their multiallelic nature is a challenge for genotyping tools designed for biallelic markers. Our investigation uses long-read, telomere-to-telomere high-fidelity genome data, and TR-specific genotyping software to identify TR loci in seven ape species. Our results revealed that distributions of TR total length and repeat unit size are conserved across all examined species, with a high frequency of primate-specific transposable elements, including Alu elements. To obtain homologous TRs between humans and the remaining species we performed liftover analysis with the human TR catalog as the target. The resulting orthologous regions were then intersected with the TR catalog of each species. In the chimp x human comparison, ~90.7% of the TR-containing regions are orthologous. By employing a phylogenetic framework across all seven species, we can access signs of past selection. TR phylogenies with exceptional species-specific long branches suggest directional selection, while TR loci subjected to balancing selection will exhibit deeper genealogies due to the maintenance of ancient polymorphisms. This comprehensive investigation will advance our understanding of TR evolution and provide valuable insights into the potential role of TRs in complex traits and rapid adaptation in primates

172 - Historic migration routes of Native American males: reconstructing distribution patterns of Y-chromosomal haplogroup Q sublineages throughout South America

Zehra Köksal

Zehra KÃ¶ksal, Claus BÃ¸rsting, Graciela Bailliet, GermÃ¡n Burgos, Elizeu Carvalho, Adriana Castillo, Carlos David Neyra Rivera, Danielle Malheiros, Marilia Brito Gomes, Beatriz MartÃ-nez, Humberto Ossa, MarÃ-a Laura Parolin, Alfredo Quiroz, Ulises Toscanini, Irina VelÃ¡zquez, Carlos Vullo, Leonor GusmÃ£o, Vania Pereira
Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen,

Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Instituto Multidisciplinario de BiologÃ-a Celular, Universidad Nacional de La Plata, CCT-CONICET-La Plata, CIC, One Health Global Research Group, Facultad de Medicina, Universidad de Las AmÃ©ricas (UDLA), Grupo de Medicina XenÃ³mica, Universidad de Santiago de Compostela, DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Department of Basic Sciences, Universidad Industrial de Santander (UIS), Universidad Privada Peruano Alemana, Laboratory of Human Molecular Genetics, Postgraduate Program in Genetics, Federal University of ParanÃ¡ (UFPR), Department of Internal Medicine, Diabetes Unit, State University of Rio de Janeiro (UERJ), Instituto de Investigaciones InmunolÃ³gicas, Universidad de Cartagena, LaboratÃ³rio de GenÃ©tica y BiologÃ-a Molecular, Facultad de Ciencias, Pontificia Universidad Javeriana, Instituto de Diversidad y EvoluciÃ³n Austral (IDEAus), Centro Nacional PatagÃ³nico, CONICET, Instituto de PrevisiÃ³n Social, Primer Centro Argentino de InmunogenÃ©tica (PRICAI), FundaciÃ³n Favaloro, Instituto de Diversidad y EvoluciÃ³n Austral (IDEAus), Centro Nacional PatagÃ³nico, CONICET, DNA Forensic Laboratory, Equipo Argentino de AntropologÃ-a Forense (EAAF), DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen

The Native Americansâ€™ (NAMs) migration routes through South America are not entirely understood. Phylogenetic studies on Y-chromosomal haplogroups defined by SNPs, give a unique view on male geographic dispersals and can be used to trace back migrations in historic times. The two main NAM founding lineages Q1b1a1a1-M848 and Q1b1a2-Z780 harbor valuable information on some of the long-lost NAM migration routes in South America.

Here, we introduce the custom-designed massively parallel sequencing panel NAM-Q-Y, which targets 1,569 variants of NAM haplogroup Q sublineages. Data were generated in 422 admixed individuals of NAM male ancestry from different regions of South America.

Sequencing data from the 422 individuals were combined with 222 publicly available genomes from Central and South American individuals, which further resolved the phylogenetic tree of the NAM haplogroup Q lineages, introducing more than 100 novel or refined branches.

The distribution of selected sublineages within Q1b1a1a1-M848 and Q1b1a2-Z780 allowed analyses of regional migration patterns that supported previous hypotheses on (1) southward movements along the Andes, (2) independent westward and eastward movements from the north of South America, and (3) southward Andean spread followed by eastward movements, where the population split into a northward (to Northeast Brazil) and a southward (to East Argentina) migrating subpopulation. Furthermore, the first clear patterns of inland movements in the Southern Cone were reported based on Y-chromosomal data.

The presented findings support and complement previously hypothesized movements, and increase our understanding of the NAMs migration routes through South America.

239 - Environmental adaptation in isolated indigenous populations

Elena S. Gusareva

Elena S. Gusareva, Ghosh Amit Gourav, Namrata Kalsi, Vladimir N. Kharkov, AndrÃ©s Moreno-Estrada, Vadim A. Stepanov, Stephan C. Schuster, Hie Lim Lim

Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, The Asian School of the Environment, Nanyang Technological University, GenomeAsia 100 consortium, Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore , GenomeAsia 100 consortium, Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore , GenomeAsia 100 consortium, Research Institute of Medical Genetics, Tomsk National Research Medical Center, Russian Academy of Sciences, GenomeAsia 100 consortium, Laboratorio Nacional de GenÃ³mica para la Biodiversidad, Unidad de GenÃ³mica Avanzada (UGA-LANGEBIO), CINVESTAV, GenomeAsia 100 consortium, Research Institute of Medical Genetics, Tomsk National Research Medical Center, Russian Academy of Sciences, GenomeAsia 100 consortium, Singapore

Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore , GenomeAsia 100 consortium, The Asian School of the Environment, Nanyang Technological University, Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, GenomeAsia 100 consortium

Humans have successfully adapted to a diverse habitats on our planet. Populations residing in extreme environments exhibit specific physiological characteristics that enable them to endure challenging conditions. These advantageous traits result from acquired genetic variations selected through the evolutionary process. Individuals with genetic adaptations to a specific environment may face potential risks if that environment undergoes rapid changes. This highlights the dynamic nature of the relationship between genetic adaptations and environmental conditions, underscoring the need for ongoing research.
Our focus is on studying indigenous populations that developed high level of specialization to their habitat environments. Our research in the highlander Andean populations of South America identified a significant selection signal in the EPAS1 gene, crucial for regulating red blood cell production and elevated oxygen fixation in high-altitude environments. Notably, this gene has been previously reported in studies on Tibetan populations. Here, we identify a novel missense mutation, rs570553380 (His>Arg), specific to Andeans, highlighting their unique genetic adaptations.
As climate change rapidly accelerates becoming warmer each year, our research extends to the study of Arctic populations. We found a selection signal in the CPT1A gene that is important for fat metabolism in northern populations. The adaptive Leu479 allele of the rs80356779 variant is prevalent in Arctic populations (frequency ranges from 75 to 90%), while absent elsewhere.
Our study illuminates the intricate interplay between genetic adaptations and environmental challenges across diverse human populations, emphasizing the urgency of ongoing research for adaptive strategies in the face of evolving conditions.

268 - Ancient genomic data reveals multiple genetic shifts in the demographic history of Colombia

Kim-Louise Krettek

Kim-Louise Krettek, Lilian Andrea Casas Vargas, Alex Mas-Sandoval, Leonardo Arias Alvis, Ella Reiter, William Usaquen, Jose-Vicente Rodriguez Cuenca, Cosimo Posth
Senckenberg Centre for Human Evolution and Paleoenvironment, University of TÃ¼bingen, Grupo de GenÃ©tica de Poblaciones e IdentificaciÃ³n, Universidad Nacional de Colombia, Dipartimento di Scienze Statistiche "Paolo Fortunati", UniversitÃ¡ di Bologna, Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leiden University Centre for Linguistics, Archaeo- and Palaeogenetics Group, Institute for Archaeological Sciences, University of TÃ¼bingen, Grupo de InvestigaciÃ³n en AntropologÃ-a BiolÃ³gica, Universidad Nacional de Colombia, Grupo de GenÃ©tica de Poblaciones e IdentificaciÃ³n, Universidad Nacional de Colombia, Senckenberg Centre for Human Evolution and Paleoenvironment, University of TÃ¼bingen, Archaeo- and Palaeogenetics Group, Institute for Archaeological Sciences, University of TÃ¼bingen, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology

The first peopling of America via the Beringian land bridge has been extensively researched. Scholars have proposed that the ancestral population that first reached North America split into two main lineages, one persisting in the northern sub-continent and the other spreading southwards to reach the southern tip of South America by at least 14,500 years ago. Recent archaeogenetic studies have begun outlining the genetic history of various South American regions such as the Andes, the Atlantic coast, and the Southern cone. However, Colombia, the entry point into the southern sub-continent, has so far been omitted from ancient genomic studies.
To illuminate Colombia's crucial role in the initial settlement of South America and reconstruct the demographic history of local populations, we generated genome-wide data from 21 ancient individuals spanning the last 6,000 years. Our genomic time-transect, encompassing a spatially confined area on the central Colombian plateau, reveals at least two shifts in genetic ancestry over time, one

associated with Mesoamerican populations and the other with present-day indigenous Colombian groups. Moreover, we unveil a complex interplay between ancient populations from Colombia and the circum-Caribbean region.
This diachronic investigation into demographic transformations within Colombia holds promise for uncovering previously unknown population movements in northern South America. Additionally, our newly generated data enriches the sparse ancient genomic record from the tropics, fostering further population genetic analyses across the American continent.

301 - SIGNATURES OF NATURAL SELECTION ON TWO MALARIA RESISTANCE GENES IN AFRO AND NON-AFRO POPULATIONS OF THE COLOMBIAN PACIFIC

Diana Carolina Ortega

Diana Carolina Ortega, Heiber CÃ¡rdenas, Guillermo Barreto
Universidad del Valle, Universidad del Valle, Universidad del Valle

The Duffy and G6PD genes in humans have been shown to be subject to selection due to the resistance or protection against malaria conferred by their variants. We evaluated signatures of natural selection on the loci of these genes in Afro and non-Afro populations of the Colombian Pacific. Sequencing of approximately 3 kb regions from 42 Afro and 21 non-Afro individuals for the Duffy gene, and 47 Afro and 22 non-Afro individuals for the G6PD gene, was conducted. Descriptive statistics and deviations from neutrality were evaluated using DnaSP. In the Afro population, the Duffy gene displayed signs of admixture, possibly due to positive selection favoring the FYBES allele. This allele was polymorphic in the total sample due to mixing with chromosomes of other ancestries lacking the allele. The non-Afro population showed reduced diversity and loss of polymorphisms in the Duffy gene, suggesting purifying selection to maintain correct gene transcription while other gene regions experienced deleterious mutations under weak purifying selection. For the G6PD gene, the Afro population exhibited patterns consistent with balancing selection, showing higher nucleotide diversity and the presence of polymorphisms at intermediate frequencies, both in coding and intronic regions. The non-Afro population showed low diversity in the G6PD gene, with chromosomes lacking polymorphisms, indicating possible purifying selection preserving the region associated with normal enzyme activity. Both populations exhibited selective signatures in line with expectations for resistance variants, though interpretation necessitates consideration of demographic and mixing factors that may influence traditional neutrality tests' outcomes.

378 - From North Asia to South America: Tracing the longest human migration through genomic sequencing

Hie Lim Kim

Hie Lim Kim
Nanyang Technological University

We conducted an analysis of genome sequencing datasets from 1,537 individuals belonging to 139 ethnic groups to shed light on the peopling of North Asia and the Americas. Our findings indicate that the West Siberian ancestry, represented by the Kets, played a major role in shaping Siberian populations, including 17 indigenous Siberian ethnicities. This is supported by the population growth of West Siberians until ~10,000 years ago. We also found that Beringian populations, including Eskimo, Chukchi, and Koryaks, were the closest ancestral groups to Native Americans, with recent gene flow between them. Furthermore, we were able to demonstrate that the earliest inhabitants of South America divided into four ancestral groups - Amazonians, Andeans, Chaco Amerindians, and Patagonians - around 13,900 years ago. This population structure corresponds closely with biogeographic boundaries in South America, suggesting the impact of environmental factors on population history. The vastness of the South American continent likely resulted in immediate spatial isolation, leading to a significant reduction in genetic diversity following migration. The Patagonian ethnicities, including the Yagan and

KawÃ©sqar, have traveled the longest path out of Africa, leading to a small effective population size due to genetic drift. This, in turn, has resulted in reduced immunogenic diversity, increasing their vulnerability to past and current pathogens.

394 - Identifying signals of natural selection in unadmixed individuals of the Mexican Biobank

Maria Fernanda Miron Toruno

Maria F Miron-Toruno, Ram GonzÃ¡lez-Buenfil, Santiago Medina-MuÃ±oz, AndrÃ©s Moreno-Estrada, Austin W Reynolds
Baylor University, Unidad de GenÃ³mica Avanzada, Cinvestav, Unidad de GenÃ³mica Avanzada, Cinvestav, Unidad de GenÃ³mica Avanzada, Cinvestav, Baylor University

Over the past decade, the exponential increase in computational power and the development of next-generation sequencing technologies have facilitated the identification of genomic regions that appear to be shaped by natural selection. Studies of natural selection represent a genomic approach to understanding human evolution and differences in disease risk across many populations worldwide. However, the limited availability of diverse cohorts with genomic data impairs the effective implementation of selection analyses in many underrepresented regions, such as Latin America. Recently, Mexico generated the Mexican Biobank (MXB), one of the largest genomic resources of nationwide coverage in the Americas. The MXB comprises over 6,000 individuals from all 32 states and 898 sampling localities across Mexico, including over 1,500 unadmixed individuals of indigenous ancestry, genotyped for 1.8 million genome-wide SNPs. In this project, we aim to investigate the genetic footprint of recent adaptive pressures on Indigenous Mexican populations by analyzing a subset of ancestry-masked samples with >90% Indigenous American ancestry from the MXB. To do this, we will compute different selection statistics, including iHS, XP-EHH, nSL, and PBS. Moreover, we will be able to connect putatively selected alleles to molecular phenotypes reported in the MXB to better understand how selection has shaped traits of biomedical relevance. To date, this project represents the most extensive selection study in a national-scale Mexican cohort and the first to pair selection analyses with biomedical data to gain insight into the genetic basis of Mexican health-related traits.

433 - Exploring Recent Adaptation in the Americas: Genome-Wide Analysis of an admixed Cohort

Franco Leonel Marsico

Franco Leonel Marsico, Silvia Buonaiuto, Akram Mohammed, Lokesh Chinthala, Chester Brown, Terri Finkel, Robert Rooney, Robert Williams, Robert Davis, Vincenza Colonna
Dept of Genetics, Genomics and Informatics, UTHSC, Dept of Genetics, Genomics and Informatics, UTHSC, Center for Biomedical Informatics, UTHSC, Center for Biomedical Informatics, UTHSC, Dept of Pediatrics, Genetics Division, UTHSC, Dept of Pediatrics, Division of Rheumatology, UTHSC, Dept of Pediatrics, Genetics Division, UTHSC, Dept of Genetics, Genomics and Informatics, UTHSC, Center for Biomedical Informatics, UTHSC, Dept of Genetics, Genomics and Informatics, UTHSC, Dept of Pediatrics, Genetics Division, UTHSC

Human migration leads to the contact of populations that have been previously separated, often exposing them to different environments. The gene flow resulting from migration is an opportunity for beneficial alleles, previously adapted in a source population, to be introduced into the derived admixed population, leading to adaptation. Since the last centuries, America has been a target for numerous migration events, leading to modern admixed populations. In this study, we analysed ten thousand participants from the Biorepository and Integrative Genomics (BIG) initiative, that enrolled over 35k participants with electronic health records.Admixture deconvolution from genetic data indicates that 50% of the participants have non-European origins; this includes 20% African and 30% with admixed origins, with two-, three-way, and multiple admixture patterns. The most frequent two-way admixture (23%) includes that of Europeans with Africans, Indigenous from the Americans, or East Asians. Leveraging the lengths of ancestry tracts, we detected a strong signal of positive selection (six standard deviations above the mean, measured by iDAT) in the Indigenous-European admixed group. This signal is

located in a 10Mb region on chromosome 1 that includes the DARC locus, a well-known example of human adaptation to malaria, among other factors. We used approximate Bayesian computation to demonstrate that this signal is compatible with a demographic scenario that implies recent admixture (12-13 generations ago) and natural selection (s=0.02-0.06). Our results highlight the importance of further studying the impact of recent gene flows in America.

439 - FUNCTIONAL GENOMIC INSIGHTS INTO THE EVOLUTION OF THE ARSENIC TOLERANCE IN ANDEAN POPULATIONS

Mario Andres Apata

Mario A Apata, Sofia Acosta, Wyatt Comin, Klevis Lasku, Bridget Diviak, Ricardo Verdugo, Mauricio Moraga, Anne Stone, Jeanne Wilson-Rawls, Melissa Wilson

Evolutionary Biology Graduate Program, School of Life Sciences, Arizona State University, Tempe, USA., Center for Evolution and Medicine, School of Life Sciences, Arizona State University, Tempe, USA, Institute of Human Origins, School of Human Evolution and Social Change, Arizona State University, Tempe, USA, Molecular and Cellular Biology Graduate Program, School of Life Sciences, Arizona State University, Tempe, USA, School of Life Sciences, Arizona State University, Tempe, USA, School of Life Sciences, Arizona State University, Tempe, USA, School of Life Sciences, Arizona State University, Tempe, USA, Molecular and Cellular Biology Graduate Program, School of Life Sciences, Arizona State University, Tempe, USA, School of Life Sciences, Arizona State University, Tempe, USA, Instituto de InvestigaciÃ³n Interdisciplinaria, Universidad de Talca, Chile., Programa de GenÃ©tica Humana, Facultad de Medicina, Universidad de Chile. Santiago de Chile., Programa de GenÃ©tica Humana, Facultad de Medicina, Universidad de Chile. Santiago de Chile., Departamento de AntropologÃ-a, Facultad de Ciencias Sociales, Universidad de Chile., School of Human Evolution and Social Change, Arizona State University, Tempe, USA, School of Life Sciences, Arizona State University, Tempe, USA, Center for Evolution and Medicine, School of Life Sciences, Arizona State University, Tempe, USA, Institute of Human Origins, School of Human Evolution and Social Change, Arizona State University, Tempe, USA, School of Life Sciences, Arizona State University, Tempe, USA, Center for Evolution and Medicine, School of Life Sciences, Arizona State University, Tempe, USA, School of Life Sciences, Arizona State University, Tempe, USA

Arsenic causes severe toxicological effects in humans, yet some human populations have evolved tolerance. Many Andean people including those in the Camarones valley in northern Chile, as well as others in northwest Argentina, have been exposed environmentally to arsenic in their food and water supply over thousands of years. These populations carry conserved single nucleotide mutations in noncoding regions of the gene for a key enzyme known as the Arsenic (+3 oxidation state) Methyltransferase or AS3MT. AS3MT metabolizes arsenic in the liver; however, these Andean specific mutations have an unknown impact on the expression of this gene. Therefore, our aim was to evaluate experimentally whether some of these mutations might alter AS3MT gene expression in the liver. To address this, we selected four candidate mutations based on (1) whether they have been evolving under positive selection, (2) population genomic analyses comparing the Camarones and other indigenous populations from South America to characterize the allele frequency of these mutations, and (3) the intersection with potential regulatory regions. Then, we cloned small genomic DNA fragments carrying the Andean and the reference alleles (European) into a luciferase reporter plasmid. HepG2 cells, a human hepatocyte tumor cell line, were transfected with the different Luciferase reporter clones and treated with different arsenic concentrations. We observed significant differences between the Andean and European alleles when cells were exposed to arsenic. Our work contributes to an integrated perspective between evolutionary genomics and molecular genetics to understand how Andean people evolved in an arsenic-rich environment.

452 - Ancient genomes from the northern Southern Cone document thousands of years of continuity of a deep South American lineage

Javier Maravall López

Javier Maravall-LÃ³pez, Josefina M. B. Motti, NicolÃ¡s Pastor, Rodrigo Nores, David Reich
Harvard University, NEIPHPA-CONICET, Facultad de Ciencias Sociales, Universidad Nacional del Centro de la Provincia de Buenos Aires, 7631, Universidad Nacional de RÃ-o Cuarto, Instituto de AntropologÃ-a de CÃ³rdoba (IDACOR), CONICET, Universidad Nacional de CÃ³rdoba, 5000, Departamento de AntropologÃ-a, Facultad de FilosofÃ-a y Humanidades, Universidad Nacional de CÃ³rdoba, 5000, , Instituto de AntropologÃ-a de CÃ³rdoba (IDACOR), CONICET, Universidad Nacional de CÃ³rdoba, 5000, Department of Genetics, Harvard Medical School, Howard Hughes Medical Institute, Harvard Medical School, Broad Institute of Harvard and MIT, Department of Human Evolutionary Biology, Harvard University

The northern Southern Cone of South America is underrepresented in modern and ancient genomic surveys. To address this gap, we assembled new genome-wide data from more than 250 ancient individuals from this world-region, spanning a time transect up to 8000 years before present. We conclude that most of the ancient individuals we report in this study are descended from a deep lineage strongly represented in the central region of present-day Argentina that was already defined by 7600 years BP and persisted as the main ancestry component of these populations for thousands of years after. We also find evidence of substructure, significant differences in community sizes and mating patterns, and comparatively small subsequent contributions of other sources, including one related to the present-day Arhuaco people and one associated with ancient California-Channel Islands individuals, replicating a similar signal previously found in Central Andes ancient individuals. Apart from the main Central Argentina-associated component, individuals from the northwesternmost region of the country carried a large proportion of Central Andes-related ancestry, and individuals from the Gran Chaco region could be modeled as being entirely descended from a source related to the present-day WichÃ- people. Our findings highlight the remarkable genetic continuity of this understudied world-region, while providing fine-grained insights into networks of contacts of ancient groups within it. We present this work on behalf of a wider group of scholars who collected, generated, and analyzed material for this study.

622 - Uncovering South American genomic diversity in the Southern Cone

Epifanía Arango Isaza

EpifanÃ-a Arango-Isaza, Patrick F. Reilly, Marcos AraÃºjo Castro e Silva, TÃ¡bita HÃ¼nemeier, Felipe I. MartÃ-nez, Serena Tucci, Kentaro K. Shimizu, Chiara Barbieri
University of Zurich, Yale University, Universitat Pompeu Fabra, Universitat Pompeu Fabra, Pontificia Universidad CatÃ³lica de Chile, Yale University, University of Zurich, University of Cagliari

South Americaâ€™s rich cultural and historical diversity contrasts with its remarkably reduced genetic diversity, caused by a recent bottleneck that occurred during the initial peopling. The extent to which this lack of genetic variation limits the power of population genetic analysis is debated. However, available genetic data from Native American ancestries in South America are mostly restricted to uniparental markers and SNP chips, the latter often designed with inadequate representation of target SNPs from the continent. South Americaâ€™s genomic diversity, therefore, still lies largely unexplored, leaving gaps in our understanding of South American demographic history.
To fill this gap we generated 20 high-coverage genomes from populations in Central-Southern Chile, focusing on the Southern Cone of South America, and analyzed them with other high-quality ancient and modern genomes. We observed that population profiles inferred from whole genomes differ from directly comparable SNP chip data, highlighting the ascertainment bias intrinsic in the SNP chip. High within-population genetic diversity in the Southern Cone is indicative of a sustained larger population size in contrast to Amazonia, where European colonization has exerted a more pronounced impact on population decline. We also detect a minor pulse of gene flow from ancient migration sources which appears specifically in the Native American ancestry of Central-Southern Chile. Our study contributes to

the growing number of genomes from South America, uncovering previously understudied genomic diversity and connecting present-day indigenous ancestries with the ancient population dynamics of the region's first migrants.

652 - Ancient Mitogenomes from Belize Provide Insights into the Mobility and Migration of Ancient Communities in Mesoamerica and the Caribbean.

Celia D Cleary

Celia D Cleary, Angelina J Locker, Lauren C Springs, Diane Z Chase, Arlen F Chase, Adrian SZ Chase, Melissa M Badillo, Adela P Vallejos, Genara Cano, Roy Rodriguez, Rick WA Smith, Austin W Reynolds

Department of Anthropology, Baylor University, Department of Anthropology, Vanderbilt University, Department of Integrative Biology, University of Texas at Austin, University of Houston, Department of Comparative Cultural Studies, University of Houston, Mansueto Institute for Urban Innovation and Department of Anthropology, University of Chicago, Institute of Archaeology, Belize, To'one Masehualoon NGO, To'one Masehualoon NGO, To'one Masehualoon NGO, Department of Sociology and Anthropology, George Mason University, Department of Women and Gender Studies, Indigenous Science, Technology, and Society Lab, Faculty of Native Studies, University of Alberta, Department of Anthropology, Baylor University

Genomic research regarding paleodemography in Mesoamerica to date has often been conducted as part of large-scale studies of North and South America. Poor preservation of ancient DNA in the neotropics, among other difficulties, has limited ancient genomic studies of mobility in this region, with paleogenomic-scale studies of population history of Mesoamerica becoming a sustained focus in the past 5 years. Here, we present mitogenomic data of seven ancient Maya Ancestors recovered from interments in the ancient Maya city of Caracol, located in the Maya Mountains of present-day Belize, and 54 contemporary individuals from Corozal, in Northern Belize. These mitogenomic data were combined with 63 previously published ancient mitogenomes recovered from sites spanning Mesoamerica and the Caribbean to reconstruct paleodemographic and mobility patterns of maternal lineages through time. This work was conducted with authorization from the Belize Institute of Archaeology in and collaboration with Maya partners in Corozal. Haplogroup frequencies and median joining haplogroup networks were generated to understand the sharing of haplotypes across time and space in this region. Bayesian fixed molecular clock models were used to estimate matrilineal phylogenetic relationships between community members included in this study. Bayesian Skyline Plots were generated to infer dynamics in matrilineal effective population size. Models resulting from these 124 mitogenomes are used here to reconstruct paleodemographic events temporally and geospatially to gain a more robust understanding of the demographic history of the region.

672 - Insights from paleogenomic analyses of Michoacan burials

Xavier Roca-Rada

Xavier Roca-Rada, Kalina Kassadjikova, Fanny Mendisco, Marie-HÃ©lÃ¨ne Pemonge3, Yassine Souilmi, Marie-France Deguilloux, Lars Fehren-Schmitz, GrÃ©gory Pereira, Bastien Llamas

Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide, SA, Australia., UCSC Paleogenomics, Department of Anthropology, University of California, Santa Cruz, California, USA., UniversitÃ© de Bordeaux, CNRS, URM PACEA De la PrÃ©histoire Ã lâ€™Actuel: Culture, Environnement et Anthropologie, France., UniversitÃ© de Bordeaux, CNRS, URM PACEA De la PrÃ©histoire Ã lâ€™Actuel: Culture, Environnement et Anthropologie, France., Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide, SA, Australia., Indigenous Genomics Research Group, Telethon Kids Institute, Adelaide, SA, Australia., UniversitÃ© de Bordeaux, CNRS, URM PACEA De la PrÃ©histoire Ã lâ€™Actuel: Culture, Environnement et Anthropologie, France., UCSC Paleogenomics, Department of Anthropology, University of California, Santa Cruz, California, USA., Centre National de la

Recherche Scientifique, UMR8096 ArchAm Archaeology of the Americas, France., Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide, SA, Australia., Indigenous Genomics Research Group, Telethon Kids Institute, Adelaide, SA, Australia., Centre of Excellence for Australian Biodiversity and Heritage, University of Adelaide, Adelaide, SA, Australia., National Centre for Indigenous Genomics, John Curtin School of Medical Research, Australian National University, Canberra, ACT, Australia.

Mesoamerica stands as a historically and culturally rich geographic area (encompassing central and southern Mexico, Belize, Guatemala, El Salvador, and neighbouring areas), characterised by a mosaic of interconnected cultures, centuries of trading and agricultural practices. Despite its significance, the ancient genetic diversity of this region has remained largely understudied, with much of its complexity potentially lost or obscured by the Spanish invasion and colonial rule. Here, we analysed 11 newly sequenced ancient genome-wide datasets obtained from two archaeological sites in the Michoacan State dated to the 7th and 14th centuries CE. Preliminary analysis of mitochondrial haplogroup frequencies provides insights into a possible turnover of maternal genetic lineages between the Epiclassic and Middle Postclassic periods, potentially linked to the arrival of UacÃºsechas and Chichimecas from northern Mexico. However, contrasting signals emerge when examining nuclear data, indicating a possible genetic continuity over time. Our study underscores the need for additional analyses and expanded sampling efforts across Mesoamerica to discern population structure, adaptation, and detailed demographic trajectories essential for a comprehensive understanding of Indigenous populations history in the region.

676 - Local adaptive retention of high amylase copy number haplotypes in Peruvian potato farmers at historical time scale

Kendra Scheer

Kendra Scheer, Luane Landau, Kelsey Jorgensen, Charikleia Karageorgiou, Lindsey Siao, Angelis Morales-Rivera, Tom Brutsaert, Abigail Bigham, Omer Gokcumen

University at Buffalo, University at Buffalo, University of California, Los Angeles, University at Buffalo, University at Buffalo, Universidad de Puerto Rico, Recinto de RÃ-o Piedras, Syracuse University, University of California, Los Angeles, University at Buffalo

Starch digestion has a profound but variable impact on the evolution of human populations. The amylase gene (AMY1), which encodes for the namesake starch-digesting enzyme, shows remarkable copy number variation among humans, linked with metabolic diseases, microbiome, and adaptation to starch-rich diets. Recent studies have resolved the complex structural haplotypes and hypothesized that haplotypes harboring increased AMY1 gene copy numbers are increased in allele frequency due to recent soft sweeps in agricultural populations. To test this hypothesis, we utilized multiple bioinformatic approaches and digital PCR to show that Pima from northern Mexico and Quechua from Peruvian Andes harbor the highest number of AMY1 gene copies among world populations. To investigate the genetic variation in Andeans we sequenced 10 Peruvian individuals using Oxford Nanopore long-read sequencing to obtain 20 amylase haplotypes. Population genetic analyses suggest that the increased amylase gene copy number variation is a genome-wide outlier between Peruvian and neighboring populations (top 0.0004761653 percent), suggesting recent local selection. Previous work has shown that Peruvians relied heavily on starch-rich potatoes for their caloric needs following their domestication in the Andes around 6000-10000 BP. Prior work has tied this reliance to selection at the MGAM locus, another starch-digesting enzyme, in both ancient and modern Andeans. Thus, we hypothesize this strong dependency on potatoes has affected both enzymes in the starch digestion pathway in Peruvian populations. Collectively, our work illustrates how local dietary traditions can facilitate the adaptive retention of specific structural variations at historical timescales.

706 - Genetic Variants Associated with Skin Pigmentation in a Mayan Population of the Yucatan Peninsula

Paola Everardo

Paola Everardo, Juan Carlos Guevara, Alicia Portela-EstÃ©vez, Georgios Athanasiadis, Antonio GonzÃ¡lez-MartÃn
Universidad AutÃ³noma de la Ciudad de MÃ©xico, Universidad Pompeu Fabra, Universidad Complutense de Madrid, Universitat de Barcelona, Universidad Complutense de Madrid

Skin pigmentation is a complex trait influenced by genetic, environmental, and evolutionary factors. The Mayan population of the Yucatan Peninsula exhibits a diverse range of skin tones, reflecting their unique genetic ancestry and environmental exposures. Understanding the genetic variants underlying skin pigmentation in this population can provide insights into the evolutionary history and adaptation to local environmental conditions.
In this study, we conducted a comprehensive genetic analysis in two Mayan communities of the Yucatan Peninsula to establish the degree of association with pigmentation phenotypes of six variants previously reported in the literature (HERC2-rs12913832*A, OCA2-rs1800404*C, BNC2 - rs2153271*T, SLC45A2 - rs16891982*C, IRF4 - rs12203592*T and TYR - rs1042602*A). The ancestral composition of individuals was estimated using the Human Origins Axiom1 microarray.
Our analysis revealed a complex genetic structure in the Mayan individuals, associated with demographic dynamics of the Mayan population with Spanish and African populations during colonial times, involving pigmentation-related genes.
Overall, our study provides insights into the genetic basis of skin pigmentation variation in the Mayan population of the Yucatan Peninsula, highlighting the role of genetic ancestry. This research contributes to a deeper understanding of human diversity and adaptation in the context of indigenous populations and sheds light on the evolutionary history of pigmentation traits in the Americas. As well as in the use of categories that fail to reflect the reality of genetic diversity and evolutionary dynamics.

874 - Ancient Human DNA from the Osmore River Valley, Peru: Demographics, Disease, and Adaptation

Maxine E McCarty

Maxine McCarty, Adele Crane, Kelly Blevins, Allisen Dahlstedt, Kelly J Knudson, Jane Buikstra, C. Eduardo Amorim, Anne C Stone
School of Life Sciences, Arizona State University, Institute of Human Origins, Arizona State University, Center for Evolution and Medicine, Arizona State University, School of Life Sciences, Arizona State University, Institute of Human Origins, Arizona State University, Center for Evolution and Medicine, Arizona State University, Department of Archaeology, Durham University, School of Human Evolution and Social Change, Arizona State University, Center for Bioarchaeological Research, Arizona State University, School of Human Evolution and Social Change, Arizona State University, Center for Bioarchaeological Research, Arizona State University, School of Human Evolution and Social Change, Arizona State University, Center for Bioarchaeological Research, Arizona State University, Department of Biology, California State University Northridge, School of Human Evolution and Social Change, Arizona State University, School of Life Sciences, Arizona State University, Institute of Human Origins, Arizona State University, Center for Bioarchaeological Research, Arizona State University, Center for Evolution and Medicine, Arizona State University

Here we present preliminary ancient human DNA data from 102 individuals dating from 385 BCE to 1476 CE and spanning the Huaracane, Tiwanaku, Tumilaca, and Chiribaya cultures in the Osmore Drainage region of southern Peru with the goal of better understanding changes in population history and the spread of disease over time. Individuals with skeletal lesions specific to tuberculosis were sampled across multiple elements including teeth, vertebral elements, and ribs. We extracted 122 total subsamples, of which 98 were built into single stranded libraries that were shotgun sequenced on the Illumina NovaSeq X Plus to 4-50 million reads. The shotgun data were mapped to a human reference genome using NF-Coreâ€™s EAGER2 pipeline. We recovered low coverage mitochondrial genomes from 20 samples (0.1X-3.7X) and low full genome coverage from 8 teeth, all of which are candidates for further targeted enrichment. Using the

recovered host DNA, we aim to explore regional demographic history, familial relationships, the spread of tuberculosis spatial-temporally, and possible evidence of adaptation.

Funding: National Science Foundation, BCS-2217953

933 - Exploring Population History in Atlantic Patagonia: Ancient DNA Insights

Patricio Pezo

Patricio Pezo, Rosa Fregel, Beth Shapiro, Carlos Bustamante, MarÃ-a Laura Parolin, TÃ¡bita HÃ¼nemeier
Instituto de BiologÃ-a Evolutiva, CSIC-UPF, Departamento de BioquÃ-mica, MicrobiologÃ-a, BiologÃ-a Celular y GenÃ©tica. Universidad de la Laguna, Paleogenomics Laboratory, University of California Santa Cruz, Department of Genetics, Stanford University, Instituto de Diversidad y EvoluciÃ³n Austral, Instituto de BiologÃ-a Evolutiva, CSIC-UPF

The Patagonia region, located at the southernmost tip of South America, represents an endpoint in the global dispersal of Homo sapiens since the out-of-Africa migration. Despite considerable advancements in our understanding of migration patterns throughout South America, unresolved debates persist concerning the population history of Patagonia and subsequent migration or admixture events. This study aims to untangle the historical processes underpinning the colonization of Patagonia while shedding light on subsequent interactions with other Native American populations. Our approach involved a comprehensive analysis of nine whole-genome ancient DNA (aDNA) samples from northeastern Patagonia, spanning between 2,000 to 310 years before the present (YBP). As an initial exploratory step, we contextualized our population alongside other ancient individuals from America to delineate the evolutionary history of these populations. Our findings reveal a robust affinity between our samples to archaeological sites associated with southern Patagonian populations, such as the Selknam, Aonikenk, and Yamana. This affinity is particularly pronounced in early northeastern samples and gradually diminishing in later samples, which exhibit closer associations with geographically proximate archaeological sites like Laguna Toro, Arroyo Seco, and Laguna Chica. These preliminary results underscore a close genetic relationship between northeastern and southern Patagonia, hinting at recent shared history. The observed trend of decreasing affinity in recent samples suggests potential isolation within the hunter-gatherers from the Atlantic Patagonia, perhaps in response to diminishing connections over time. These initial insights substantially enrich our understanding of the evolutionary narrative of the Patagonian population in Argentina.

1030 - The MUC19 gene in Denisovans, Neanderthals, and Modern Humans: An Evolutionary History of Recurrent Introgression and Natural Selection

David Peede

David Peede, Fernando A. Villanea, Eli J. Kaufman, Valeria AÃ±orve-Garibay, Kelsey E. Witt, Viridiana Villa-Islas, Roberta Zeloni, Davide Marnetto, Priya Moorjani, Flora Jay, Paul N. Valdmanis, MarÃ-a C. Ã•vila-Arcos, Emilia Huerta-SÃ¡nchez
Department of Ecology, Evolution, and Organismal Biology, Brown University, Center for Computational Molecular Biology, Brown University, Institute at Brown for Environment and Society, Brown University, Department of Anthropology, University of Colorado Boulder, Division of Medical Genetics, Department of Medicine, University of Washington School of Medicine, Licenciatura en Ciencias GenÃ³micas, Escuela Nacional de Estudios Superiores Unidad Juriquilla, Universidad Nacional AutÃ³noma de MÃ©xico, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico, Center for Computational Molecular Biology, Brown University, Department of Genetics and Biochemistry, Clemson University, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico, Department of Neurosciences â€œRita Levi Montalciniâ€•, University of Turin, Department of Neurosciences â€œRita Levi Montalciniâ€•, University of Turin, Department of Molecular

and Cell Biology, University of California, Berkeley, Center for Computational Biology, University of California, Berkeley, Laboratoire Interdisciplinaire des Sciences du NumÃ©rique, UniversitÃ© Paris-Saclay, Division of Medical Genetics, Department of Medicine, University of Washington School of Medicine, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico, Department of Ecology, Evolution, and Organismal Biology, Brown University, Center for Computational Molecular Biology, Brown University

All humans carry a small fraction of archaic ancestry across the genome, the genetic legacy of Neanderthals and Denisovans. While the effects of Neanderthal ancestry on human fitness have been thoroughly explored, there are fewer examples of adaptive introgression from Denisovans. Here, we propose a candidate gene for adaptive introgression, MUC19, for which some modern humans carry the Denisovan haplotype. MUC19 encodes a member of the gel-forming mucin protein family related to lacrimal and salivary gland function in humans and has been previously proposed as a top candidate for positive selection in Indigenous American populations. We find the Denisovan-like haplotype of MUC19 at its highest frequency in admixed American populations. Interestingly, the Denisovan harbors nine missense mutations in this region, all of which are fixed in modern-day individuals harboring the introgressed Denisovan-like haplotype. Furthermore, after conducting formal tests of selection, we find that the Denisovan-like haplotype of MUC19 is responsible for the signals of positive selection in admixed American populations. Lastly, we find that late Neanderthals also carry the Denisovan-like haplotype in the heterozygous state, while the earlier Altai Neanderthal is more diverged than expected from the Denisovan in this region. Our results suggest a complex pattern of multiple introgression events throughout human history at MUC19, first from Denisovans into late Neanderthals, followed by Neanderthal introgression into modern humans, and then selection for the introgressed haplotype in the Americas, which may have played a unique role in the evolutionary history of Indigenous American populations.

1043 - Unraveling transpacific immigrations to Peru and the Philippines between the mid-16th and early 20th centuries

Zara Paulina Martínez-Sánchez

Zara P MartÃ-nez-SÃ¡nchez, Carmina Barberena-Jonas, Santiago G Medina-MuÃ±oz, Consuelo D Quinto-CortÃ©s, Alexander J Mentzer, Alexander G Ioannidis, AndrÃ©s Moreno-Estrada, Mashaal Sohail
Centro de Ciencias GenÃ³micas (CCG), Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), Laboratorio Nacional de GenÃ³mica para la Biodiversidad (UGA-LANGEBIO), CINVESTAV, Laboratorio Nacional de GenÃ³mica para la Biodiversidad (UGA-LANGEBIO), CINVESTAV, Laboratorio Nacional de GenÃ³mica para la Biodiversidad (UGA-LANGEBIO), CINVESTAV, The Wellcome Centre for Human Genetics, University of Oxford, Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Department of Biomedical Data Science, Stanford University, Laboratorio Nacional de GenÃ³mica para la Biodiversidad (UGA-LANGEBIO), CINVESTAV, Centro de Ciencias GenÃ³micas (CCG), Universidad Nacional AutÃ³noma de MÃ©xico (UNAM)

The migratory stories of people who crossed the Pacific, thanks to maritime routes that connected Latin America and Asia since the mid-16th century with the Manila Galleon have been overlooked despite often carrying a narrative of discrimination and even slavery. This study dives into the history of some of these migrations by combining historical and computational methods for ancestry inference such as Admixture and Gnomix, MAAS-MDS to marcate specific origins, and Tracts to time migrations, using the 1000 Genomes Project, Human Genome Diversity Project, Oceania Genome Variation Project and a Philippine dataset. Particularly, we focus on migrations from East Asia to Peru, estimating two pulses of migration that occurred in the mid-19th and early 20th centuries that are consistent with the records of migrations from China and Japan to Peru, in addition to a third migration pulse estimated to occur in the early 18th century. We also map Asian ancestries in these individuals to China, Japan and Myanmar. Focusing on the west side of the Pacific, we also study migrations from Latin America to the Philippines based on historical information of transpacific migrations through the Manila Galleon

during the Colonial Period (Mawson, 2016), finding the presence of American Indigenous ancestry in varying proportions among individuals from different regions of the Philippines, larger than those observed in other East Asian cohorts. This way, our work investigates and makes visible a piece of the migratory history that occurred between Latin America and Asia in the last 500 years.

1103 - Reconstructing Western and Northern Mexico's Past

Roslyn Curry

Roslyn Curry, Jakob Sedig, Jose Luis Ponzo

Harvard University, Harvard University, n/a, Instituto Nacional de AntropologÃ-a e Historia

Though ancient DNA has revolutionized our understanding of the past, America/Mexico remain understudied. The Proyecto de investigaciÃ³n de poblaciones antiguas en el norte y occidente de MÃ©xico (PIPANOM) addressed this imbalance through the examination of over 400 ancient individuals from central, western and northern Mexico. Data from these individuals have revealed population structure that is broadly associated with geographic distance across what is now modern Mexico, Belize, Guatemala, and the southwest United States. These data have also provided new insights into long-standing questions about migration and interaction of different archaeologically defined cultures in key eras of Mexicoâ€™s past, such as the movement of people during the Postclassic period. Along providing new insights in an area that has been understudied with aDNA, PIPANOM has helped correct institutional imbalances in archaeological and paleogenetic research. PIPANOM is an archaeologist-led project that has brought together more than 20 collaborators. As part of the project, there have been numerous workshops, trainings, and visits bringing together Mexican and American archaeologists, geneticists, researchers, analysts, and students. As such, PIPANOM provides a roadmap for equitable large-scale projects that produce robust scientific results.

1110 - Evidence for genetic continuity, contacts, and changing social organization in the Central Andes from ancient DNA

Alison R Barton

Alison R Barton, Weston C McCool, Valda G Black, Lars Fehrenâ€‘Schmitz, Jennifer L Kennedy, David E Reich

Harvard University, University of Utah, Washington State University, University of California Santa Cruz, Binghamton University, Harvard University, Harvard Medical School, Howard Hughes Medical Institute

Previous work with populations from the Central Andes has suggested a high level of genetic continuity in the region but has also highlighted evidence of movement of peoples and changing kinship structures. In this work, we examine 315 new samples from ancient individuals from modern Peru and Bolivia dating from 8600-200 BP. We find similar evidence of genetic continuity in the Andes based primarily on North-South geography. However, identity-by-descent (IBD) analysis supports the hypothesis of genetic contacts between these groups locally and in a few cases with peoples in what is today Colombia and Ecuador, with at least 3 individuals sharing greater than 50 cumulative cM of IBD segments longer than 8 cM with individuals from these areas. Moreover, we find additional support for the claim that consanguineous pairings increased during the Late Intermediate Period (LIP, 1000-1476 CE) likely due to changing social structures centered around the ayllu system that were reported historically. We observe in increase (Fisherâ€™s exact test, p= 0.00061) from 17% (4/23 samples) to 65% (90/161) during the LIP that subsequently drops (Fisherâ€™s exact-test, p = 1.3E-07) after this period to 9% (3/35). We also present the first large pedigrees from this time period from 3 valleys in the Nazca highlands representing samples from 10 different chullpas. Five of these tombs with the densest sampling suggest burial practices were strongly tied to biological kinship and illustrate family units in which we appear to see this social structure in place.

149 - Disentangling the Evolution of Grape Crop Wild Relatives

Christopher J Fiscus

Christopher J Fiscus, JonÃ¡s A Aguirre-Liguori, Jacob B Landis, Garren RJ Gaut, Abraham Morales-Cruz, Brandon S Gaut
University of California, Irvine, Universidad AutÃ³noma de YucatÃ¡n, Cornell University, University of California, Irvine, Lawrence Berkeley National Laboratory, University of California, Irvine

The Vitis genus comprises ~80 species distributed throughout North America and Eastern Asia that thrive in a variety of environments from lush forests to dry deserts. The rapid adaptive radiation of this genus makes it an ideal system to examine the genetic and environmental factors contributing to species divergence. Understanding adaptation in wild Vitis also benefits the breeding and improvement of the crop species V. vinifera, as wild species possess alleles pre-adapted to abiotic and biotic stress that can be utilized agronomically. To study the evolution of Vitis, we cataloged genetic variation in over 550 individuals from 56 species, with special emphasis on species endemic to North America. We first used this dataset to clarify species relationships, which have been obscured due to rampant introgression and hybridization between species. We detected several clades of hybrid origin, suggesting complex speciation histories. Next, we identified genotype-environment associations (GEA), identifying alleles associated with either warm/wet or cool/dry climates, and we examined the evolutionary history of these alleles through recent evolutionary time. Finally, we leveraged our dense sampling of V. arizonica, a species endemic to the American Southwest, to model the relationship between climatic and genetic variation using novel machine learning algorithms that were modified to account for collinear predictors, which otherwise distort predictor importance. Our work contributes to ongoing efforts to decipher the genetic basis of environmental adaptation and supports initiatives aimed at bolstering the sustainability of essential food crops through harnessing the potential of crop wild relatives.

154 - Genetic diversity of Zea diploperennis and strategies for its conservation

José Carlos Moreno Juárez

JosÃ© Carlos Moreno JuÃ¡rez, Alicia Mastretta-Yanes, Ana Laura Wegier, Diana MarÃ-a Rivera, JosÃ© de JesÃºs SÃ¡nchez GonzÃ¡lez
JardÃ-n BotÃ¡nico, Instituto de BiologÃ-a, UNAM, ComisiÃ³n Nacional para el Conocimiento y Uso de la Biodiversidad, JardÃ-n BotÃ¡nico, Instituto de BiologÃ-a, UNAM, Departamento de Ciencias BÃ¡sicas, TecnolÃ³gico Nacional de MÃ©xico, Campus Tlajomulco, Centro Universitario de Ciencias BiolÃ³gicas y Agropecuarias, Universidad de Guadalajara

Zea diploperennis, a perennial teosinte species and wild relative of maize, exhibits a distinctive range and is characterized by unique agricultural traits. These traits hold substantial potential for future maize breeding programs. Additionally, it holds cultural significance for traditional farmers in the Sierra de ManantlÃ¡n region. However, Z. diploperennis populations face threats from land-use changes and the decline of traditional agriculture, closely linked to its existence. Consequently, a comprehensive assessment of the conservation status of its genetic diversity is imperative. This evaluation, aligned with the Kunming-Montreal Global Biodiversity Framework ratified in 2022, utilized Genotyping-by-Sequencing (GBS) data from all known Z. diploperennis localities. Genetic clustering analysis revealed two distinct clusters corresponding to Jalisco and Nayarit localities, indicating significant genetic structure. Current and historical effective population sizes were estimated, revealing both clusters with effective population sizes below 500 and a recent decline in effective sizes. This suggests a present low genetic diversity, further diminished in recent generations. Urgent in situ conservation strategies are warranted for both clusters to mitigate potential inbreeding depression.

283 - Evolution history and the importance of genomic diversity facing climate change. The case of Agave marmorata Roezl., a microendemic agave used for mezcal production.

Karen Yazmin Ruiz Mondragon

Karen Yazmin Ruiz Mondragon, Luis Eguiarte
Instituto de EcologÃ-a, Instituto de EcologÃ-a

Genetic diversity holds the key to adaptation processes that are extremely important through climatic and environmental changes.
Agave marmorata is an endemic species which is distributed in the states of Oaxaca and Puebla, Mexico and is widely used to produce a mezcal. We assessed the genomic diversity and differentiation with RADseq method and 29,101 high quality SNPs, in A. marmorata, with four different types of management (cultivated, live fence, nursery and wild). We examined several demographic scenarios of the evolutionary history in this species, and possible projections into the future with various models of climate change.
We estimated high levels of genomic diversity (HS=0.229) and relatively high inbreeding (FIS=0.106, Fhat3=0.190). The cultivated samples harbored less genetic diversity than the wild plants. Furthermore, we estimated very low differentiation between cultivated and wild plants (FST=0.037).
We identified two main differentiated genetic groups, one in the East and another in the West of the distribution area. This genetic structure possibly derived from contractions during the Pleistocene (~216,879.75 years BP), and the formation of two refuges in small areas with climatic stability. We estimated a recent expansion of ~1,618 years BP, with larger effective population size (Ne=8,099). The future climate change models indicated negative effects on the distribution of A. marmorata, with a reduction in suitable habitat (from -0.87 to -1.08 in Current vs. 2041-2060 and Current vs.2061-2080, respectively). Finally, we proposed conservation measures for the different types of management, considering the biotic and abiotic interactions of Agave marmorata.

296 - The utility of landrace environmental data for climate-adaptive maize breeding

Forrest Li

Forrest Li, Daniel J Gates, Edward S Buckler, Sarah J Hearne, Jeffrey Ross-Ibarra, Daniel E Runcie
Department of Evolution and Ecology and Department of Plant Sciences, University of California Davis, Department of Plant Sciences, University of California Davis, USDA-ARS, Robert W. Holley Center for Agriculture and Health, Ithaca, NY and Institute for Genomic Diversity, Cornell University, International Maize and Wheat Improvement Center (CIMMYT), Department of Evolution and Ecology, Center for Population Biology, and University of California Davis, Genome Center, University of California Davis, Department of Plant Sciences, University of California Davis

Maintaining crop production yields in the face of climate change is a major challenge facing plant breeding today. Considerable adaptive genetic variation to improve breeding programs exists in ex situ germplasm collections of traditional landraces, but effective utilization of this diversity requires identifying adaptive alleles or testing the agronomic performance of varieties in different environments.
To evaluate approaches for identifying environmentally adaptive alleles, we collected climatic variables from the environment-of-origin of each of nearly 4000 previously genotyped traditional maize landraces representing a breadth of maize diversity. We evaluated three different approaches of using genetic and climate data to identify useful diversity based on phenotypic predictive ability in field trials. First, we performed gene-environment association (GEA) studies identifying genetic loci associated with environmental heterogeneity, such as hsftf9, a putative heat shock protein that may provide adaptive benefit in high temperatures, as well as the locally adaptive large scale inversion Inv4m, previously characterized as locally adaptive to highland conditions. Second, we used existing genotype data to predict optimal climate niches that may inform selection of landraces better adapted to target environments as well as demonstrate the limitations of prediction of environmental variables across spatial ranges. Third, we tested the ability of genomic prediction using population structure and high-resolution SNP data for landrace collections.
We find that genomic prediction capturing population structure and genetic relationship continues to have the highest predictive ability compared to models employing climate-of-origin variables or models focused on large-effect loci associated with climate.

302 - Traditional crops as reservoirs of genetic diversity: the case of pumpkin Cucurbita pepo ssp. pepo in Mexico

Carmina NA Martínez-González

Carmina MartÃ-nez-GonzÃ¡lez, Guillermo SÃ¡nchez-de la Vega, JosuÃ© Barrera-Redondo, Emanuel VillafÃ¡n de la Torre, CuauhtÃ©moc Jacques HernÃ¡ndez, Erika Aguirre-Planter, Rafael Lira-Saade, Luis E. Eguiarte

UBIPRO, Facultad de Estudios Superiores Iztacala, Universidad Nacional AutÃ³noma de MÃ©xico, UBIPRO, Facultad de Estudios Superiores Iztacala, Universidad Nacional AutÃ³noma de MÃ©xico, Max Planck Institute for Biology, Instituto de EcologÃ-a A.C. (INECOL), Red de Estudios Moleculares Avanzados (REMAv), Centro de BiotecnologÃ-a GenÃ³mica, Departamento de EcologÃ-a Evolutiva, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, UBIPRO, Facultad de Estudios Superiores Iztacala, Universidad Nacional AutÃ³noma de MÃ©xico, Departamento de EcologÃ-a Evolutiva, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico

Landrace crops (razas or variedades criollas in Spanish) are considered the second most important reservoir of genetic diversity for crop improvement, only after the wild crop relatives. Landrace crops are the product of centuries of selection by farmers and have locally evolved in response to both the environmental conditions of each region and different social and cultural aspects. In Mexico, the traditional landraces of pumpkins and squashes of C. pepo ssp. pepo were domesticated, cultivated and managed under the traditional cultivation system â€' the milpaâ€' starting at least 10,000 years ago, which has resulted in the development of high morphological and genetic diversity. We analyzed 8,643,037 SNPs obtained by whole genome sequencing at high coverage from 96 individuals of different landraces of C. pepo ssp. pepo representing all Mexico and Guatemala. We also included 13 individuals of the ancestor of this cultivated species, C. pepo ssp. fraterna, from the only three wild populations that we found. The traditional landrace varieties of Mexico exhibit a genetically differentiated group from their wild populations and have higher levels of genetic diversity (Cultivated Ï€= 0.155; Wild Ï€= 0.058), possibly due to the very small population sizes of the extant wild populations. We found a clear geographic differentiation among the landraces, forming four groups. We also detected 28 genes under selection. We conclude that in contrast to what is generally reported in the literature, the traditional landraces of C. pepo ssp. pepo are the most important reservoir of genetic diversity for this species.

367 - The improvement of shattering traits in key climate-resilient under-utilised crops using Crispr-Cap9 technology.

Amy C. Jackson

Amy C. Jackson, George Burton, Emma Wallington, Caspar Chater, Rafal Gutaker

RBG Kew, RBG Kew, NIAB, RBG Kew, University of Sheffield, RBG Kew

There are many examples of underutilized crops that display drought-resistant traits which could play a crucial role in improving food security, especially in regions with challenging agricultural environments. However, they often display a lack of domestication characteristics, such as non-shattering, which hinders their potential as a global sustainable food item for this changing world. With increasing efforts to publish genomic resources for under-utilised crops, genetic improvement of agronomic traits is increasingly possible. Here we focused on white fonio (Digitaria exilis), a small grain cereal crop native to West Africa and rice bean (Vigna umbellata), a legume cultivated across Asia. Through orthologous searches, we have identified homologous seed-shattering genes, SEED-SHATTERING1 (Sh1) and QUANTITATIVE SHATTERING1 (qSh1) in fonio and MYB26 gene in rice bean. Phylogenies and multiple sequence alignment were performed for each gene between closely related species to identify mutations, conserved regions, and gene structure. We aim to develop a reduced-shattering phenotype in both species using Crispr-Cas9-mediated genome editing technology. This poster highlights the utility of homolog search for accelerated domestication of crops with underutilized potential.

496 - Domestication gradient in Capsicum annuum chile pepper’s center of origin: implications under climate change scenarios drawn from niche models and genomic diversity.

Natalia Elena Martínez-Ainsworth

Natalia E MartÃ-nez-Ainsworth, Hannah Scheppler, Alejandra Moreno-Letelier, Vivian Bernau, Michael Kantar, Kristin Mercer, Lev JardÃ³n-Barbolla

National Autonomous University of Mexico (UNAM), Ohio State University, Ohio State University, National Autonomous University of Mexico (UNAM), Iowa State University, University of Hawai'i, Ohio State University, National Autonomous University of Mexico (UNAM)

Among crops domesticated in Mexico, chile pepper (Capsicum annuum) offers an ideal model to analyze the intersection of domestication status and management regimes over diverse environmental conditions and how these categories might face the challenges imposed by anthropogenic climate change. Here we generate a large C. annuum collection (n=1226) spanning four domestication categories throughout Mexico. In addition, we recovered coordinates from bibliography, public repositories and developed proxies to perform detailed ecological niche models for each category under present day conditions as well as eight future climate change scenarios, three time points and two emissions forecasts. Our main findings include an expansion to higher altitudes and retraction from dry-hot areas of the early domesticate landraces followed by a regaining of these lost areas by commercial varieties purportedly through irrigation. Environmental envelopes differed across domestication categories, notably for temperature seasonality and precipitation of the warmest quarter. There were several key geographic areas suitability was lost in both wild and landrace under future niche suitability projections this is a potential threat to geneflow between these groups which has helped maintain genetic diversity. We further obtained genome-wide genotyping (GBS) data on a sample of 61 wild, 45 semiwild, 164 landraces and 5 commercial samples. By comparing genomic diversity estimates and exploring environmental structuring of such diversity we have found that backyard and milpa management practices are a key potential source of climate adaptive traits and thus should be targeted by conservation efforts.

517 - Improving genome quality through artificial truncating purifying selection using heat shock: case of common carps (Cyprinus carpio L.)

Aleksandr Kuzmin

Aleksandr Kuzmin, Valeria Lobanova, Viktoria Skripskaya, Polina Brodovaya, Evgenii Vinogradov, Dmitrii Balashov, Nikolai Mugue, Konstantin Popadin

Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University, Branch for the freshwater fisheries of the Federal State Budget Scientific Institution "Russian Federal Research Institute of Fisheries and oceanography" Laboratory of genetics and fish, Branch for the freshwater fisheries of the Federal State Budget Scientific Institution "Russian Federal Research Institute of Fisheries and oceanography" Laboratory of genetics and fish, Federal State Budget Scientific Institution "Russian Federal Research Institute of Fisheries and oceanography" Laboratory of genetics and fish breeding (VNIRO), Immanuel Kant Baltic Federal University

The process of domestication is associated with a decrease in effective population size, which in turn leads to the accumulation of slightly deleterious variants (SDV) due to genetic drift. In the current study, we analyzed the case of common carp, Cyprinus carpio L.
We established and tested a heat shock-induced truncating purifying selection, which preferentially eliminates mutation-burdened embryos. Experimentally we observed that the fitness of mutation-burdened versus wild-type carp lines drops stronger after a heat shock of 38 degrees during 40 minutes within the pharyngula stage. Despite the heat shock reducing the burden of SDVs in the genome of carp, it can temporarily reduce the adaptability of the first generation. So in (i)the long term scenario, the selected replacement-brood stock will

consist of second-generation offspring (F2), avoiding the associated epigenetic impacts of temperature. (ii)The short term scenario of genome purification, which we call â€œHeat Shock Quality Assessmentâ€• - an experimental method to assess the genetic quality of breeder fish uses temperature shock on a sample of larvae, with survival rates indicating breeder quality. High survival rates in the shocked larvae indicate that the control group is suitable for stocking and distribution in fish farms.
This technology might be implemented into cultivating fish populations to restore and maintain the high quality of the gene pool, which will contribute to maintaining the genome quality of domesticated species of common carp at a high level.

626 - The distribution of self-incompatibility systems in cultivated angiosperms, facing global changes

Miriam Monserrat Ferrer

Miriam Monserrat Ferrer, Mirley Arlyn Verde-CÃ¡ceres, Sara Victoria Good
Universidad AutÃ³noma de YucatÃ¡n, Universidad AutÃ³noma de YucatÃ¡n, The University of Winnipeg, University of Manitoba

Theoretical and experimental studies have highlighted that angiosperms harbouring self-incompatibility (SI) systems are more susceptible to inbreeding depression when a transient breakdown in the SI system is achieved either due to selective forces (e.g. mate limitation) or environmental effects (e.g increased temperature). Domesticated plants are usually thought to be self-compatible, although, in a recent survey of 5686 hermaphroditic angiosperms, we found that 443 out of 907 cultivated species displayed self-incompatibility, partial self-compatibility, or self-sterility. Here, we summarize existing experimental work examining the effect of increased temperature on the breakdown of SI, as well as review the molecular and genetic basis of SI systems in cultivated species. Next, we present an analysis of the association of cultivated species with i) phylogenetic distribution, ii) presence of SI, self-compatibility (SC), partial self-compatibility (PSC) or self-sterility (SS); iii) growth form, iv) propagation method and v) use of the crop. We discuss how clonal propagation of individuals enables the maintenance of self-incompatible and partially self-compatible cultivars even for fruit production, but these maintaining fertility and SI lines present unique challenges not present for self-compatible cultivars. Finally, we contextualize this work within the ongoing dialogue about the particular vulnerabilities of self-incompatible cultivars in the context of global change, particularly the impact that increasing temperature.

744 - Genomic landscape of introgression in Indonesian cattle and its implication for their adaptive potential

Sabhrina Gita Aninta

Xi Wang, Casia Nursyifa, Laura Bertola, Genis Garcia-Erill, Sabhrina Gita Aninta, Anubhab Khan, Josiah Kuja, Kristian HanghÃ¸j, Jonas Meisner, Thomas BÃ¸gild, Corey Bradshaw, Alam Persada Putra, Sendi Dwi Priyono, Yuli A. Tribudi, Jiang Yu, Johannes Lenstra, Reagan Sims, Darren E. Hagen, Michael P. Heaton, Timothy P. L. Smith, Mikkel-Holger S. Sinding, Laurent Frantz, Greger Larson, Dedy Duryadi, Muhammad Agil, Bambang Purwantara, Rasmus Heller
Department of Biology, University of Copenhagen, Department of Biology, University of Copenhagen, Department of Biology, University of Copenhagen, Department of Biology, University of Copenhagen, Department of Biology, University of Copenhagen, Department of Biology, University of Copenhagen, Department of Biology, University of Copenhagen, Department of Biology, University of Copenhagen, Department of Biology, University of Copenhagen, Department of Biology, University of Copenhagen, College of Science and Engineering, , Laboratorium Genetika Konservasi Hewan, Institut Pertanian Bogor, Faculty of Biology, Universitas Gadjah Mada, Faculty of Agriculture, Universitas Tanjungpura, College of Animal Science and Technology, Northwest A & F University, Department Population Health Sciences, Utrecht University, Department of Agricultural Sciences, Texas State University, Animal & Food Sciences, Oklahoma State University, U.S.

Department of Agriculture, U.S. Department of Agriculture, Department of Biology, University of Copenhagen, Faculty of Veterinary Medicine, University of Munich, School of Archaeology, University of Oxford, Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Pertanian Bogor, Fakultas Kedokteran Hewan, Institut Pertanian Bogor, Fakultas Kedokteran Hewan, Institut Pertanian Bogor, Department of Biology, University of Copenhagen

Admixture between species is increasingly recognised to be prevalent in many species, especially within domesticated cattle breeds as local people aim to improve their genetic variation. In Indonesia, cattle breeds are generally known to be in high genetic diversity due to having an admixed ancestry from zebu (Bos indicus) and banteng (Bos javanicus). We performed whole genome sequencing of 233 bovids, representing local cattle from Aceh, Pesisir, Pasundan, Jabres, Madura, Sumba Ongole, Javan banteng, as well as Bali cattle in Indonesia and Australia. After comparing with other cattle where published genomic data is available, Indonesian cattle breeds are found to have the highest genetic diversity of any cattle populations worldwide. This is largely due to amounts of banteng introgression, reaching up to ~30% in the Madura breed. We also found regions with pronounced peaks and valleys of introgression, suggesting selection has been acting on the introgressed haplotypes. While the genomic landscape of introgressed ancestry was significantly correlated between breeds, extreme outliers were predominantly private to each breed, consistent with breed-specific artificial selection pressures. However, we found evidence that some phenotypes, such as coat colour, might experience convergent adaptive introgression in all breeds with banteng introgression. These findings unveil much of the unknown part of the evolutionary history of Asian cattle and processes occurring after introgression into a livestock species, in addition to discovering a rich source of previously unknown genetic variation in cattle that is relevant for genetic improvement.

799 - The domestication process of two Agave species used for mezcal production with different propagation methods and evolutionary histories.

Irene Martínez Velasco

Irene MartÃ-nez Velasco, Jaime Gasca Pineda, Anastasia Klimova, Karen Ruiz MondragÃ³n, CÃ©cile Truchot Taillefer, Guillermo SÃ¡nchez de la Vega, Erika Aguirre Planter, Rafael Lira, Alfonso Valiente Banuet, Luis E. Eguiarte Fruns
Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, Laboratorio de Recursos Naturales, Unidad de BiotecnologÃ-a y Prototipos (UBIPRO), Facultad de Estudios Superiores Iztacala, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico

In Mesoamerica, in situ management of wild plants is common and plants with different use to humans can exhibit varying degrees of domestication and be subject to diverse practices. In economically important species, such as agaves, direct extraction of wild plants along with different management practices alters the genetic composition of the populations. Currently, many Agave species are intensively propagated for mezcal production. Two of them, Agave karwinskii in the states of Puebla and Oaxaca, and Agave cupreata from Guerrero and MichoacÃ¡n are facing strong extraction in their wild populations and artificial selection processes. The prevailing propagation method of many Agave species-- including A. karwinksii-- involves underground rhizomes, facilitating the establishment of commercial plantations. However, species like A. cupreata do not produce basal shoots or bulbils, depending entirely on seed production, making plant nurseries indispensable for mass plant production. We analyzed populations of both species under different degrees of management, from wild to commercially exploit. We studied 300 individuals of A. karwinskii, and 153 of A. cupreata using genomic data

from GBS, with 12,888 and 15,880 SNPs, respectively, to obtain diversity statistics and describe the genetic differentiation patterns. We found a high expected heterozygosity in A. karwinskii (0.299) and A. cupreata (0.236), as well as a significant decrease in genetic diversity associated with the degree of agricultural management in both species. Our results indicate that, in addition to the propagation method, the type and direction of anthropogenic pressure are significant modifiers of their genomic biodiversity populations

827 - Comprehensive analysis of the microbial communities of traditional agave fermentations, a starting point for their conservation

Angelica Jara Servin

AngÃ©lica Jara-Servin, Luis D Alcaraz, Sabino I Juarez-Serrano, AarÃ³n Espinosa-Jaime, Ivan Barajas, Lucia Morales, Alexander DeLuna, Antonio HernÃ¡ndez-LÃ³pez, Eugenio Mancera
Faculty of Science, National Autonomous University of Mexico, Faculty of Science, National Autonomous University of Mexico, Department of Genetic Engineering, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, National School of Higher Education, National Autonomous University of Mexico , National School of Higher Education, National Autonomous University of Mexico , International Laboratory for Human Genome Research, National Autonomous University of Mexico , Unidad de GenÃ³mica Avanzada (Langebio), Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, National School of Higher Education, National Autonomous University of Mexico , Department of Genetic Engineering, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional

Agave spirits are traditional Mexican beverages with increasing international appreciation. They are produced throughout the country by an artisanal process that relies on the open fermentation of cooked agave juice from a variety of agave plant species. Most emphasis to preserve and manage the production of agave spirits has been placed on the plants and production practices, while the microbial communities involved in the fermentation remain largely unexplored. To bridge this gap of knowledge we performed 16S and ITS amplicon sequencing of almost 100 fermentation tanks from 42 different distilleries throughout Mexico. The distilleries encompass contrasting production practices, climatic conditions, and biogeographic characteristics. We identified thousands of OTUs clustered into hundreds of bacterial genera and fungal species. Lactic acid bacteria and ascomycetous yeasts were the most prevalent microorganisms, in agreement with previous reports, but most taxa had never been associated with this environment. Despite the differences in the distilleries, a core microbiome of 15 bacterial genera and 13 fungal species was identified. The only variable that was associated with the composition of both bacterial and fungal microbiome is the distillery, suggesting that microbial composition is determined by the local production practices and characteristics of each site. Our results represent the most comprehensive analysis of microbial communities of agave fermentations to date, contributing key knowledge for its management and conservation.

948 - Genomic studies on the domestication of chayote (Sechium edule ssp. edule), and its relationship with wild species (Sechium spp.).

Cécile Truchot Taillefer

CÃ©cile Truchot Taillefer, Irene MartÃ-nez Velasco, Carmina MartÃ-nez GonzÃ¡lez, Erika Aguirre Planter, Jaime Gasca Pineda, Guillermo Sanchez de la Vega , Rafael Lira Saade, Luis Eguiarte
Instituto de EcologÃ-a UNAM, Instituto de EcologÃ-a UNAM, Facultad de Estudios Superiores Iztacala UNAM, Instituto de EcologÃ-a UNAM, Instituto de EcologÃ-a UNAM, Facultad de Estudios Superiores Iztacala UNAM, Facultad de Estudios Superiores Iztacala UNAM, Instituto de EcologÃ-a UNAM

The cultivated chayote belongs to the genus Sechium P. Browne of the Cucurbitaceae family. It comprises 11 species divided into two

clades. The Mexican clade encompasses six taxa native to Mexico, including five wild taxa (S. chinantlense, S. compositum, S. hintonii, S. mexicanum, S. edule ssp. sylvestre), along with the cultivated chayote, S. edule ssp. edule. The remaining species of the genus belong to the Central American clade. Chayote cultivation, which has a history of at least 500 years, is believed to have been domesticated in Mexico or in Central America. Yet, the soft protective covering of the seed is subject to degradation, preventing archaeological evidence to determine the timing of chayote domestication with precision and its relationship to wild species. Previous studies using different methods yielded unclear results regarding the phylogenetic relationships within the genus. In this study, we disentangle the relationships within the Mexican clade through phylogenomic and coalescent-based species delimitation approaches, based on genome-wide single nucleotide polymorphisms (SNPs). Using a reduced representation method (GBS) and an already published chayote genome, we obtained 4,185 SNPs markers. Our analysis revealed current gene flow between cultivated and wild S. edule in the state of Veracruz. In addition, we found mixed ancestry in S. chinantlense and S. edule ssp. sylvestre, suggesting hybridization between them in the past. Our data support the idea that the center of origin of chayote is in the state of Oaxaca. Our study is useful for developing future conservation and management plans, including genetic improvement.

960 - Disentangling the genetic bases of adaptation to a warmer world

Jennifer Molinet

Jennifer Molinet, Rike Stelkens
Stockholm University, Stockholm University

Accelerating climate change underscores the urgency of comprehending evolutionary responses to heat stress. In this context, one of the most significant challenges facing organisms today is global warming. There is an urgent need to gain a better understanding of the long-term evolutionary effects of rising temperatures on biodiversity. The primary objective was to assess the capacity of populations to adapt to heat stress. For this purpose, we conducted a long-term evolution experiment using the powerful microbial model system, budding yeast (Saccharomyces spp.). We selected eight species with distinct temperature preferences, ranging from 15 to 35Â°C, isolated from diverse ecological and geographic locations around the world. Multiple populations of each species evolved for 600 generations under increasing temperature conditions, ranging from 25 to 40 Â°C. We observed that not all species have the ability to adapt to high temperatures, while cold-tolerant species did not exceed 34Â°C. In contrast, heat-tolerant species surpassed 36Â°C. However, growth rate measurements at the final evolved temperature revealed that cold-tolerant species exhibited the greatest increase in fitness compared to their non-evolved strains. Finally, we performed comparative genomics analyses to dissect the genetic architecture of thermal adaptation. Whole-genome sequencing demonstrated signatures of selection that explained this greater fitness. The large genomic diversity and evolutionary time frames covered by this work provide the unique opportunity to study responses to future climate change, in real-time, with high replication, and across multiple species.

124 - Does mitonuclear coevolution affect genome-wide signatures of speciation?

Shady A Kuster

Shady A Kuster, Molly Schumer, Justin C Havird, Daniel B Sloan
Colorado State University, Stanford University, University of Texas at Austin, Colorado State University

Coevolution between mitochondrial and nuclear-encoded mitochondrial (n mt) proteins is said to contribute disproportionately to speciation processes through the formation of genetic incompatibilities. During secondary contact, incompatibilities can emerge when mitochondrial haplotypes are placed into cellular contexts without their coevolved n mt proteins. However, the extent to which mitonuclear coevolution affects nuclear genome evolution and speciation is unknown. An increasingly common approach used to identify

signatures of mitonuclear coevolution is to divide the genome into gene products that function in mitochondria and those that do not. The efficacy of this approach has not been extensively tested, but a few studies have found evidence of mitonuclear coevolution contributing to reproductive isolation in this fashion. Here, we leverage a swordtail fish (Xiphophorus) species hybrid system in which a lethal mitonuclear genetic incompatibility has been identified to test the hypothesis that large numbers of n-mt loci are involved in mitonuclear incompatibilities. We partition the nuclear genome into loci that form complexes with mitochondrial gene products, loci that are targeted to mitochondria (but not in mitonuclear complexes), and loci that have no known function involving mitochondria. Linkage disequilibrium analyses show no statistical difference between the three groups, despite selection against the mitonuclear incompatibility. We additionally compare our findings to previous analyses in hybrid systems that similarly partition nuclear loci to determine when mitonuclear coevolution produces detectable genome-wide signatures. It may be that these genome-wide scans are insufficient when used on such broad functional groups, and their efficacy may be situational.

159 - The danger of flying power plants: life-history traits shape the evolution of bird mtDNA

Yury Gusarov

Yury Gusarov, Valentina Burskaya, Andrey Bushuev, Alina Mikhailova, Bogdan Efimenko, Konstantin Gunbin, Konstantin Popadin
Immanuel Kant Baltic Federal University, University of Antwerp, Lomonosov Moscow State University, Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University

A>G mutation in the mitochondria heavy chain is one of the most discussed types of mutational signatures: previous studies prove its association with mtDNA chemical damage, caused by elevated energy production level. Here we study A>G mutation patterns in birds - creatures with highest energy consumption in the world. We demonstrate a highly expected excess of A>G mutation frequency in birds compared to mammals. In order to understand biochemical mechanisms, which govern A>G mutations in birds, we describe a list of associations between A>G mutation frequencies and life history traits of birds.
First, we show that unlike mammals, birds demonstrate no connection between frequency of A>G mutations and most obvious chemical damage correlates: body mass, lifespan or basal metabolic rate (BMR). It is a surprising result, which we interpret as a sign of a highly optimized electron-transport chain, which produces a very stable level of chemical damage agents in a wide range of â€œregularâ€• conditions. However, according to our results, there is a list of â€œirregularâ€• conditions, which cause significant change of A>G mutation frequencies. Decrease of A>G mutation is caused by loss of flight: this change is in line with drastic decrease of energy consumption. Increase of A>G mutation rate we have found in diving birds (diving hypoxia is known to produce chemical damage) and in long range migrators (which are known for outstanding peak metabolism levels). The results are significant after PGLS correction and show strong phylogenetic inertia.

192 - Mitochondrially-mediated RNA interference, a fast-evolving retrograde signaling system affecting nuclear gene expression

Federico Plazzi

Federico Plazzi, Youn Le Cras, Alessandro Formaggioni, Marco Passamonti
University of Bologna, University of Bologna, UniversitÃ© Paris CitÃ©, University of Bologna, University of Bologna

Short noncoding RNAs play roles in diverse regulatory processes in eukaryotes, with miRNAs being among the most well-characterized classes. A particularly overlooked, yet fascinating regulatory network in eukaryotic cells is mito-nuclear crosstalk. Recently, miRNA-like elements of mitochondrial origin, known as smithRNAs, were discovered in the bivalve species Ruditapes philippinarum. These RNA molecules, originating in the organelle, have been shown to play regulatory roles on nuclear genes in vivo. Notably, miRNA genes can

readily evolve de novo compared to protein-coding genes: our study aims to estimate the probability of a newly arisen smithRNA finding a suitable target in the nuclear transcriptome. Simulations using transcriptomes from twelve bivalve species indicate that this probability is (i) high and (ii) not species-specific, estimated at one in a hundred million (1Ã—10âˆ’8) when allowing for five mismatches between the smithRNA and the 3â€™ mRNA, but even more mismatches are allowed in animals. Conversely, we also provide estimates on the probability of a mitochondrial genome evolving a smithRNA-like element as a function of the nucleotide composition, genomic architecture, and strand usage bias. We suggest that novel smithRNAs may evolve readily as exaptations of pre-existing mitochondrial genome architecture, characterized by common and constitutive secondary structures. Furthermore, the ability to evolve novel smithRNAs may have played a crucial role in mito-nuclear interactions during eukaryote origin and evolution.

197 - The emerging role of mitochondrially encoded small interfering RNAs to regulate nuclear gene expression.

Marco Passamonti

Marco Passamonti, Diego Carli, Alessandro Formaggioni, Iuri Icaro, Youn Le Cras, Giovanni Marturano, Giorgio Muneretto, Francesco Nardi, Oscar Wallnoefer, Federico Plazzi

University of Bologna, University of Bologna, University of Bologna, University of Bologna, Institut Pasteur , University of Siena, University of Bologna, University of Siena, University of Bologna, University of Bologna

Mitochondria seem to be connected to many more cell functions than ATP production. The possibility that mitochondrial DNA (mtDNA) can act on nuclear gene expression has been suggested only recently: mtDNAs have been found to produce small noncoding RNAs (sncRNAs), long non-coding RNAs (lncRNAs) and peptides, all of them suggested or demonstrated to interact via different pathways with the nucleus. Our research aims at characterizing the retrograde mitochondrial-to-nucleus signaling by mitochondrially encoded small RNAs (called small mitochondrial highly transcribed RNAs, smithRNAs), in a comparative way. We have already predicted the presence of many and proved functionality of two smithRNAs in Ruditapes philippinarum, in which they are likely involved in gonad formation and sex determination. While we have developed a bioinformatic pipeline to detect and analyze smithRNAs in metazoans, we present new data on: 1. their phylogenetic distribution among Metazoa; 2. their role in mito-nuclear incompatibilities and speciation; 3. their evolvability in the context of the polycistronic maturation of mtDNA; 4. their possible ways to escape mitochondria to deliver their function in the cytoplasm; and 5. their maturation processes, to understand whether they can be ascribed to the known classes of small interfering RNAs (i.e. miRNA, siRNA or piRNA), or to a new unknown one. Overall, smithRNAs are emerging as a fast-evolving generalized new form of retrograde signaling and, potentially, a common feature among metazoans, adding a brand-new functionality level of mitochondria in the eukaryotic cell.

200 - SmithHunter: a novel, unified workflow for the identification of candidate smithRNAs and their targets

Giovanni Marturano

Giovanni Marturano, Diego Carli, Claudio Cucini, Antonio Carapelli, Federico Plazzi, Francesco Frati, Marco Passamonti, Francesco Nardi

Department of Life Sciences, University of Siena, Department of Biological, Geological and Environmental Sciences, University of Bologna, Department of Life Sciences, University of Siena, Department of Life Sciences, University of Siena, NBFC, National Biodiversity Future Center, Department of Biological, Geological and Environmental Sciences, University of Bologna, Department of Life Sciences, University of Siena, NBFC, National Biodiversity Future Center, Department of Biological, Geological and Environmental Sciences, University of

Bologna, Department of Life Sciences, University of Siena, NBFC, National Biodiversity Future Center

SmithRNAs (Small MITochondrial Highly-transcribed RNAs) are a novel class of small RNA molecules that are encoded in the mitochondrial genome and regulate the expression of nuclear transcripts. Initial evidence for their existence came from the Manila clam Ruditapes philippinarum, where they have been described and whose activity has been validated through RNA injection experiments. Current evidence on the existence of these RNAs in other species is scanty and based on small RNA sequencing only. In order to evaluate the potential of this phenomenon as a novel signaling pathway in the broader context of mito-nuclear cross talking, a comprehensive investigation across Metazoa is imperative.

We therefore present SmithHunter: a novel workflow specifically designed for smithRNAs. Sequence data (from small RNA sequencing) uniquely mapping to the mitochondrial genome are clustered into putative smithRNAs and pre-filtered based on abundance, presence in replicate libraries and end-conservation. Surviving sequences are subsequently matched to the untranslated regions of nuclear transcripts to identify possible targets. To display SmithHunter structure and functionalities, as well as the effect of parameter customization, the original Manila clam data has been re-analyzed. Furthermore, preliminary results are presented on the presence of smithRNAs in a selection of metazoan species.

With this systematic approach we wish to contribute to our understanding of the functional implications of retrograde mitochondrial RNAi across animal lineages. Are smithRNAs an â€˜odd feature of an odd systemâ€™ or a novel signalling pathway of general interest, with implications as far as the origin of the eukaryotic cell?

202 - Investigating the impact of naturally occurring heteroplasmy on mitochondrial functions in gills of the blue mussel (Mytilus edulis)

Ariane Pouliot-Drouin

Ariane Pouliot-Drouin, Stefano Bettinazzi, Sophie Breton
UniversitÃ© de MontrÃ©al, University College London, UniversitÃ© de MontrÃ©al

The process of oxidative phosphorylation (OXPHOS) in mitochondria is essential for energy production, relying on genetic contributions from both the mitochondrial and nuclear genomes. Mitochondrial DNA, usually maternally inherited, undergoes strong selection pressures to maintain homoplasmy and efficient coevolution, since mitonuclear incompatibilities due to mitochondrial heteroplasmy can severely affect bioenergetic efficiency and predispose cells to diseases. However, natural heteroplasmy has been reported in about a hundred bivalve mollusk species where paternal mitochondrial DNA (M-mtDNA) is passed to male descendants and is exclusively kept in male gametes. While most tissues strictly contain maternal mitochondrial DNA (F-mtDNA), occasional presence of M-mtDNA has been observed in somatic tissues of males and, on rarer occasions, in females. This research project aimed to explore the impact of this heteroplasmy on mitochondrial functions in gills of the blue mussel Mytilus edulis by studying enzyme activities associated with energy production. Heteroplasmy levels were assessed via qPCR in both male and female gill samples and heteroplasmy was found to be much more frequent than anticipated. Spectrophotometric analysis evaluated the activity of key mitochondrial enzymes coded by the mitochondrial and nuclear genomes. A negative response to heteroplasmy was observed for all enzymes tested but only in females. Once normalized by mitochondrial content, trends of higher activities following the increase in heteroplasmy were observed. Silenced mitochondria containing M-mtDNA in females could explain our results as well as a general downregulation of OXPHOS and Krebs cycle. Further experiences will confirm theses hypotheses.

203 - The alternative mystery of the mitochondrial genome

Ludovic Nadeau-Lachance

Ludovic Nadeau-Lachance, Thierry Choquette, Hajar Hosseini Khorami, Annie Angers, Sophie Breton
UniversitÃ© de MontrÃ©al, UniversitÃ© de MontrÃ©al, UniversitÃ© de MontrÃ©al, UniversitÃ© de MontrÃ©al, UniversitÃ© de MontrÃ©al

Mitochondria possess a circular genome distinct from the nuclear genome. In animal species, the mitogenome is generally described as being limited to 37 genes, thirteen of which produce proteins involved in energy production. However, this â€œmitochondrial dogmaâ€• has recently fallen following the discovery of genes inside canonical genes in the human mtDNA. Among these recent discoveries, we have demonstrated the existence of a gene in an alternative reading frame inside the mitochondrial protein-coding gene nd4: mitochondrial alternative nd4 or mtaltnd4. We found that the protein MTALTND4 impacts cell and mitochondrial physiology, possibly by interacting with C1QBP, a protein that is known to regulate cellular energy metabolism through modulation of mitochondrial translation or through protein-protein interactions. We followed up on these discoveries related to MTALTND4 by studying its effects on gene expression. Our results suggest that the peptide downregulates the expression of several genes involved in metabolic or biosynthetic processes. The discovery of MTALTND4 sheds light on the intricate nature of the mitochondrial genome and enhances our understanding of mitochondrial biology.

205 - Patterns of gene loss in endosymbionts at the origin of eukaryotes

Aidan Pierce

Aidan Pierce, Nick Lane, Andrew Pomiankowski
University College London, University College London, University College London

Understanding the formation of the eukaryotic cell remains a major challenge for evolutionary biology. A central question relates to the transition of endosymbiont to organelle and why modern cases of endosymbiosis donâ€™t follow the same patterns of gene loss and genome size reduction as seen in mitochondria. Using computer simulations we provide insights into patterns of endosymbiont gene loss in the context of multi-level selection (on the eukaryote itself and competition between its constituent organelles). In doing so, we identify how endosymbiont population size is a key factor that can determine the types of genes retained by endosymbionts, and highlight how epistasis within and between endosymbionts impacts gene retention. These results provide a context for organelle gene loss at the origin of eukaryotes.

207 - Key differences in the mitochondrial mutation spectrum and heteroplasmy levels between laboratory and natural populations of Caenorhabditis elegans

Alexandre A. Schifano

Alexandre A. Schifano, Ulfar Bergthorsson, Vaishali Katju
Uppsala University, Uppsala University, Uppsala University

Random mutations display biases in their relative frequencies with important consequences for genome structure and composition. While laboratory studies have provided insights into the spectrum of spontaneous mutations, the laboratory environment of optimal growth conditions may produce its own biases. Hence, further investigation is necessary to determine if laboratory results are truly representative of natural populations. We analyzed the mitochondrial (mtDNA) genomes of 1,524 Caenorahbditis elegans natural isolates comprising 550 unique mtDNA haplotypes to investigate patterns of mtDNA evolution in the wild. Ancestral reconstruction was used to polarize 3,321 mutations (92 indels, 3,229 SNPs) and the results were compared to mutations identified in experimentally evolved laboratory populations

under minimized selection. The frequency of mutations is associated with both codon position and site degeneracy in a manner consistent with strong purifying selection on mtDNA protein-coding sequences. There is significant variation in the per site nonsynonymous and synonymous substitutions between genes and nd1, nd2, and nd4 in ETC complex I are particularly enriched in nonsynonymous substitutions. Whereas laboratory evolved lines exhibit a uniform distribution of heteroplasmic frequencies of mtDNA variants, natural isolates possess a bimodal distribution comprising (i) a majority of variants having reached fixation and (ii) possibly deleterious heteroplasmic variants in lower frequencies. Although some of the patterns of mtDNA mutational bias are similar between the laboratory and wild populations, we also observe significant differences. In particular, transversions typically associated with oxidative damage are less common at four-fold degenerate sites in natural populations than in the laboratory.

225 - Do the sex-specific mitochondrial genes play a role in sex determination in DUI bivalves?

Julie Brémaud

Julie BrÃ©maud, AlizÃ©e Debelli, Annie Angers, Bernard Angers, Sophie Breton
DÃ©partement de Sciences Biologiques, UniversitÃ© de MontrÃ©al, Groupement de Recherche Interuniversitaire en Limnologie, DÃ©partement de Sciences Biologiques, UniversitÃ© de MontrÃ©al, DÃ©partement de Sciences Biologiques, UniversitÃ© de MontrÃ©al, DÃ©partement de Sciences Biologiques, UniversitÃ© de MontrÃ©al, Groupement de Recherche Interuniversitaire en Limnologie, DÃ©partement de Sciences Biologiques, UniversitÃ© de MontrÃ©al, Groupement de Recherche Interuniversitaire en Limnologie

While mitochondria are maternally inherited in almost all animals, there is a group of about a hundred bivalve species that have a biparental transmission of mitochondria named DUI (Doubly Uniparental Inheritance). In these species, offspring inherit their mitochondria from both parents, but only male embryos will retain the paternal mitochondria in their germ cells. Even more intriguing, the paternal (M-mtDNA) and maternal (F-mtDNA) mitochondrial genomes have additional genes absent in other animal species (orphan genes or ORFans): M-orf (Male-specific ORF) in M-mtDNA and F-orf (Female-specific ORF) in F-mtDNA. It has been hypothesized that the existence and maintenance of such genes could imply a function in sex determination, but this remains to be demonstrated. Immunoblotting and immunocytochemistry methods were performed to determine the tissue and temporal expression patterns of the M-ORF and F-ORF proteins and their sub-cellular localization. Co-immunoprecipitations allowed us to find their interaction partners. Our results show that F-ORF is expressed in all tissues of both sexes in the freshwater mussel Venustaconcha ellipsiformis and marine mussel Mytilus edulis, while M-ORF is only expressed in the gonadal tissues of males sexually mature. The two proteins co-localize with the mitochondria in the gametes (F-ORF in eggs and M-ORF in sperm), and M-ORF is also found in the sperm acrosome. Both proteins possess interaction partners with functions in gamete-gamete interaction. These preliminary results suggest non-conventional functions for mtDNA-encoded proteins.

374 - Origins and implications of mitochondrial genes expansion in bivalves.

Diego Carli

Diego Carli, Federico Plazzi, Enrico Massimiliano Negrisolo, Marco Passamonti
University of Bologna, University of Bologna, University of Padova, University of Bologna

Bivalvia are a class of molluscs with various oddities in their mitochondrial genomes, which are partly associated to the peculiar mitochondrial segregation mechanism that is exclusive to this group. Many of them host a large number of unassigned regions and elongations that are uncommon in animal mitochondrial genomes. To re-evaluate the annotation and function of mitochondrial proteins in bivalves, we established a comparative approach, paying particular attention to nuclear OXPHOS compatibility and gene boundaries. Since mutations involving start and stop codons can significantly affect the size of a gene, we initially investigated the genesis of

extensions by assessing oversupply and lack of such codons throughout genomes. Next, we used AlphaFold and related scoring methods to determine the potential influence of elongations into OXPHOS core proteins on their interactions with nuclear-encoded subunits. In this work, we share some insights into our attempts to re-examine the mitochondrial genetics of bivalves. Our goals are twofold: to give a comprehensive assessment of the group's mitochondrial gene features and to elucidate the causes, effects, and conservation of mitochondrial expansion in metazoans.

384 - Clues on the biogenesis of mitochondrial short non-coding RNAs from in silico and in vitro approaches

Alessandro Formaggioni

Alessandro Formaggioni, Federico Plazzi, Marco Passamonti
University di Bologna, University of Bologna, University of Bologna

The network of regulatory systems that has arisen from the genomic interactions between the nuclear and mitochondrial genomes is, for the most part, poorly known. A new class of short non-coding RNA was recently discovered in four phylogenetically distant species (i.e., Mus musculus, Drosophila melanogaster, Caenorhabditis elegans and Ruditapes philippinarum). These small RNAs are transcribed in the mitochondrial genome and, according to in silico and in vivo investigations, appear to target nuclear transcripts. However, how these Small Mitochondrial Highly Transcribed RNAs (smithRNAs) can regulate different cellular processes is not clear yet. In the present work we identified a possible smithRNA maturational pathway by analyzing which proteins interact with this class of small RNAs.
Analyzing publicly available RNA immunoprecipitation (RIP) and cross-linking immunoprecipitation (CLIP) libraries revealed that in Homo sapiens, proteins involved in miRNA maturation (Ago2, Drosha, and DGCR8) interact with mitochondrial tRNAs, where the majority of smithRNAs are located. Furthermore, the protein-mitochondrial tRNA interaction was also observed in D. melanogaster and M. musculus Ago2, and C. elegans ERGO1, all of which are members of the Argonaute family and typically interact with nuclear short non-coding RNAs. According to our results, it is likely that proteins involved in the maturation of nuclear small RNAs have been coopted for the smithRNA maturational pathway.
Moreover, we conducted in vitro experiments in which biotinylated smithRNA sequences were immunoprecipitated alongside the interacting proteins. These interacting proteins were identified and characterized through mass spectrometry, aiming to depict a possible pathway of smithRNA biogenesis.

397 - Evidence of convergent evolution in the nuclear and mitochondrial OXPHOS genes across Squamata deep lineages

Oscar Wallnoefer

Oscar Wallnoefer, Alessandro Formaggioni, Federico Plazzi, Marco Passamonti
University of Bologna, University of Bologna, University of Bologna, University of Bologna

The oxidative phosphorylation system (OXPHOS) is composed of five complexes, with subunits encoded by both the mitochondrial genome (mtOXPHOS) and the nuclear genome (nucOXPHOS). mtOXPHOS and nucOXPHOS subunits tightly interact; indeed, signs of coevolution between these two sets of genes have been observed in Metazoa.
Squamata is particularly intriguing from a mitochondrial point of view: the phylogenetic signal inferred from mitochondrial markers is radically discordant compared to nuclear markers. According to mitochondrial topology, lizards of the family Agamidae (Acrodonta) are in sister relationship with the crown-group of snakes (Serpentes), rejecting the monophyly of Iguania (Acrodonta + Pleurodonta), which is supported by most phylogenetic studies on Squamata. It has been proposed that the convergence between the early lineages of snakes

and agamids is due to a common selective pressure on OXPHOS.

In this study, we aim to investigate whether convergent evolution between these two lineages (i.e., Serpentes and Agamidae) is also detectable in nucOXPHOS. We annotated 88 nuclear OXPHOS genes from 56 squamatan genomes encompassing 24 families, along with the 13 mitochondrial OXPHOS genes. Phylogenetic trees were inferred from both datasets, and we evaluated signals of convergent evolution between Agamidae and Serpentes.

The evolutionary rate correlation analysis revealed a clear coevolutionary signal between nucOXPHOS and mtOXPHOS, which is stronger when considering only subunits in close contact within the complexes. Moreover, some nucOXPHOS genes resulted under positive selection along the Serpentes and Acrodonta lineages, with clues of convergent substitutions observed between the two.

422 - The stick insect genus Bacillus and the role of small mitochondrial highly transcribed RNAs (smithRNAs) in hybridization and speciation.

Iuri Icaro

Iuri Icaro, Diego Carli, Alessandro Formaggioni, Federico Plazzi, Marco Passamonti

B.I.G.E.A. Department, University of Bologna, B.I.G.E.A. Department, University of Bologna, B.I.G.E.A. Department, University of Bologna, B.I.G.E.A. Department, University of Bologna, B.I.G.E.A. Department, University of Bologna

The stick insect genus Bacillus displays complex reproductive dynamics, with species such as B. grandii, B. rossius, and B. atticus engaging in various modes of reproduction, including bisexual and facultative parthenogenetic forms. Hybridization among these species has given rise to novel hybrids like B. whitei and B. lynceorum. Notably, mitochondrial DNA analysis has revealed asymmetrical hybridization events, with B. rossius consistently acting as the maternal parent. Recent research focused on a class of small non-coding RNAs called smithRNAs, encoded by the mitochondrial genome and implicated in regulating nuclear gene expression. The distribution and characteristics of smithRNAs were examined across parental Bacillus species. Results indicated a significant presence of smithRNA in B. grandii and B. atticus, while B. rossius, the mitochondrial donor species for hybrids, exhibited minimal quantities of smithRNA. These findings suggest a potential role for smithRNAs in cyto-nuclear interactions and the evolution of isolating mechanisms driving speciation within the Bacillus genus. The study underscores the importance of understanding the molecular mechanisms underlying hybridization and speciation processes, particularly in taxa exhibiting reticulate evolution. Bacillus stick insects serve as an experimental model for exploring the intricate relationships between sexual and clonal taxa, shedding light on the evolution of reproductive strategies and species boundaries.

564 - Has the human mitochondrial genome evolved under environmentally mediated natural selection?

Finley Grover-Thomas

Finley Grover-Thomas, Lucy van Dorp, Francois Balloux, Aida AndrÃ©s, Florencia Camus

University College London, University College London, University College London, University College London, University College London

Physiologists and classical geneticists would likely disagree on the adaptive potential of the small, but important, mitochondrial genome (mtDNA). The long-standing assumption in evolutionary and population genetics that positive selection rarely acts on mitochondrial genomes is, however, being slowly eroded away by increasing volumes of new evidence from model organisms and natural populations alike. In contrast, evidence of local adaptation in human mitochondrial DNA is still debated, with much of the discussion occurring in the early 2000s. Over a decade later, the quality and size of genomic and environmental datasets has dramatically improved, as have our methods to detect signatures of selection within genomes. Consequently, we can now address the question of local adaptation in human

mitochondrial genomes with much increased power.

Using publicly available genomic and remote sensing climate databases (NCBI assemblies and WorldClim) we have curated a collection of 23,791 complete human mitochondrial genomes from 80 countries paired with associated bioclimatic metadata, including temperature metrics and precipitation levels. This dataset was used in an established generalised linear model approach that identified sites where the distribution of genotypes was better explained by models which included bioclimatic variables. This would be unexpected under neutral evolution. We further inspected identified candidates of local adaptation in the context of mitochondrially relevant stressors (diet, pathogen load, and thermoregulation), and examined changes to protein structure and mitochondrial physiology. Together, our work establishes the role of environmentally mediated natural selection on human mitochondrial evolution, highlighting the contribution of mitochondrial genomes to adaptive process.

648 - How can the mitogenome help to elucidate the evolutionary history of Kelp Gull?

Gisele P M Dantas

Gisele PM Dantas, Lorena V Pinto
PUC Minas, PUC Minas

The genetic composition of a population can be affected by some factors that interfere with the process of transmitting the alleles to future generations, such as genetic drift and selection. The mitochondrial DNA (mtDNA) has been extensively used in evolutionary and phylogenetic studies. However, some studies have indicated selection in the mitochondrial genome, which will be important to interpretation of evolutionary history with accuracy. Thus, this study search evaluates the evolutionary history of Kelp Gull through the complete mitogenome of Kelp Gull (Larus dominicanus) from South American and Antarctica. Additionally, previous studies proposed the hypothesis that the mitochondrial DNA of this species would have undergone a selective sweep, that is, intense loss of genetic diversity in the molecule due to selection in one of its genes. For this, we sequenced 21 complete genomes of Kelp Gull, 6 from Brazil, 5 from Argentina, and 10 from the Antarctic Peninsula. We did not observe haplotype sharing between localities, indicating high haplotype diversity but low nucleotide diversity. There is a clear indication of population expansion reflected in the Network, in addition to positive selection signals as indicated in the genes encoding mitochondrial DNA, which showed low variability and few haplotypes, except for the NADH2 and NADH6 genes. Our results showed that mitochondrial genes are under directional selective pressure, with few mitochondrial genes â€“ NADH2 and NADH6, showing to be variable, so a selection followed by recent population expansion, with population structuring, is the most likely scenario.

663 - Integration of large and diverse angiosperm DNA fragments into Asian Gnetum mitogenomes

Shu-Miaw Chaw

Shu-Miaw Chaw, Chung-Shien Wu, Ruiiiang Wang
Biodiversity Research Center, Academia Sinica, Biodiversity Research Center, Academia Sinica, South China Botanical Garden, Chinese Academy of Science

Horizontal gene transfer (HGT) events have rarely been reported in gymnosperms. Gnetum is a gymnosperm genus comprising 25â€’35 species sympatric with angiosperms in West African, South American, and Southeast Asian rainforests. Only a single acquisition of an angiosperm mitochondrial intron has been documented to date in Asian Gnetum mitogenomes. We wanted to develop a more comprehensive understanding of frequency and fragment length distribution of such events as well as their evolutionary history in this genus. We sequenced and assembled mitogenomes from five Asian Gnetum species. These genomes vary remarkably in size and foreign

DNA content. We identified 15 mitochondrion-derived and five plastid-derived (MTPT) foreign genes. Our phylogenetic analyses strongly indicate that these foreign genes were transferred from diverse eudicotsâ€”mostly from the Rubiaceae genus Coptosapelta and ten genera of Malpighiales. This indicates that Asian Gnetum has experienced multiple independent HGT events. Patterns of sequence evolution strongly suggest DNA-mediated transfer between mitochondria as the primary mechanism giving rise to these HGT events. Most Asian Gnetum species are lianas and often entwined with sympatric angiosperms. We therefore propose that close apposition of Gnetum and angiosperm stems presents opportunities for interspecific cell-to-cell contact through friction and wounding, leading to HGT. Our study reveals that multiple HGT events have resulted in massive amounts of angiosperm mitochondrial DNA integrated into Asian Gnetum mitogenomes. Gnetum and its neighboring angiosperms are often entwined with each other, possibly accounting for frequent HGT between these two phylogenetically remote lineages.

684 - Hints of functional necessity of common repeat in mtDNA

Victor Shamanskiy

Victor Shamanskiy, Konstantin Gunbin, Konstantin Popadin

Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University

Ageing is associated with accumulation of somatic mutationsâ• . This process is especially pronounced in mitochondrial genomes (mtDNA) of postmitotic cells, where accumulation of large-scale somatic mitochondrial deletions is associated with both healthy aging and age-related mitochondrial encephalomyopathies.

We just recently considered the influence on the formation of deletions of such factors as structure secondary structure of the human mitochondrial genome (https://doi.org/10.1186/s12915-023-01606-1). In our follow-up studies we defined and test two hypotheses: that increased longevity and decreased the occurrence of mitochondrial encephalomyopathies may be owing disruption mtDNA common repeat by synonymous variants, and that the colocalization of direct and nested inverted repeats shape the distribution of deletions. More deep mechanistic understanding of the somatic mtDNA deletion formation can shed light on human age-related diseases and help to understand haplogroup - specific risks of different mtDNA-related encephalomyopathies. In this study, we focused on the process of repeat occurrence in the mitochondrial genome of mammals.

We analyzed both changes in repeats during evolution of mammalâ€™s mitochondrial genome, in particular the general repeat, and the conservativeness of regions adjacent to the repeats.

The results show that there is no evidence of negative selection against direct repeats in long-lived mammals, and at the same time, we see cases of the appearance and consolidation in the population of long perfect direct repeats that lead to the majority of deletions, which raises many questions. Supported by RSFâ„–21-145-20143.

741 - Mitochondrial mutation spectrum in Chordates: damage versus replication signatures, causes, and dynamics

Dmitrii Iliushchenko

Dmitrii Iliushchenko, Bogdan Efimenko, Konstantin Gunbin, Konstantin Popadin

Immanuel Kant Baltic Federal University, ., Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University

Detecting mutational signatures in diverse biological samples, particularly in oncology, requires the analysis of mutational spectra. It has resulted in the identification of approximately 60 unique mutational signatures linked to various mutagens. However, the mutational

processes within mitochondrial DNA (mtDNA) are not as well characterised. This project aims to bridge this knowledge gap by analysing mitochondrial mutagenesis through a comprehensive 192-component mutational spectrum, utilising 86,149 polymorphic synonymous mutations from the mitochondrial CytB gene across 967 chordate species. Our analysis has revealed insightful patterns in mtDNA repair and mutation processes, which can be classified into three predominant signatures: (i) symmetrical mutations distributed uniformly across both DNA strands, primarily induced by gamma DNA polymerase, which account for about 50% of all mutations; (ii) asymmetrical, heavy-strand-specific C to T mutations, accounting for approximately 30%; and (iii) asymmetrical, heavy-strand-specific A to G mutations, comprising about 20%, which appear to be influenced by metabolic rates and age-related factors. It is hypothesised that these asymmetrical signatures result from strand-specific damage in conjunction with suboptimal base excision repair mechanisms on the mtDNA's lagging (heavy) strand. Understanding these intricate mutational mechanisms is critical for developing solutions aimed at preventing somatic mtDNA mutations, which could potentially reduce the progression of age-related ailments.
Supported by RSFâ„–21-145-20143.

792 - Mitonuclear genotypes as integrators of epistasis and GxE for complex traits in Drosophila

David M Rand

David M Rand, Leah Darwin, Faye Lemieux, Yevgeniy Raynes
Brown University, Brown University, Brown University, Brown University

Proper cellular function requires coordinated expression of the 37 mitochondrial-encoded and >1200 nuclear-encoded genes that have been interacting for more than 1 billon years. Disruption of these complex interactions can cause a variety of diseases and influence adaptive evolution in heterogeneous environments. To advance our understanding of how these complex genetic interactions influence organismal fitness we have constructed a panel of 90 mitonuclear genotypes built from 22 different mtDNAs (10 Zimbabwe, 10 Bejing, from D. melanogaster plus D. simulans and D. yakuba mtDNAs) placed onto two nuclear chromosomal backgrounds (OreR and DGRP375) in duplicate. Each genotype was cultured on standard Drosophila diet and a diet containing 25 mM rotenone, a natural pesticide that inhibits the activity of mitochondrial complex I, NADH dehydrogenase. Climbing speed, flight ability, development time and body weight were measured in adult males and females from each diet. This design allows for partitioning of mtDNA effects, nuclear genotype effects, diet environment, and their interactions in models that test for GxGxE effects. There were strong mitonuclear epistatic effects across the phenotypes, with further modification of epistasis by the rotenone treatment. The mtDNA, nuclear and environmental effects were different for each phenotype. The outgroup mtDNAs of D. simulans and D. yakuba had climbing and flight performances comparable to D. melanogaster mtDNAs despite 100s of substitutions in these mtDNA genomes. Overall, the analyses identify specific mtDNA and nuclear genome pairs that show sensitivity or resistance to rotenone, amenable to further genetic mapping of epistatic and environmental interactions.

798 - Influence of G4 Structures on Mutation Spectra in mammalian mtDNA

Viktoria Skripskaya

Viktoria Skripskaya, Bogdan Efimenko, Konstantin Popadin
Center for Mitochondrial Functional Genomics, Immanuel Kant Baltic Federal University, Center for Mitochondrial Functional Genomics, Immanuel Kant Baltic Federal University, Center for Mitochondrial Functional Genomics, Immanuel Kant Baltic Federal University

The G-quadruplex (G4) structures in nucleic acids, prevalent in guanine-rich regions, play a key role in cellular regulation, and act as molecular switches for transcription and replication in mitochondrial DNA (mtDNA). G4s can induce polymerase stalling, suggesting they may influence mutagenesis around G4s. Our analysis of G4 sequences in mtDNA from human cancers

(https://doi.org/10.1038/s41588-019-0557-x) and mouse somatic tissues (https://doi.org/10.7554/eLife.83395) shows an increased Ch>Th rate (h means heavy strand of mtDNA) and a decreased Gh>Ah rate compared to nearby regions used as control. Notably, a significant increase of Th>Gh transversion within G4 regions was observed in human cancers. We hypothesize that the observed mutational bias may emerge during the replication, at a stage when the double-stranded DNA is temporarily separated. This separation could facilitate the formation of G4s on one strand, potentially halting the replication machinery until these structures are resolved. Meanwhile, the complementary strand that does not form G4s remains in a single-stranded state, leaving it susceptible to damage and subsequent mutagenesis. Additionally, the increased rates of Th>Gh observed in cancers might be attributable to the malfunctioning of the translesion synthesis (TLS) polymerase, which is capable of bypassing DNA lesions, but has low fidelity. Consequently, our findings hint at a potential mutational bias in the G4 regions of mammalian mtDNA, warranting further investigation.
This project was supported by the Russian Science Foundation (RSF) grant â„– 21-75-20143

858 - Earth, Wind, and... Water?: A Look into Mitonuclear Coevolution across Birds

Dan Alexandru Tatulescu

Dan A Tatulescu, Joseph J Dubie, Justin C Havird
The University of Texas at Austin, The University of Texas at Austin, The University of Texas at Austin

Mitonuclear coevolution is a phenomenon involving mitochondrial and nuclear genomes where changes in one of the genomes drive complementary changes in the other genome. This preserves functionality of oxidative phosphorylation as the nuclear genome encodes most of the proteins involved in these pathways and must work in conjunction with proteins encoded by the faster evolving mitochondrial genome. With a wide range of lifestyles of varying energetic needs, birds provide an interesting group in which to examine how mitonuclear coevolution plays out. More specifically, here we aimed to determine how evolutionary rate correlations (ERCs) between nuclear and mitochondrial genes vary in birds with different means of locomotion and metabolic needs, i.e., in different metabolic backgrounds. To do so, we gathered the nuclear and mitochondrial genomes of 105 birds with representatives from every order, making sure to include birds with terrestrial, aquatic, and flighted lifestyles. We analyzed these genomes using a combination of Orthofinder and ERCnet to obtain ERCs for every gene-by-gene comparison within each bird. We will present on patterns that may come up in our ERC comparisons and discuss them in relation to differences in lifestyle. To clarify, we will present on the general hypothesis that ERCs should be strongest between mitochondrial genes and nuclear genes encoding products with mitochondrial functions. We will also examine whether this pattern is stronger in birds with more energetically demanding lifestyles, typically meaning flighted lifestyles.

871 - Temperature sensitivity of mtDNA mutational spectrum: insights from cold mammals

Alina G. Mikhailova

Alina G. Mikhailova, Alexander Kuptsov, Konstantin Popadin
Immanuel Kant Baltic Federal University, Severtsov Institute of Ecology and Evolution RAS, Immanuel Kant Baltic Federal University

Mitochondrial genomes of chordata animals reflect their life-history traits. Neutral mutations in mtDNA can be informative about how mutagenesis is linked with different physiological and ecological parameters of an organism, thus revealing internal mutation signatures specific to mitochondria. Initially, comparing the mitochondrial mutation spectra between different classes of chordates we have shown that frequencies of A>G substitutions are sensitive to organismal longevity and body temperature(10.1093/nar/gkac779,10.1101/2020.07.25.221184), precisely to level of metabolism and emanating from it chemical damage(10.1101/2023.12.08.570826). Based on our previous results we hypothesised that mammals with a reduced body temperature and metabolic rate have decreased frequencies of A>G on heavy strand and thereby increased fraction of A in four-fold degenerate sites

compared to warmer ones. We split all species into â€œcolderâ€• and â€œwarmerâ€• groups, taking into account not only the body temperature but also such life-history traits as hibernation and daily torpor. We observed that there is correlation between frequencies of A>G and temperature, making the temperature-sensitive mutational signature weakly pronounced in mammals in comparison with longevity-sensitive. However, additional comprehensive study should elucidate the effect of metabolism in warm-blooded animals, since we observed unusual trends. For example, fraction of A in Xenarthra is low (in line with increased A>G) compared to hibernators, despite their extremely low metabolic rate, while naked mole-rat, being a cold-blooded mammal with extended lifespan has lower frequencies of A>G. Supported by RSFâ„–21-145-20143.

897 - Evolutionary Dynamics of Mitochondrial Sequences Present in Nuclear Telomere-to-Telomere (T2T) Genomes of Apes

Edmundo Torres-González

Edmundo Torres-GonzÃ¡lez, Saswat K Mohanty, Jessica M Storer, Kateryna D Makova, T2T Primates Consortium
Department of Biology, The Pennsylvania State University, Huck Institutes of the Life Sciences, The Pennsylvania State University, Department of Biology, The Pennsylvania State University, Huck Institutes of the Life Sciences, The Pennsylvania State University, University of Connecticut, Department of Biology, The Pennsylvania State University, Center for Medical Genomics, The Pennsylvania State University, Center for Medical Genomics, The Pennsylvania State University, Genome Informatics Section, National Human Genome Research Institute, Department of Genome Sciences, University of Washington School of Medicine

Nuclear sequences of mitochondrial origin (NUMTs) are remnants of the endosymbiotic origin of mitochondriaâ€”where mitochondrial DNA (mtDNA) are incorporated into the nuclear genome. Novel NUMT insertions occur 1/10,000 births in humans and are often indistinguishable from mtDNA, confounding studies of low-level heteroplasmy as well as genome assemblies and alignments. Despite the importance of NUMTs, they have been critically under-analyzed due to the lack of complete genome assemblies. To fill this gap, we used telomere-to-telomere (T2T) genomes for seven ape species (bonobo, chimpanzee, human, gorilla, Sumatran and Bornean orangutans, and siamang), to perform a comprehensive analysis of NUMT insertion and evolution. The number of NUMTs increased, in T2T vs. previous assemblies, across all species. The highest NUMT contentâ€”total number of bp annotated as NUMTsâ€”was observed in bonobo and chimpanzee, consistent with prior studies, but also in Bornean orangutan. Interestingly, we found significantly higher NUMT content in Bornean orangutan as compared to its sister species, Sumatran orangutan, suggesting different rates of NUMT insertion over 1 MY of divergence. Siamang, a gibbon species, shows a dramatically lower NUMT content compared to great apes. Next, we studied the dynamics of NUMT insertions using multi-species alignments and identified shared and lineage-specific NUMTs. Furthermore, we analyzed the sequence context of NUMTs and their proximity to transposable elements (TEs) to glean potential insertion and propagation mechanisms, respectively. This study provides a complete picture of NUMT evolutionary dynamics in the available T2T ape genomes, which should be augmented by intra-species comparisons.

903 - Effect of mutagens on the stability of mitochondrial genome

Natalia Ree

Natalia Ree, Nadezhda Potapova, Nikita Van Leiden, Konstantin Gunbin, Konstantin Popadin
Immanuel Kant Baltic Federal University, Institute for Information Transmission Problems, Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University

The question of factors that affect the mutation rate of the mitochondrial genome is still unanswered. There is a discrepancy about the role of reactive oxygen species produced by mitochondria in mtDNA instability. The issue is whether mitochondrial membranes are

allowing mutagens from the environment to pass through and cause mutations. Our study aims to investigate mutagenic agents of various natures on the stability of mtDNA. We analyzed 2448 files of WGS data from Kucab et al., 2019 study in which IPS human cells were exposed to 79 known or suspected environmental carcinogens or were kept in different control mediums. Cram files were realigned on hg38, then only reads aligned on chrM and unmapped were selected and trimmed with TRIMMOMATIC, then aligned on hg38 with BWA MEM, somatic mutation calling was performed with MitoHPC. Mutations with AF<0.5 were selected. We showed that not all agents affected mitochondrial ploidy (16 mutagens increased and 27 decreased mtDNA content), their action was not associated with the chemical nature of agents or their biological role. Also only 25 agents increased the quantity of mutations in mtDNA at about 2 fold comparing to water, in 6 cases it was concentration-dependent dependence. In most cases the dominant mutation was G>A, the most outstanding mutagen signature was for DBAC, which is an adductor forming agent.
Conclusion: Mutations in mitochondrial genomes are unlikely to be influenced by mutagens in the environment, but some of them cause changes in mitochondria copies and increase the overall number of mutations.

931 - Codon usage in mitochondrial DNA: dissecting the contributions of mutagenesis and selection

Eldar Khaibulin

Eldar Khaibulin, Alina G. Mikhailova, Konstantin Gunbin, Konstantin Popadin
Center for Mitochondrial Functional Genomics IKBFU, Vavilov Institute of General Genetics, Center for Mitochondrial Functional Genomics IKBFU, Vavilov Institute of General Genetics, Center for Mitochondrial Functional Genomics IKBFU, Center for Mitochondrial Functional Genomics IKBFU

Mitochondrial DNA (mtDNA) codon usage is shaped by mutagenesis and selection, yet their relative impacts are not fully understood. Our study investigates two hypotheses: whether the neutral mutational spectrum driven by chemical damage is the dominant factor in mtDNA evolution, or if it has evolved through natural selection to improve translation efficiency.
We conducted simulations to analyze mtDNA codon usage as a function of mutagenesis, reconstructed from polymorphisms, and selection, taking into account two codon-anticodon pairing types: Watson-Crick and wobble. Our model accurately predicts the mtDNA codon usage across hundreds of chordates.
Our findings reveal that over 50% of the codon-anticodon pairings adhere to the Watson-Crick type. This proportion of pairings exceeds observed levels at saturation in our simulations and in actual codon usage, albeit with variation across genes.
Our analysis revealed minimal changes in tRNA anticodons, underscoring their evolutionary conservation. Notably, species possessing atypical anticodons demonstrated divergent codon usage patterns compared to species with typical anticodons.
Analysis of mutational spectra, derived exclusively from synonymous substitutions within the gnomAD and HelixMT databases, showed concordance between them. Moreover, we observed significant differences between the mutational spectra of homoplasmic and heteroplasmic variants. These differences remain pronounced when the data are stratified into allele frequency categories.
Our analyses support both hypotheses concerning the factors driving mtDNA evolution, underscoring the complexity of these mechanisms. This necessitates further investigation, which is currently underway.
This project was supported by the Russian Science Foundation grant No. 21-75-20145.

1120 - The genetic and phenotypic correlates of mtDNA copy number in the Mexican Biobank

Amara Shaukat

Amara Shaukat, Maria Jose Palma, Lourdes Garcia Garcia, AndrÃ©s Moreno Estrada, Arslan Zaidi, Mashaal Sohail
Centro de Ciencias Genomics, UNAM, Centro de Ciencias Genomics, UNAM, Centro de Ciencias Genomics, UNAM, Instituto Nacional de

Salud PÃºblica, Laboratorio Nacional de GenÃ³mica para la Biodiversidad (UGA-LANGEBIO), CINVESTAV, Department of Genetics, cell biology, and Development, University of Minnesota, Institute of Health Informatics, University of Minnesota, Centro de Ciencias Genomics, UNAM

Mitochondria are important components of the cell and mitochondrial dysfunction can be involved in complex disease through alteration in mitochondrial copy number. We analyzed ~4400 individuals from the Mexican Biobank (MXB) to investigate the genetic and phenotypic correlates of mtDNA copy number (mtCN). We computed mtCN using array data and investigated the role of blood cell type variation in mtCN variation. We computed polygenic scores for 11 blood cell types using trans-ethnic GWAS summary statistics. Cumulatively, they explained a small fraction of the mtCN variation (adj R2 = 0.0029) with only Eosinophil (WB-cell type) significantly associated with mtCN, which we used as a covariate for all subsequent analyses. We performed a genome-wide association study on mtCN and identified three variants rs533353962, rs372129830 and rs867763743 significantly associated with mtCN that are low-frequency variants in MXB (MAF ~0.01) and are rare or absent elsewhere in the world. These fall in SNAPIN shown to regulate mitochondrial homeostasis at synapses, INF2 shown to be involved in stimulating mitochondrial division through mechanisms leading to mitochondrial membrane constriction, and PKD1, mutations in which have been implicated in polycystic kidney disease likely through mitochondrial abnormalities. Further, we performed a pheWAS for mtCN with 20 complex traits and observed mtCN to be significantly associated with Creatinine, Triglycerides, Cholesterol, HDL and LDL levels, Diastolic blood pressure and Rheumatoid arthritis (FDR < 0.05). Overall, these results indicate a significant role of the mitochondrial genome through copy number influencing various complex diseases in a population-specific manner.

73 - Distinctive bacterial community structure of sunflower soils from two geographical locations

Blessing Chidinma Nwachukwu

Blessing C Nwachukwu, Ayansina S Ayangbenro, Olubukola O Babalola
University of the Witwatersrand, North-West University, North-West University, North-West University

Agricultural practices are aim to enhance nutrients in the fields or crops, soil quality, and crop productivity; though, the impacts of ecological degradation and loss have affected ecosystem functioning. Crop rotation and crucial rhizosphere-microorganism networks may improve a stable ecosystem and functions. When comparing the sunflower rhizosphere to surrounding bulk soils, the application of next-generation sequencing technique will aid in the characterization of all bacterial species present. Investigations were conducted on the predictive functional traits of the bacterial community structure of sunflower rhizosphere and bulk soils grown under various farming practices at two locations in the North West Province of South Africa. DNA was extracted from the rhizosphere and bulk soils from the crop rotation sites (rhizosphere soils from Lichtenburg [LTR] and bulk soils from Lichtenburg [LTB]) and the monocropping sites (rhizosphere soils from Krayburg [KRPR] and bulk soils from Krayburg [KRPB]), and sequenced using 16S amplicon sequencing. Utilizing bioinformatics tools, the sequenced dataset was analyzed. In bulk soils, Firmicutes and Actinobacteria predominated, whereas Proteobacteria and Planctomycetes dominated in rhizospheric soils. There were significant (P<0.05) differences in predicted functional groups and bacterial structure at the phylum and family levels across the soils at each site. Bacterial distribution across the study locations were influenced by physicochemical factors. This study provide evidence on the abundance of bacterial structures in the sunflower soils and their predictive functions at the growing stage. It is recommended that future research be done on the imminent crop production and management for improved yields.

95 - Magellanic penguin feather-dependent ectosymbiont bacteria discovered by genomes assembled from metagenomes

Brenda Paulina Vázquez Ramírez

Brenda P. VÃ¡zquez, Manuel Ochoa-SÃ¡nchez, Ulises Rodriguez-Cruz, Valeria Souza
Licenciatura en Ciencias GenÃ³micas, UNAM, Institute of Ecology, UNAM, Institute of Ecology, UNAM, Institute of Ecology, UNAM

The microorganisms associated with the Magellanic penguin plumage are believed to be acquired through the environment with which they come into contact, which would imply that they do not depend on the penguin to subsist. An alternative interpretation is that the microbiome is feather-dependent and therefore characteristic of the penguin species. To discern between these possibilities, assembled genomes were obtained from shotgun metagenomic samples and subsequently analyzed for the metabolic pathways present in the microorganisms found. All genomes had at least one essential pathway absent. These results suggest a Magellanic penguin-dependent ectosymbiont microbiome.

199 - Origin and molecular evolution of an obligate digestive symbiosis

Marleny García Lozano

Marleny GarcÃ-a Lozano, Christine Henzler, Miguel Ã•ngel GonzÃ¡lez-Porras, InÃ¨s Pons-Guillouard, Aileen Berasategui, Christa Lanz, Heike Budde, Kohei Oguchi, Yu Matsuura, Yannick Pauchet, Shana Goffredi, Takema Fukatsu, Donald Windsor, Hassan Salem
Mutualisms Research Group, Max Planck Institute for Biology, Mutualisms Research Group, Max Planck Institute for Biology, Mutualisms Research Group, Max Planck Institute for Biology, Mutualisms Research Group, Max Planck Institute for Biology, Max Planck Institute for Biology, Max Planck Institute for Biology, National Institute for Advanced Industrial Science and Technology, Tropical Biosphere Research Center, University of the Ryukyus, Department of Insect Symbiosis, Max Planck Institute for Chemical Ecology, Department of Biology, Occidental College, National Institute for Advanced Industrial Science and Technology, Smithsonian Tropical Research Institute, Mutualisms Research Group, Max Planck Institute for Biology

Timing the acquisition of a beneficial microbe relative to the evolutionary history of its host can shed light on the adaptive impact of a partnership. Here, we investigated the onset and molecular evolution of an obligate symbiosis between Cassidinae leaf beetles and Candidatus Stammera capleta, a Gammaproteobacterium. Residing extracellularly within foregut symbiotic organs, Stammera upgrades the digestive physiology of its host by supplementing plant cell wall-degrading enzymes. We observe that Stammera is a shared symbiont across tortoise and hispine beetles that collectively comprise the Cassidinae subfamily, despite differences in their folivorous habits. In contrast to its transcriptional profile during vertical transmission, Stammera elevates the expression of genes encoding digestive enzymes while in the foregut symbiotic organs, matching the nutritional requirements of its host. Despite the widespread distribution of Stammera across Cassidinae beetles, symbiont acquisition during the Paleocene (~62 mya) did not coincide with the origin of the subfamily. Early diverging lineages lack the symbiont and the specialized organs that house it. Reconstructing the ancestral state of host-beneficial factors revealed that Stammera encoded three digestive enzymes at the onset of symbiosis, including polygalacturonaseâ€”a pectinase that is universally shared. Although non-symbiotic cassidines encode polygalacturonase endogenously, their repertoire of plant cell wall-degrading enzymes is more limited compared with symbiotic beetles supplemented with digestive enzymes from Stammera. Highlighting the potential impact of a symbiotic condition and an upgraded metabolic potential,Stammera-harboring beetles exploit a greater variety of plants and are more speciose compared with non-symbiotic members of the Cassidinae.

236 - The Influence of Aspergillus Starter Cultures on Microbes Found in Miso

Kimberly Lynette Acevedo

Kimberly L Acevedo, John G Gibbons
University of Massachusetts Amherst, University of Massachusetts Amherst

This research explores the intricate interplay within the microbial community of miso fermentation, focusing on the co-culture of A. sojae and challengers, a combination of various different prokaryotes and Saccharomyces cerevisiae. Aspergillus sojae (AS), renowned for its proteolytic prowess in soy fermentation, serves as a parallel model to A. oryzae/A. flavus, presenting an intriguing hypothesis of AS as a distinct population within Aspergillus parasiticus (AP), adapted to the food environment. This study investigates species interactions between wild and domesticated Aspergillus starter cultures and the challengers commonly found in miso fermentation. Through pair-wise dual assays, we measure the growth and inhibitory effects of Aspergillus strains when co-cultured with these challengers. Identification of differences between domesticated and wild cultures prompts a dual-RNA-seq analysis in strains displaying variations. Our hypothesis posits that the absence of secondary metabolite production in A. sojae will lead to heightened growth of challengers in co-culture. This study provides a glimpse
into the unexplored territory of microbial dynamics in miso fermentation. Join us at the poster presentation to unravel the collaborative journey of A. sojae and challengers, shedding light on the sensory and physical characteristics of the final fermented product. The outcomes not only contribute to our understanding of microbial interactions but also offer insights into optimizing miso production and enhancing the nuances of this cherished fermented food.

285 - Benchmarking methods for inferring ecological interactions from ancient eDNA data

Fiona Callahan

Fiona Callahan, Kaiyuan Li, Rasmus Nielsen
University of California, Berkeley, University of California, Berkeley, University of California, Berkeley

With increasing availability of ancient and modern environmental DNA sequencing, whole-community species occurrence or abundance data over time and space is becoming more available. These data can be used is to infer associations among species and associations between species and abiotic environmental factors, to understand causal factors affecting community structure including species extirpiation. Here, we use simulated data to evaluate five common methods for this type of inference. We find that under a variety of reasonable data simulation scenarios, all methods have high false discovery rates. Additionally, all methods are outperformed by simple linear or logistic regression techniques under many realistic scenarios. We conclude that continued methods development and higher sample sizes are needed to leverage the potential of spatio-temporal eDNA data for understanding causal determinants of community structure.

292 - Dynamics of bacterial recombination in the human gut microbiome

Zhiru Liu

Zhiru Liu, Benjamin H Good
Stanford University, Stanford University

Horizontal gene transfer is a ubiquitous force in microbial evolution. Previous work has shown that the human gut is a hotspot for gene transfer between species, but the more subtle exchange of variation within species -- also known as recombination -- remains poorly characterized in this ecosystem. Here, we show that the genetic structure of the human gut microbiome provides an opportunity to measure recent recombination events from metagenomic sequences of fecal samples, enabling quantitative comparisons across diverse commensal species that inhabit a common environment. By analyzing recent recombination events in the core genomes of 29 human gut bacteria, we observed widespread heterogeneities in the rates and lengths of transferred fragments, which are difficult to explain by existing models of ecological isolation or homology-dependent recombination rates. We also show that natural selection helps facilitate the spread of genetic variants across strain backgrounds, both within individual hosts and across the broader population. These results

shed light on the dynamics of in situ recombination, which can strongly constrain the adaptability of gut microbial communities. Furthermore, we discuss the implications of these observations on the coevolution (or "codiversification") of the gut microbiota with modern human populations.

299 - Monkey's bacteria: knowing the gut microbiome of wild mexican primates

Diego Zubillaga-Martín

Diego Zubillaga-MartÃ-n, Ella VÃ¡zquez, Brenda SolÃ³rzano, Luisa FalcÃ³n

Insituto de EcologÃ-a, UNAM, Instituto de EcologÃ-a, UNAM, ENES MÃ©rida, UNAM, Instituto de EcologÃ-a, UNAM

Recently, we have discovered amazing relations between the microbiome and the evolution and ecology of their hosts. For instance, the influence of the gut microbiome in health and behavior of non-human primates. Despite this, little is known about the composition of the gut microbiome in wild populations of New World monkeys. MÃ©xico has three species of wild primates: two howler monkeys (Alouatta palliata, A. pigra) and one spider monkey (Ateles geoffroyi). Howlers are a model for gut microbiome studies in neotropical primates. Their microbiome varies naturally and is influenced mostly by age, diet, and seasonality. Notably, there is rather little information for Mexican howler populations and less so for spider monkeys. We evaluated the diversity and abundance of gut bacterial communities of 155 individuals from wild populations of these three species, encompassing most of its distribution range. We extracted DNA from fresh stool samples and amplified and analyzed the 16S rDNA V4 region. Results suggest that phylogeny is the principal factor in the structure of these species gut microbiome, while diet and geographic distribution are also important. The most abundant microbial Phyla are shared among species but differing at finer taxonomic levels. Ateles geoffroyi has the most differentiated microbial communities, harboring more rare taxa. We also identified significant differences in relation with sex in Alouatta pigra, but not in A. palliata. The gut microbiome abundance and diversity patterns per species are consistent with their diet. Our study provides the first overall microbiome data for these three Mexican primate species.

316 - Deciphering the molecular genetic basis of deep and long diving in cetaceans

RAN TIAN

RAN TIAN

Nanjing Normal University

Cetaceans exhibit a remarkable diversity in diving depth and duration, implying a range of adaptive strategies to stressors, such as hydrostatic pressure and oxidative stress. Here, morphologic, comparative genomic, and engineering experiments were performed to dissect the genetic basis of cetaceans with prolonged diving capacities. We found that maximum dive duration and dive depth all increased with skull size, myoglobin concentration and blubber thickness of cetaceans, which may be an adaptation to support oxygen storage and thermoregulation during extreme hypoxia and cold exposure in deep diving. We identified 115 diving phenotype related genes, 290 positively selected genes and 109 rapidly evolving genes in deep/long diving cetaceans. The pervasive molecular adaptations of these genes are involved in lipid and cholesterol metabolism, lung injury and oxidative stress tolerance, helping deep diving cetaceans to thrive at high pressures and hypoxia. We further provided functional evidences that unique changes in XDH and ATP8B1 might be beneficial for enhancing antioxidative capacity and reducing lung inflammation in deep/long diving sperm whale, supporting the positive role of these genes function in oxidative stress and decompression sickness adaptation, respectively. Our findings offer novel insights into the genetic basis underpinning the adaptive strategies of cetaceans to their extreme diving and may have implications for understanding decompression sickness and hypoxia resistance.

453 - The emergence of a resistance mechanism to a synthetic antimicrobial in the resistome: when genomic mobilization is enough

Stella Cellier-Goetghebeur

Stella Cellier-Goetghebeur, ClaudÃ¨le Lemay-St-Denis, Maxime St-Aubin, Joelle N. Pelletier
Department of Biochemistry and Molecular Medicine, UniversitÃ© de MontrÃ©al, PROTEO, The QuÃ©bec Network for Research on Protein, Function, Engineering and Applications, CGCC, Center in Green Chemistry and Catalysis, Department of Biochemistry and Molecular Medicine, UniversitÃ© de MontrÃ©al, PROTEO, The QuÃ©bec Network for Research on Protein, Function, Engineering and Applications, CGCC, Center in Green Chemistry and Catalysis, Department of Biochemistry and Molecular Medicine, UniversitÃ© de MontrÃ©al, PROTEO, The QuÃ©bec Network for Research on Protein, Function, Engineering and Applications, CGCC, Center in Green Chemistry and Catalysis, Department of Biochemistry and Molecular Medicine, UniversitÃ© de MontrÃ©al, PROTEO, The QuÃ©bec Network for Research on Protein, Function, Engineering and Applications, CGCC, Center in Green Chemistry and Catalysis, Department of Chemistry, UniversitÃ© de MontrÃ©al

While resistance to natural antibiotics is ancient, this is not the case for the emergence of mechanisms of resistance to antimicrobials of synthetic origin. Trimethoprim is one such antimicrobial: this synthetic molecule introduced in the 1960s targets the ubiquitous and essential bacterial dihydrofolate reductase enzymes (FolA). However, Type B dihydrofolate reductase (DfrB) enzymes were identified in the 1970s as intrinsically providing resistance to trimethoprim. When FolA is inhibited by trimethoprim, DfrB acts as a substitute to maintain active metabolism. Although FolA and DfrB catalyze the same reaction, they are evolutionary unrelated. The origin of DfrB enzymes is unknown; here, we examine the mechanisms that led to DfrB enzyme evolution and emergence in the clinical resistome.
Our recent work reveals that DfrB enzymes are present not only in pathogenic bacteria but also in environmental bacteria not subjected to anthropogenic pollution. Here, through curation and analysis of a dataset of 82 experimentally characterized DfrB enzymes identified in a wide variety of environments such as permafrost and the rhizosphere, we demonstrate that the emergence of DfrB enzymes in bacteria predates the introduction of trimethoprim in clinical settings. Our genomic analyses reveal that DfrB-mediated trimethoprim resistance was recruited at least once to the clinical setting, from agricultural bacteria. This led to the emergence of a distinct population of dfrB genes defined by â€œclinical traitsâ€• rather than â€œenvironmental traitsâ€•. Our observations allow proposing a model that recapitulates evolutionary trajectories leading to the emergence of DfrB enzymes in the modern resistome.

462 - Quantifying negative selection on introgressed fragments in human gut bacteria

Sophie Jean Walton

Sophie Jean Walton, Dmitri A Petrov, Benjamin H Good
Stanford University, Stanford University, Stanford University

Introgression is ubiquitous across the tree of life, enabling the exchange of genetic information between otherwise isolated groups of individuals. In sexually reproducing organisms, previous work has shown that negative selection against introgressed fragments can contribute to the genetic isolation of species and subspecies. However, the evolutionary forces acting on introgressed fragments in bacteria, which reproduce clonally and exchange genetic material through horizontal gene transfer, remain poorly characterized. Recent work has demostrated that the rates of genetic exchange can be reduced between groups of ecologically and genetically similar strains. However, it remains unclear to what extent negative selection or other molecular barriers to recombination (e.g. bacterial defense systems) are driving this genetic isolation. To distinguish between these mechanisms, we developed a new approach for quantifying the strength of negative selection on horizontally transferred fragments, which compares the frequency spectra of individual transfer events and synonymous mutations on the core genome. We detect these events by leveraging a hidden Markov model that infers the site specific

ancestry of genomes that reccombine via horizontal gene transfer. We applied this approach to a collection of Phocaeicola vulgatus genomes from hundreds of human hosts, which were previously shown to be divided into two genetically isolated clades. Our preliminary results suggest that horizontally transferred fragments from the other clade tend to be deleterious in their new genetic background. This provides new evidence that selective forces (e.g. genetic incompatibilities) may contribute to the genetic isolation of this prominent member of the human gut microbiome.

635 - A universal and constant rate of gene content change traces pangenomes to LUCA

Katharina Trost

Katharina Trost, Michael R. Knopp, Jessica L.E. Wimmer, Fernando D.K. Tria, William F. Martin
Institute of Molecular Evolution, Heinrich Heine University, Institute of Molecular Evolution, Heinrich Heine University, Institute of Molecular Evolution, Heinrich Heine University, Los Alamos National Laboratory, Institute of Molecular Evolution, Heinrich Heine University

Prokaryotic genomes constantly undergo gene flux via lateral gene transfer, generating a pangenome structure consisting of a conserved core genome surrounded by a more variable accessory genome shell. Over time, flux generates change in genome content. Here we measure and compare the rate of genome flux for 5,655 prokaryotic genomes as a function of amino acid sequence divergence in 36 universally distributed proteins of the informational core (IC). We find a clock of gene content change. Long-term average rate of gene content flux is remarkably constant across all higher prokaryotic taxa sampled. A linear regression predicts between 0% (Chlamydiae) and 30-33% gene content differences (Alphaproteobacteria, Gammaproteobacteria, Clostridia) for conspecific genome pairs having identical IC sequences, corresponding to the accessory genome proportion for well-sampled species. The data suggest that pangenome structure is a general feature of prokaryotic genomes and that it has been in existence since the divergence of bacteria and archaea.

712 - The evolution of Arabidopsis thaliana-associated Pseudomonas and Sphingomonas

Haim Ashkenazy

Haim Ashkenazy, Or Shalev, Talia Karasov, Derek Lundberg, Manuela Neumann, Detlef Weigel
Max Planck Institute for Biology TÃ¼bingen, Max Planck Institute for Biology TÃ¼bingen, Max Planck Institute for Biology TÃ¼bingen, Max Planck Institute for Biology TÃ¼bingen, Current address: Swedish University of Agricultural Sciences, Max Planck Institute for Biology TÃ¼bingen, Max Planck Institute for Biology TÃ¼bingen

Members of the bacterial genera Pseudomonas and Sphingomonas are often highly abundant in the leaf-associated plant microbiota. While some species or isolates have been described as pathogenic, others are beneficial to host plants or merely commensals. In this study, we focus on the evolution of a local collection of plant-associated Pseudomonas and Sphingomonas strains, to better characterize the factors governing plant colonization by members of these taxa.

We collected 1,524 Pseudomonas isolates and 425 Sphingomonas isolates from wild Arabidopsis thaliana plants, over several years and sites near TÃ¼bingen, Germany. We sequenced and assembled the genomes of these bacterial isolates and inferred their pan-genome, based on ortholog groups. Out of the 72,397 and 56,923 ortholog groups specified for Pseudomonas and Sphingomonas, respectively, only 1.3% and 6.9% belonged to the bacterial â€˜core genomeâ€™, while 36.3% and 40.7% were unique to individual isolates - suggesting high variability in isolatesâ€™ gene content.
Next, we grouped closely-related strains (determined by their genome-wide average nucleotide identity, ANI) and compared the gene content within and between groups. By comparing these closely-related isolates, we discovered genomic islands that varied among

isolates and we characterized their function and evolutionary dynamics. We further identify genomic islands shared between Pseudomonas and Sphingomonas isolates.

Taken together, our results demonstrate that focusing on the evolution of closely-related strains, which are co-colonizing the same plant population, can shed light on the complex host-microbe and microbe-microbe interactions shaping the plant microbiome.

802 - Effect of glyphosate use on microbial communities in a black leptosol of Yucatan, Mexico

Anell Del Carmen García-Romero

Anell del C GarcÃ-a-Romero, Miriam M Ferrer, HÃ©ctor Estrada-Medina
Universidad AutÃ³noma de YucatÃ¡n , Universidad AutÃ³noma de YucatÃ¡n, Universidad AutÃ³noma de YucatÃ¡n

The herbicide glyphosate inhibits EPSPS enzyme involved in shikimate biosynthesis in plants, bacteria, archaea and fungi. This herbicide is applied in higher doses to those recommended in Yucatan. To identify the effects of glyphosate doses variation in bacteria and fungi communities of a black Leptosol, we conduct an exploratory greenhouse experiment and environmental DNA using a metagenomic approach. The DNA was extracted from soil treated with commercial dose (x), 0.5x, 1.5x and 0x glyphosate samples. The V3-V4 16SRNA and ITS2 regions were amplified and sequenced using NovaSeq 6000 Illumina. QIIME was used to process uncut sequences, unmerged, classified and quantify different taxa among samples. Phyloseq and DESeq2 was used to compare the microbiota enrichment of glyphosate treated samples. Haloferula helveola (bacteria) and Arxiella dolinchadrae (fungi) were enriched in 1.5x glyphosate tratment. Those species are common in soils, and if they are glyphosate resistant maybe used in bioremediation processes. Colletotrichum truncatum (fungi) was also enriched in 1.5x treatment but is potentially pathogenic. A low enrichment was observed in 1x doses with no taxa remarkable in both cases. Future studies, including isolation of colonies by selective cultures and metabolomics and shotgun sequences would help to understand the species potential in bioremediation projects and their function in the soil health

803 - Bacterial microbiota during embryonic development: a maternal inheritance

Elizabeth Selene Gómez Acata

Nina Montoya Ciriaco, Elizabeth S. GÃ³mez Acata, Luc Dendooven, Stephanie Hereira Pacheco, Anibal H. de la Vega PÃ©rez, Yendi E. Navarro Noya
Doctorado en Ciencias BiolÃ³gicas, Centro Tlaxcala de BiologÃ-a de la Conducta, Universidad AutÃ³noma de Tlaxcala, Tlaxcala, Mexico, Laboratorio de Interacciones BiÃ³ticas, Centro de InvestigaciÃ³n en Ciencias BiolÃ³gicas, Universidad AutÃ³noma de Tlaxcala, Tlaxcala, Mexico, Laboratorio de Interacciones BiÃ³ticas, Centro de InvestigaciÃ³n en Ciencias BiolÃ³gicas, Universidad AutÃ³noma de Tlaxcala, Tlaxcala, Mexico, Laboratorio de Interacciones BiÃ³ticas, Centro de InvestigaciÃ³n en Ciencias BiolÃ³gicas, Universidad AutÃ³noma de Tlaxcala, Tlaxcala, Mexico, EstaciÃ³n CientÃ-fica La Malinche, Centro Tlaxcala de BiologÃ-a de la Conducta, Universidad AutÃ³noma de Tlaxcala, Tlaxcala, Mexico, Laboratorio de Interacciones BiÃ³ticas, Centro de InvestigaciÃ³n en Ciencias BiolÃ³gicas, Universidad AutÃ³noma de Tlaxcala, Tlaxcala, Mexico

The bacterial microbiota plays a crucial role in different aspects of animal hosts, including nutrient acquisition, energy balance, physiology, reproduction, immunity, behavior, life history and organ development of the host. However, our understanding of the microbiota's role during embryonic development in non-mammalian viviparous vertebrates is still limited. In this study, samples from the intestine, mouth, cloaca, and the aseptic ventral skin from maternal individuals of the viviparous lizard Sceloporus grammicus Wiegmann, 1828 were taken to compare their microbiota with that found in the embryo's amniotic environment (amniotic fluid, membrane and extraembryonic yolk) at the last stage of development. We used metabarcoding of the 16S rRNA to explore the bacterial microbiota in the

gastrointestinal tract and amniotic environment of lizard embryos and mothers. We found higher bacterial diversity in maternal tissues than in embryonic ones, and 78% of the embryonic amplicon sequence variants (ASVs) were found in the maternal bacterial microbiota. The embryonic bacteria were more similar to those found in the mouth and aseptic ventral skin of the mother than those in the cloaca and intestine. Additionally, the bacterial composition was found to be more similar among embryos than with their mother. Our findings provide evidence of vertical microbiota transfer during embryonic development in non-mammalian viviparous vertebrates and suggest that the transmission of bacteria from the mother to the embryos is controlled during early embryonic stages.

838 - COMPARISON OF THE MICROBIOTA OF A RED SOIL UNDER DIFFERENT DOSES OF GLYPHOSATE.

María alejandra Ocaña Ek

MarÃ-a Alejandra OcaÃ±a , Miriam Monserrat Ortega , Hector Medina
Universidad Autonoma de YucatÃ¡n , Universidad Autonoma de Yucatan , Universidad Autonoma de Yucatan

Glyphosate inhibits the enzyme 5-enolpyruvyl shikimate-3-phosphate synthase (EPSPS), present in plants, bacteria, fungi, and protozoa. This enzyme is vital in these organisms due to its involvement in amino acid biosynthesis and metabolic processes. The exposure to glyphosate causes metabolic disruption in several soil microbial communities. To evaluate the effects of this herbicide on the microbiota, an exploratory study was conducted using a metagenomic approach. Glyphosate was applied using a commercial dose (x), 0.5x, 1.5x and 0x on pots with red Leptosol in a open greenhouse. DNA was extracted and amplified for Illumina sequencing of V3-V4 16SRNA and ITS2 regions. QIIME was used to process uncut sequences, unmerged, classified and quantify different taxa among samples. Phyloseq was used to compare the microbiota enrichment of glyphosate treated samples with Deseq2. The enriched bacteria species for the 1.5x dose of glyphosate were Clostridium paraputrificum,; Gemmiger formicilis, Catenibacterium mitsuokai that has been found in vertebrates species and are potentially pathogens. Enriched fungi were for the 1.5x dose of glyphosate were Trichoderma asperellum, and unknown species from the genus Trechispora and Microascus that are associated with plants, mosses, and soil. Fungi and bacteria may be resistant to glyphosate and has the potential to be used in restoration programs. For this and future studies, these enriched organisms can be isolated and cultured and their functions evaluated using omics tools such as shutgun and metabolomics.

972 - Bridging Nutritional Gaps: The Role of Microbiomes in host productivity

Luis D Alcaraz

Gerardo Mejia, Angelica M Jara-ServÃ-n , Luis D Alcaraz
Faculty of Sciences, National Autonomous University of Mexico, Faculty of Sciences, National Autonomous University of Mexico, Faculty of Sciences, National Autonomous University of Mexico

The symbiosis connection between the host's nutritional status and its microbiome is essential for good health. We investigated some effects of the host's nutritional status and its repercussions on the composition and functioning of the microbiome. We found that under nutrient-limiting conditions, the importance of the microbiome becomes evident. We explored this using model tomato plants and their phenotypes and metagenomic and 16S rRNA gene sequencing. By inoculating tomato plants with diverse soils as a microbial source, we compared the results of the root microbiome in soils and simplified hydroponic systems. This approach allowed us to identify plant growth-promoting microbial communities. Our results indicate that it is possible to select specific microbial communities that can improve plant resilience and growth, even under nutrient-limiting conditions. We also contribute to the One Health concept by connecting microbes in production systems. Our results highlight the impact of the microbiome on host adaptation to nutritional stress, opening clues about the structuring and mechanisms of interaction with the host.

1022 - Phylogeny-aware strain profiling limits false positive detections and quantifies divergence time of novel strains

Evan B Qu

Evan B Qu, Jacob S Baker, Laura Markey, Tami D Lieberman

Massachusetts Institute of Technology, Massachusetts Institute of Technology, Massachusetts Institute of Technology, Massachusetts Institute of Technology

Reference-based strain profiling methods are well-understood to be limited in the event of novel diversity but provide advantages over reference-free methods in samples with low coverage and multiple strain colonization. We present PHLAME, a novelty-inclusive method to profile bacterial strains up to their most recent shared ancestor along a reference phylogeny. PHLAME leverages the read distribution across clade-specific SNPs and a Bayesian mixture model to distinguish between lowly abundant strains represented in the reference set and highly abundant strains that diverge from the known phylogeny. By masking known clades from the phylogenies of different bacterial species, we show that posterior estimates of divergence time inferred from metagenomics track well with true branch lengths, indicating that many fixed mutations are acquired sequentially along branches of a tree. Using a ground-truth dataset of 4,055 isolate genomes and 33 metagenomes generated from the same samples, we show that existing strain profiling methods infer more false than true positives as the number of reference genomes and profiling resolution increases, whereas PHLAME limits the false detection rate to <0.05 across profiling resolutions while retaining comparable recall to other methods. We demonstrate the utility of PHLAME in detecting novel phylogenetic associations using public skin metagenomes, including the identification of recently emerged clades of C. acnes that are geographically restricted yet highly prevalent, suggesting limitations to the rate of spreading of this bacteria. PHLAME will enable researchers to conduct rigorous epidemiology and evolution studies from rapidly accruing metagenomic data.

55 - Inferring demographic and selective histories from population genomic data using a two-step ABC approach in species with coding-sparse genomes: an application to human data

Vivak Soni

Vivak Soni, Jeffrey D Jensen

Arizona State University, Arizona State University

The demographic history of a population, and the distribution of fitness effects (DFE) of new mutations in functional genomic regions, are fundamental factors dictating genetic variation. Although both demographic and DFE inference has been performed extensively in humans, these approaches have either been limited to simple demographic models involving a single population and a single size change, or, where a complex population history has been inferred, without accounting for the confounding effects of selection at linked sites. Taking advantage of the coding-sparse nature of the genome, we propose a two-step Approximate Bayesian Computation (ABC) approach in which coalescent simulations are first used to infer a complex four-population out-of-Africa demographic model, utilizing large non-functional genomic regions that are likely free from the effects of background selection. We then use forward-in-time simulations to perform DFE inference in functional regions, conditional on the complex demography inferred in the first step, and utilizing expected background selection effects in the estimation procedure. Throughout, recombination and mutation rate maps were used to account for the underlying empirical rate heterogeneity across the human genome. Furthermore, within this ABC framework it is possible to utilize and fit multiple aspects of the data, including the site frequency spectrum, linkage disequilibrium, and divergence information. This approach represents a generalized framework for such large-scale inference in species with coding-sparse genomes, whereas those characterized by coding-dense genomes require more computationally-intensive joint and simultaneous inference of demographic and selective parameters.

60 - Enrichment of hard sweeps on the X chromosome across six Drosophila species

Mariana Harris

Mariana Harris, Bernard Y Kim, Nandita Garud

UCLA, Stanford University, UCLA

The X chromosome is hemizygous in males, leaving it fully exposed to the effects of natural selection and, thus, potentially subject to different evolutionary dynamics than autosomes. This is of particular interest given the role of sex chromosomes in local adaptation, speciation, and sexual dimorphism. Here, we investigate the differences in the mode and tempo of adaptation between the X chromosome and autosomes. Specifically, we test the hypothesis that hard sweeps, expected when adaptation is gradual, are more common on the X chromosome due to there being a smaller effective population size and a reduction in standing genetic variation resulting from more efficient selection. By contrast, we predict that soft sweeps, expected when adaptation is rapid, are more common on the autosomes. Recently, we found an enrichment of hard sweeps on the X chromosome relative to the autosomes in a North American population of D. melanogaster. Now, we generalize these findings by analyzing diversity patterns across six Drosophila species, where we find regions with steep reductions in diversity and elevated haplotype homozygosity on the X chromosome as compared to autosomes. To assess if these signatures are consistent with positive selection, we simulate a broad range of evolutionary scenarios, spanning variations in demography, mutation rate, recombination rate, background selection, hard sweeps, and soft sweeps. We find that the diversity patterns observed on the X are most consistent with hard sweeps, suggesting that hard sweeps have played a significant role in shaping diversity patterns on the X chromosome across multiple species.

220 - The impact of natural selection on patterns of genetic diversity

Jesús Abad Guzmán-López

JesÃºs A. GuzmÃ¡n-LÃ³pez, Parul Johri, Diego Ortega-Del Vecchyo

Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico, Santiago de QuerÃ©taro, QuerÃ©taro, MÃ©xico, Department of Biology, University of North Carolina, Chapel Hill, NC 27599 Department of Genetics, University of North Carolina, Chapel Hill, NC 27599 Integrative Program for Biological & Genome Sciences, University of North Carolina, Chapel 11 Hill, NC 27599, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, Universidad Nacional AutÃ³noma de MÃ©xico, Santiago de QuerÃ©taro, QuerÃ©taro, MÃ©xico

Synonymous mutations are changes in protein-coding DNA that do not modify the amino acid; these mutations often occur in the third base of a codon due to the genetic code redundancy. Hence, synonymous mutations are assumed to have a neutral or nearly neutral effect on the fitness of a population. Evidence has shown that a high proportion of synonymous sites are under strong selection in species such as Drosophila melanogaster. Here we perform extensive genomic simulations with the software SLiM and subject them to different degrees of selection, to analyze the impact of natural selection on various summaries of genetic diversity such as the site frequency spectrum (SFS). We discuss the impact of our results on estimates of genetic diversity in various species and compare them with empirical data.

237 - Amino acid physicochemical property convergence may underlie adaptive functional convergence.

Shanshan Chen

Shanshan Chen, Zhengting Zou

Key Laboratory of Zoological Systematics and Evolution Institute of Zoology, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Key Laboratory of Zoological Systematics and Evolution Institute of Zoology, Chinese Academy of Sciences

Many studies have proposed various methods to detect the molecular basis for adaptive functional convergences between species, often attributing them to the convergence of amino acid states in functionally related proteins. Here we group amino acids into physicochemically similar property classes, to test whether convergence of amino acid properties, not states, may also underlie organismal or functional adaptive convergence. Investigating three previous case studies of different species and methodology, we found functionally related genes with putatively adaptive property convergence, providing evidence for our hypothesis, and emphasizing the importance of considering amino acid property when studying adaptive sequence convergence.

267 - Detecting natural selection in Holocene Europe with multi-locus genotype identity scans

Devansh Pandey

Devansh Pandey, Mariana Harris, Nandita R Garud, Vagheesh M Narasimhan
The University of Texas at Austin, University of California Los Angeles, University of California Los Angeles, The University of Texas at Austin

Large sample size ancient DNA studies now offer insights into human evolution by examining changes in genomes over time. Recent efforts using time-stratified aDNA for selection scans mainly focused on single locus approaches. As different types of statistics often provide complementary information about selective events, we carried out multi-locus genotype scans for natural selection on a time transect of 708 samples over the past 7,000 years of European history. We used the G12 statistic that had been designed for unphased diploid data and validated its performance on aDNA using simulations. We incorporated the missingness, ascertainment bias, DNA damage, and unphased random allele calling typical of aDNA processing utilizing a demographic model that captures broad features of the allele frequency spectrum of European genomes. With these simulations as well as using positive control loci that have been previously identified and functionally validated in modern Europeans, we found that the multi-locus statistic G12 performed significantly better than an allele frequency-based selection statistic, SweeepFinder2 that has also been recently used on ancient DNA. Applying our approach to the ancient DNA time transect, we identified 14 signals of selection across the four time periods with half of the observed signals seen in the earliest time period, not detectable in the latest time period. Taken together, our results suggest that selective events in European prehistory, including from times when humans first began to domesticate animals, have been obscured by neutral processes such as drift or by demographic events including admixture

484 - Joint modeling of demography and background selection reveals complex patterns of polymorphism

Gustavo Valadares Barroso

Gustavo Valadares Barroso, Aaron Ragsdale
University of Wisconsin-Madison, University of Wisconsin-Madison

Classic background selection theory describes the deterministic effect of strong negative selection on diversity at linked neutral sites. However, the model breaks down when drift allows deleterious variants to segregate more freely. Recent models have incorporated weak selection, but predictions remain inaccurate for 2Ns values in the vicinity of â€“4 (the peak of interference among deleterious variants). This is also the most interesting region of parameter space, where the reduction in diversity at neutral sites (B-values) is strongest. Therefore, reliable models of linked selection depend on accurate prediction of B-values under interference. Here we develop a new

method (moments++) to achieve this goal, based on the Hill-Robertson system of two-locus statistics. We model the effect of selection acting on weakly and moderately constrained loci linked to a neutral locus at arbitrary recombination distance. We treat the joint effect of multiple selected loci independently, but we show interference among them is well captured through local rescaling of mutation, recombination and selection in an iterative procedure that converges quickly. We demonstrate by forward simulations that our formulation leads to B-value predictions that are substantially more accurate than the current state-of-the-art, bridging the gap between weak and strong selection. Moreover, moments++ is more flexible than existing methods and supports any demographic history, as well as mutation and recombination maps. We couple complex demography with variable recombination, mutation and selection intensities along the genome to show that patterns of diversity can be misinterpreted by other models.

567 - Evolutionary Analysis of Plasmodium falciparum Across the Asian and African continents

Daniel Garcia Ruiz

Daniel Garcia Ruiz, Hsiao-Han Chang

Bioinformatics Program, Taiwan International Graduate Program, Institute of Statistical Science, Academia Sinica, Institute of Bioinformatics and Structural Biology, College of Life Sciences and Medicine, National Tsing Hua University,, Institute of Bioinformatics and Structural Biology, College of Life Sciences and Medicine, National Tsing Hua University

Malaria remains one of the most burdening infectious diseases despite control programs, the development of vaccines, and the availability of anti-malaria drugs. While the majority of malaria cases occur in the African continent, it is noteworthy that resistance to several anti-malaria drugs initially emerged in Asia. The current study aims to understand Plasmodium falciparum's evolutionary history, including population structure and natural selection, by performing a comparative analysis between the African and Asian continents. Variant information from the Plasmodium falciparum dataset 7 was used to estimate nucleotide diversity between populations and to perform principal component analysis (PCA) and will be used to estimate pairwise dN/dS. Antigenic genes and genes involved in drug resistance show different patterns of nucleotide diversity across different continents. Kelch 13 and MDR1, genes associated with artemisinin and multi-drug resistance as well as the antigenic genes SERA2 and TRAP show reduced diversity in the Asian continent. Contrarily, PCA shows that most of the variation occurs in the Asian continent, specifically in the Greater Meakong subregion. Our preliminary results indicate the presence of distinct evolutionary pressures in the Asian and African continents, however, dN/dS estimation is needed to better characterize the selection forces acting upon Plasmodium falciparum across different continents. We anticipate that our study will provide valuable insights to address the challenges posed by drug resistance in malaria treatment and elimination efforts.

629 - Detecting hybrid incompatibilities using the two locus site frequency spectrum

Vitor C Sousa

Alexandre Blanckaert, Vitor C Sousa

cE3c â€“ centre for ecology, evolution and environmental changes & CHANGE - Institute for Global Change and Sustainability, Department of Animal Biology, CiÃªncias â€“ Faculty of Sciencies of the University of Lisbon, cE3c â€“ centre for ecology, evolution and environmental changes & CHANGE - Institute for Global Change and Sustainability, Department of Animal Biology, CiÃªncias â€“ Faculty of Sciencies of the University of Lisbon

Disentangling the genetics of speciation remains one of the major questions in evolutionary biology, due to their complexity and diversity. Many recent studies have shown the ubiquitous nature of hybridization, highlighting the importance of natural hybrid populations in the speciation process. Hybridization has a dual role: on one hand, it can act as a source of genetic variation while on the other hand, it can

result in reduced fitness due to genetic incompatibilities that act epistatically. Indeed, hybrid populations have the particularity of expressing genetic incompatibilities (known as Dobzhansky-Muller incompatibilities), leaving potential unique signatures of epistasis and selection on those genes. Here, to capture the signatures of hybrid incompatibilities due to epistasis between a pair of interacting loci, we propose to extend a classical population genetic summary statistic, the Site Frequency Spectrum (SFS), to a pair of regions of the genome. The resulting Two Locus Site Frequency Spectrum (2L-SFS) captures not only the polymorphism within regions but also the linkage disequilibrium existing both within and between genomic regions. We derived a Composite Likelihood Method to compute the probability of observing a given 2-SFS, conditioned on the marginal SFS to correct for the effect of selection on the polymorphism within a region. Preliminary results indicates that we this approach allows to reliably detect the two regions involved in genetic incompatibilities. Our approach uses the whole genome as reference, allowing this approach to be used without the need to reconstruct the underlying demographic history.

631 - A new cross-population selection scan using genealogies uncovers tens of millennia of canid evolution

Mattias Sherman

Mattias Sherman, Jasmin Rees, Leo Speidel

Genetics Institute, University College London, Genetics Institute, University College London, Department of Genetics, University of Pennsylvania, Genetics Institute, University College London, Ancient Genomics Lab, Francis Crick Institute

The identification of the traits and genes under selection in populations is a central question in evolutionary biology, and has been key to understanding how environmental pressures lead to genomic adaptations and phenotypic changes over time. However, challenges have hindered progress in this field, including the confounding effects of population structure and admixture, as well as the signature-diluting effect of complex traits. Recent developments in genome-wide genealogical methods now allow us to efficiently produce more informative statistics, which can be used to boost traditional approaches used for detecting natural selection. Here we focus on cross-population selection scans that look out for particularly differentiated genomic regions to quantify selection acting on a subpopulation. In most relevant cases, groups are only recently diverged compared to average coalescence times between them, and so the ability to estimate cross population divergence at individual genomic loci is hampered by the inclusion of many uninformative old coalescence events. We have developed a novel method to overcome this issue by establishing a time cutoff which masks older coalescence events, dated by reconstructing whole genome genealogies, from the analysis. We demonstrate a large statistical boost in cross-population selection scans in humans. We then apply our method to the Dog10K consortium data. Key questions remain unanswered relating to wolf evolution and dog origins, including responses to climatic changes over past millennia and adaptive pathways leading to domestication. Our analysis uncovers several new, strong signatures of selection relating to wolf evolution and dog domestication.

814 - NeMu: A Comprehensive Pipeline for Accurate Reconstruction of Neutral Mutation Spectra from Evolutionary Data

Bogdan Efimenko

Bogdan Efimenko, Konstantin Popadin, Konstantin Gunbin

Center for Mitochondrial Functional Genomics, Immanuel Kant Baltic Federal University, Center for Mitochondrial Functional Genomics, Immanuel Kant Baltic Federal University, Center for Mitochondrial Functional Genomics, Immanuel Kant Baltic Federal University

One of the most important characteristics of each contemporary model of molecular evolution is the assumption that mutations occur in

a constant manner. However, in nature mutations are determined by the interaction of DNA replication and repair efficiency, which are highly diverse because of environmental changes like exposure to mutagens or naturally selected changes in DNA repair. Although the importance of mutational processes is well-recognized in human cancer research, progress in the reconstruction and utilization of mutational spectra at comparative- or intra-species levels remains limited due to data scarcity. Traditional mutation accumulation experiments used to estimate neutral mutational spectra are not only laborious and generally confined to model organisms, but they also might not be entirely neutral, thus impeding progress in filling this knowledge gap. To address this shortfall, we present the NeMu pipeline, available at https://nemu-pipeline.streamlit.app. By harnessing the vast amount of sequences from existing databases, NeMu efficiently and accurately reconstructs the neutral mutational spectrum. The pipeline encompasses multiple steps: automated sequence sampling from nucleotide databases, phylogeny and precise ancestral states reconstruction, extraction of substitutions and inference of the neutral spectrum. This efficient tool enables the identification of species-specific mutational signatures, advancing research into species-specific spectrum variation and allowing in future for their integration into more realistic evolutionary models for understanding drivers of evolution.

This project was supported by the Russian Science Foundation (RSF) grant â„– 21-75-20143

818 - GHIST: the Genomic History Inference Strategies Tournament

Ryan N Gutenkunst

Travis J. Struck, Katie E Lotterhos, AndrÃ©s Moreno-Estrada, Peter L Ralph, Adam Siepel, Ryan N Gutenkunst
University of Arizona, Northeastern University, CINVESTAV, University of Oregon, Cold Spring Harbor Laboratory, University of Arizona

In population genomics, there are a dizzying array of potential data analysis approaches to infer population history, aspects of natural selection, or other evolutionary properties from data. Although methods developers try to evaluate their approaches, those evaluations can be unconsciously biased or may not reflect the experiences of real-world users. GHIST aims to create an annual forum in which the community can test inference approaches in an unbiased fashion. Each year, the GHIST organizers will release simulated population genomic data sets and host a competition to infer various aspects of the processes that generated those data. From the competitorsâ€™ solutions, the community will learn what approaches perform well or poorly in particular circumstances. Our first competition is now live, with four demographic history inference tasks of escalating expected difficulty. We look forward to your submissions!

850 - Population genetic consequences and causes of mutation rate heterogeneity in Arabidopsis thaliana

Mariele Elizabeth Lensink

Mariele E Lensink, Daniel Kliebenstein, Grey Monroe
UC Davis, UC Davis, UC Davis

Apparent mutation rates are known to vary across organism genomes due to differential damage, repair, and selection. We set out to study this variation through the lens of population genetics in two ways. First, we developed a novel evolutionary algorithm to estimate the relative contributions of mutation rate and selection heterogeneity to shaping patterns of genomic variation observed in natural populations of Arabidopsis thaliana. Results from this model suggest that mutation rate heterogeneity is a significant driver of genetic variation in populations, which should inform future evolutionary models. We then tested if the predicted mutation rate and selection heterogeneity are consistent with classical mutation modifier theory, informed by the drift barrier hypothesis, mutation accumulation studies, and experimentally established mechanisms of differential DNA repair activity across the Arabidopsis genome. These analyses established the potential for natural selection to act on mechanisms by which DNA repair proteins are recruited to histone modifications

enriched in gene bodies. In summary, population genomics analysis and theory provide multiple frameworks to study the causes and consequences of intragenomic mutation rate variation with implications for understanding the mutational origins of genetic diversity.

855 - Extending Flex-sweep with Random Forest to infer time and strength of selection

Jesus Murga-Moreno

Jesus Murga-Moreno, David Enard
University of Arizona, University of Arizona

The contribution of selective sweeps to genomic adaptation remains an active area of research in population genetics. Machine learning (ML) has recently surpassed most typical approaches, such as classical summary statistics, maximum likelihood, and ABC. Although many existing methods can detect selective sweeps, most are statistically limited to specific sweep types, and subtle signals are particularly challenging to detect. One method, Flex-sweep, is a Swiss-knife army that detects sweeps of different types, strengths, and ages, allowing us to better discern past selective pressures such as ancient epidemics. Although Flex-sweep and other ML approaches can detect and classify sweeps, only a few methods can quantify the strength and timing of selection. We extend Flex-sweep with Flex-date, a Random Forest approach to infer the strength and timing of selection. We exploit Flex-sweep's classifications, summary statistics, and theoretical differences between sweep to reduce the prior space parameter when performing the inferences. Because different summary statistics capture different sweep patterns, our method relies on a two-step approach: i) classify Flex-sweep outputs by age and ii) predict parameters through the selected sweep age model. By prioritizing each model's most important summary statistics, we manage the high-dimensional data and focus on the most informative statistics for further parameter estimation. Flex-date performs robust inferences on the strength and timing of selective sweeps covering ancient, complete and partial sweeps. We compared Flex-date to recent methods for inferring sweep parameters including ARG-based methods.

973 - Minus the error: making selection detection robust to residual alignment errors

Avery Selberg

Avery Selberg, Nathan Clark, Maria Chikina, Sergei L. Kosakovsky Pond
Temple University, University of Pittsburgh, University of Pittsburgh, Temple University

Advancements in computational methods for genomic data analyses are crucial for unraveling the evolutionary history and adaptation of diverse species, yet errors in multiple sequence alignments (MSAs) are known to bias many comparative methods. In the context of natural selection analyses, specifically codon evolutionary models, excessive rates of false positives result. A characteristic â€œsignatureâ€• of error-driven findings is unrealistically high estimates of dN/dS (e.g., >100), affecting only a small fraction (e.g., ~0.1%) of the alignment. Despite the widespread use of codon models to assess alignment quality, their potential for error correction remains unexplored. We present BUSTED-E (â€œminus the errorâ€•): a novel method designed to detect positive selection while concurrently identifying alignment errors. This method is a straightforward adaptation of the BUSTED flexible branch-site random effects model used to fit distributions of dN/dS, with an important modification: it integrates an error-sink component representing an â€œabiologicalâ€• evolutionary regime (dN/dS > 100). With the error-sink component, the need for alignment pre-filtering is eliminated. In large-scale analyses, BUSTED-E reduces unrealistic rates of positive selection detection and provides the option for masking errors in the MSA for downstream analyses. We conducted a genome-wide scan of avian and mammalian genes utilizing BUSTED-E, resulting in a substantial decrease in the percentage of positively selected genes (from 40% to 4%), which were notably enriched in immune-related functionsâ€”highlighting BUSTED-Eâ€™s ability to detect meaningful biological patterns even in the presence of alignment errors.

986 - Integrating epidemiological and population genetic models to gain insights into Plasmodium vivax biology

Shyamalika Gopalan

Shyamalika Gopalan, Jillian Grassia, Amy Goldberg

Clemson University, Duke University, Duke University, Duke University

Genomic data are increasingly being generated to understand the evolution and epidemiology of human pathogens. Despite progress in studies of non-recombining viruses and bacteria, gleaning actionable insights from eukaryotic parasite data remains difficult, largely due to their complex life cycles. Here, we develop a novel framework for jointly modeling the epidemiological dynamics and genomic evolution of the malaria parasite Plasmodium vivax. We build a two-step model across the parasiteâ€™s haploid and diploid stages within the human and mosquito, respectively. First, we use a stochastic Ross-Macdonald-style model to simulate a series of human-mosquito transmission events under a given set of epidemiological parameters. Then, we use SLiM to simulate the within-human and within-mosquito stages of parasite genome evolution as they recapitulate these transmission events, including recurrent transmission bottlenecks and periods of exponential growth. By capturing the processes generating neutral genetic variation in parasites, we are able to explain empirically observed phenomena, test the relative contributions of life stages to summaries of variation, and provide neutral baselines for common summary statistics. For example, for parasite populations thought to be at constant population size, most empirical studies observe strongly negative genome-wide average Tajimaâ€™s D values. We show that, unlike for human populations, this is an expected feature under a neutral model of P. vivax evolution. Our findings have important implications for the interpretation of summary statistics and overall patterns of genetic variation used for the inference of demography and selection in eukaryotic pathogens.

1057 - MSS: A codon model with Multiple classes of Synonymous Substitutions.

Hannah Elaine Verdonk

Hannah E Verdonk, Jody Hey, Alyssa Pivirotto, Vitor Pavinato, Sergei LK Pond

Temple University, Temple University, Temple University, Temple University, Temple University

A fundamental goal of evolutionary biology is to identify drivers of natural selection. Existing codon evolutionary models examine selection on genes by comparing the rates of synonymous (dS) and nonsynonymous substitutions (dN). However, such models overlook the nuanced variation in the relative rates for individual synonymous substitutions, driven by selection on codon usage. Crucially, the relative rates are known to correlate with key biological traits like gene expression, and can inform researchers about the underlying factors driving selection. To explore selection on codon usage, we introduce the Multiclass Synonymous Substitution (MSS) model, a novel codon evolutionary model designed to separate the monolithic synonymous rate class into its individual components. The MSS model estimates the relative rate for each pair of synonymous codons, enabling precise comparison of synonymous substitution rates within a gene or across a set of genes. Ranking synonymous pairs by their relative rates allows us to explore which pairs are under selection (i.e., those with a substitution rate <1). We applied the MSS model to several well-studied genuses with prior evidence of selection on codon usage. The MSS ranking of synonymous substitution rates in Drosophila aligns with average SNP density across codon pairs, an independent measure of selection on codon usage. Our model also recapitulates different substitution patterns in highly divergent taxonomic groups, such as Enterobacteria, Caenorhabditis, and Saccharomyces. With the MSS model, we can now study selection on synonymous substitutions in diverse taxa, independent of any a priori assumptions about the forces driving that selection.

443 - Spatial and temporal analyses identify two introgression events between brown and American black bears

Emily E. Puckett

Emily E. Puckett
University of Memphis

Brown bears (Ursus arctos) colonized North America from Eurasia in two distinct and temporally separated waves. Once in North America they encountered endemic American black bears (U. americanus) during range expansions from eastern Beringia southwards into the interior of the continent. The establishment of sympatry between these species provided the opportunity for hybridization and introgression, which was previously identified using D-statistics. Both species have broad spatial ranges that should limit the extent of introgression, such that it is found primarily between sympatric populations. I used whole genome sequencing from samples collected across the ranges of both bear species to test for spatial variability in introgression. I identified two pulses of introgression between brown and American black bears, and demonstrate the introgressed segments occur across spatially structured lineages in both species. The first pulse occurred 240 â€“ 130kya, near the estimated intraspecific divergence times of both species. This pulse was also estimated to occur prior to the establishment of sympatry in western North America, and thus implicates another lineage mediating gene flow between Eurasian brown bears and North American black bears. The second pulse occurred only between western black bears and North American brown bears. Introgressed genomic segments occur in both species, yet North American brown bears contain 2-4% of black bear segments, compared to 0.2-0.8% of brown bear segments in American black bears. This asymmetry may be linked to range expansion of brown bears across North America, but may also implicate adaptive introgression.

896 - Reconstructing the recent demography of Canada's beluga populations to inform conservation in light of the warming Arctic

Claudio Müller

Claudio MÃ¼ller, Evelien deGreef, Colin Garroway
University of Manitoba, University of Manitoba, University of Manitoba

The Arctic is the fastest-warming region on the planet. As such, ice-adapted species face considerable threats associated with current and forecasted levels of climate warming. In this study, we provide a holistic overview of recent demographic changes in Canadaâ€™s beluga populations. Belugas are adapted to the cold waters of the Arctic, with thick skin and blubber, with a dependency on ice for protection and annual migration, dictated by ice formation and break up. Using 140 whole genomes from individuals sampled across North Americaâ€™s east coast, we find significant levels of population structure which reflects their wintering ground more strongly than their summer habitats. Deep-time demographic history reconstruction reveals these populations splitting from one large population around the peak of the last ice age. While these populations have all had largely stable effective population size for most of their existence, reconstruction of their recent demographic history shows them all peaking around 75 generations ago, after which they start to decline. The small populations in Cumberland Sound, James Bay and St. Lawrence show very pronounced decrease in numbers around that time. Current effective population size estimates are very low in all but one of these populations. This might be the result of increased interbreeding with a neighboring population during the winter. Recent declines in effective population size suggest that adapting to ongoing climate warming may be challenging.

148 - Origin and diversity of the wild cottons (Gossypium hirsutum) of Mound Key, Florida

Weixuan Ning

Weixuan Ning, Karen M Rogers, Chuanyu Hsu, Zenaida V Magbanua, Olga Pechanova, Mark A Arick II, Ehsan Kayal, Guanjing Hu, Daniel

G Peterson, Joshua A Udall, Corrinne E Grover, Jonathan F Wendel
Iowa State University, Florida Department of Environmental Protection, Mississippi State University, Mississippi State University, Mississippi State University, Mississippi State University, Iowa State University, Chinese Academy of Agricultural Sciences, Agricultural Genomics Institute at Shenzhen, Mississippi State University, USDA/Agricultural Research Service, Iowa State University, Iowa State University

Deciphering genetic diversity within wild forms of modern crops is essential for understanding domestication and the possibilities of wild germplasm utilization. Gossypium hirsutum is a predominant source of natural plant fibers and the most widely cultivated cotton species. Wild forms of G. hirsutum are challenging to distinguish from feral derivatives, and truly wild populations are uncommon. Here we characterize a population from Mound Key Archaeological State Park, Florida using genome-wide SNPs extracted from 25 individuals over three sites. Our results reveal that this population is genetically dissimilar from other known wild, landrace, and domesticated cottons, and likely represents a pocket of previously unrecognized wild genetic diversity. Furthermore, our results suggest this population may be truly wild, refuting the notion that it could be a naturalized derivative of cultivars used by the Calusa Native Americans, the former inhabitants of Mound Key before European colonization. The unexpected level of divergence between the Mound Key population and other wild cottons suggests that the species may harbor other relictual and genetically distinct populations that are geographically scattered in suitable habitats throughout the Caribbean. Our work thus has broader conservation genetic implications and suggests that further exploration of natural diversity in this species is warranted.

226 - Unraveling the evolutionary history of indica rice post domestication

Yu-Chen Lin

Yu-Chen Lin, Michael Purugganan
New York University, New York University

Asian rice comprises two primary subspecies: japonica and indica, and it is cultivated across a broad latitudinal range and diverse local ecosystems. Archaeological and population genomic evidence indicates that indica rice underwent domestication in the Ganges Valley approximately four millennia ago, potentially spreading to Southeast Asia and China through introgression with japonica rice. A previous study revealed that the dispersal patterns of indica rice remain puzzling, largely due to widespread gene flows and a weak population structure. In this study, we employed an extensive sampling of indica rice landraces to scrutinize the population structure and demographic history of indica rice post-domestication. Our preliminary findings highlight a robust population structure linked to geographical origins, with admixture analysis suggesting two admixture events. Inferring selection and allele frequency trajectories through genomic data of the geographic subpopulations promises to offer insights into the evolutionary history of indica rice.

318 - A uniform bacterial community inhabits the rhizosphere across cultivars of the common bean (Phaseolus vulgaris).

Víctor González

Victor GonzÃ¡lez, Griselda LÃ³pez-Romo, Rosa I SantamarÃ-a, Patricia Bustos, Francisco EcheverrÃ-a, Luis Roberto Reveles-Torres
Centro de Ciencias GenÃ³micas-UNAM, Centro de Ciencias GenÃ³micas, Centro de Ciencias GenÃ³micas, Centro de Ciencias GenÃ³micas, Centro de Ciencias GenÃ³micas, INIFAP, INIFAP

The success of common bean (Phaseolus vulgaris) crops relies, among other factors, on the microbial species inhabiting the rhizosphere. Although still inadequately characterized, beyond the established role of nitrogen-fixing bacteria, current metagenomic techniques allow

us to catalog a large number of microorganisms hosted in the rhizosphere. Here, we studied the bacterial communities in the rhizosphere of nine common bean cultivars used in agriculture. Bulk soil and rhizosphere samples from P. vulgaris cultivars were collected in situ from plots with and without a cultivation history. In both plots, bacterial diversity in the bulk soil exceeded that in the rhizosphere. Diversity and taxonomic composition analyses confirmed the dominance of Proteobacteria in the rhizosphere. Furthermore, a uniformly distributed bacterial community encompassing various cultivars was present in the rhizosphere. The results of comparing bulk soil-rhizosphere metagenomes of various cultivated plants, including maize, wheat, tomato, cucumber, Arabidopsis, and common bean cultivars, showed a distinct rhizosphere effect for the latter. Specifically, common bean cultivars exhibited a different pattern in the presence of bacterial genera known to promote plant growth, such as Rhizobium, compared to the other plants studied. These findings suggest that the rhizosphere effect of common bean cultivars may be unique and deserve further investigation. Metagenome-assembled genomes (MAGs) reconstructed from metagenomes confirmed the presence of diverse biosynthetic gene clusters that are involved in defense, signaling, and plant growth promotion. The results underscore the significance of the prior soil microbiome for the subsequent cultivars in facilitating bacterial colonization of the rhizosphere.

356 - Unraveling the population structure and origins of pigmented rice

Yoon Kyung Lee

Yoon Kyung Lee, Nese Sreenivasulu, Rafal M. Gutaker

Royal Botanic Gardens, Kew, International Rice Research Institute, Royal Botanic Gardens, Kew

Unpolished rice (Oryza sativa L.) provides nutritional benefits such as dietary fiber, essential vitamins, and minerals. The coloration in the rice pericarp, caused by the accumulation of anthocyanins and proanthocyanidins, not only adds to the nutritional value, but also has cultural significance. Aside from the rice domestication gene Rc, the mutation which led to the emergence of white rice, the subsequent origins of pigmented rices with various range of hues from red to purple, their evolutionary history and distribution is underexplored. To shed light on this, we conducted a comprehensive population structure analysis using genomic data from a diverse panel of 360 pigmented and unpigmented rice varieties. The study determines whether pigmented rice varieties share a common ancestry or if they have evolved independently in different geographic regions of Asia. To achieve this, Principal Component Analysis projection was employed, over the panel composed of three representative subgroups, including indica, tropical japonica, and subtropical japonica varieties. Further studies on the genetic diversity among different subpopulations may indicate adaptation to different environments. Our findings are expected to provide insights into the intricate relationship among genetic diversity, environmental factors, and human selection pressures. These insights will inform future breeding strategies aimed at achieving desirable nutritional values.

358 - Genetic diversity of Northeast Asian rice and selection signatures for eating quality trait

Yoon Kyung Lee

Yoon Kyung Lee, Hee-Jong Koh, Rafal M. Gutaker

Royal Botanic Gardens, Kew, Seoul National University, Royal Botanic Gardens, Kew

Understanding the genetic basis of eating quality traits in rice (Oryza sativa L.) is critical for breeding programs aimed at improving consumer satisfaction. In this study, we investigated the genetic diversity and selection signatures associated with eating quality traits in a diverse panel of 284 rice accessions from South Korea, Japan, China, Taiwan, and USA, comprising 171 improved varieties and 113 landraces. Population structure analysis revealed that most accessions from this panel belonged to the temperate japonica subgroup. Further subdivision within this population revealed three distinct clusters: landraces, improved cultivars, and recently bred improved cultivars. Phenotypic evaluation was conducted to characterize the eating quality trait by measuring the glossiness of cooked rice using

the Toyo taste meter. Subsequently, a selection scan for genetic determinants associated with eating quality traits was performed using XP-CLR, focusing on 25 accessions exhibiting high eating quality and 25 with low eating quality. The study enhances our understanding of the genetic architecture of grain quality traits in japonica rice. It also highlights the complexities and nuances encapsulated within varietal diversity. The study identifies genomic regions with molecular signatures for selection associated with superior eating quality, which can inform targeted breeding strategies to enhance these attributes.

390 - Tracing Domestication History in Potato: Admixture and Ancestry Inference in a Polyploid Species

Sergio Tusso

Sergio Tusso, Korbinian Schneeberger
LMU, LMU

Understanding the history of crop domestication is fundamental to agricultural advancement and biodiversity conservation. In this project, we delve into the complex and unclear domestication history of potato in the Andean region of South America. The domestication of potato likely involved admixture from multiple species spanning the geographic range from Mexico to Argentina, during initial domestication 10000 years ago and subsequent plant breeding programs in more recent times. However, given the tetraploid nature of potato, traditional methods face significant challenges in the genomic inference of ancestral admixture. To address these challenges, we present a novel approach utilizing an extended dataset comprising resequencing data from 840 samples, including globally distributed cultivars and 84 wild and landrace species. Leveraging linkage information, we identify ancestral groups within cultivar samples, revealing clear introgressed blocks from multiple species and variation along the genome, indicative of a complex evolutionary history. Surprisingly, despite the evident admixture from multiple species, we also observe markedly reduced haplotype diversity, which has important implications for potato agricultural research and biodiversity conservation efforts. By demonstrating the application of linkage disequilibrium for ancestry inference, our study not only provides novel insights into the intricate history of potato, but also presents a potential alternative to for admixture inference of other polyploid species.

410 - Human and ant-domesticated plants harbour different disease resistance gene repertoires compared to their wild relatives

Noah Bourne

Noah Guy Bourne, Guillaume Chomicki
University of Sheffield, Durham University, Durham University

Plant domestication is a pivotal innovation driving the success of human civilizations; however, the domestication process often includes unintended phenotypic and or genomic consequences for the plants involved. These include susceptibility to diseases in which their wild relatives are resistant. Convergently ant-domesticated epiphytes are susceptible to disease upon ant-exclusion compared to their non-farmed (wild analogues) relatives. The genomic basis of this apparent discrepancy is linked to domesticated plants having differing resistance gene (R-gene) repertoires to their respective wild relatives. Although due to the lack of consistent annotations in previous studies it is unclear whether domestication repeatedly leads to reduced R-gene repertoires and thus unequal disease resistance capability across the phylogeny of domesticated plants. Pan-genome approaches highlight that many R-genes in plants exist in the auxiliary portion of the genome with their diversity being maintained in a population by strong balancing selection which is often negated during domestication. This provides an evolutionary explanation to why reduced R-gene repertoires persist in crop plants. Using high quality genomes of crop plants and their wild relatives, including ant-domesticated Squamellaria and their non-farmed relatives, we use state of

the art deep learning approaches to provide consistent annotations of R-gene repertoires. Among the 15 crop/wild relative pairs covering 10 plant families we see significant differences in the repertoire of specific R-gene types. Additionally, the inclusion of ant-domesticated Squamellaria provides an exceptionally unique insight into the evolution of domesticated plants over a larger timescales.

464 - Domestication is nothing new: Maize domestication leveraged standing genetic variation

Regina Anne Fairbanks

Regina A Fairbanks, Jeffrey Ross-Ibarra

University of California, Davis, University of California, Davis

Domestication often involved radical morphological changes as plants adapted to novel anthropogenic environments. Understanding the origin of genetic variants underlying these changes has long been of interest. Despite the dramatic morphological differences between maize and its wild relative teosinte, research on a handful of well-described genes has found that maize domestication relied on standing genetic variation in ancient teosinte. But, researchers had concluded that one key trait and the genetic variation that causes it â€“ exposed kernels and a mutation in tga1 â€“ could exist only in cultivated populations. If true, the causative mutation in tga1 must have appeared de novo in cultivated populations. However, the puzzling absence of tga1 in many genome-wide scans for selection, typically well-powered for strong selection on de novo mutations, prompted us to consider that the causative mutation in tga1 actually predates domestication. Leveraging a large dataset of maize and teosinte genomes, we first identified the maize allele of tga1 in modern teosinte, suggesting that the purportedly deleterious allele can persist in the wild as standing genetic variation. Second, we conducted population genetic simulations and found that haplotype diversity patterns and distribution of mutational ages near tga1 can be best explained by selection on standing variation. Finally, we estimated the mutational age of the causative allele in tga1, finding that it predates even conservatively early estimates for the beginning of maize domestication. We conclude that, in contrast to several decades of thought, maize domestication drew upon standing genetic variation in tga1.

508 - Evolutionary genomics of historical varieties of Vitis vinifera in Mexico

Diana Margarita Mojica Muñoz

Diana Margarita Mojica MuÃ±oz, Luis Delaye, Antonio HernÃ¡ndez LÃ³pez

Departamento de IngenierÃ-a GenÃ©tica, Cinvestav Irapuato, Departamento de IngenierÃ-a GenÃ©tica, Cinvestav Irapuato, ENES Unidad LeÃ³n

The grapevineâ€™s cultivars in America dates back to Spanish conquest of the Aztec Empire. Thanks to the demand of wine for evangelization purposes, vineyards spread with the establishment of missions throughout America. In Baja California, Mexico, the most important viticulture region in the country, historical vineyards still prevail, which have shown good yield despite the conditions of climate change. However, it is unknown which varieties remain in these vineyards. In order to understand the origin of grapevines introduced by the Spaniards during the conquest, whose descendants are still cultivated in Baja California, four samples (E1, E11, E12 and E15) collected from historical vineyards in Baja California were sequenced and compared with whole-genome sequences previously publicated of wild vitis species and domesticated grapevines from Europe, Asia, and North America. We realize a phylogenetic analyzis, a population structure analyzis and the gene flow using the Patterson's D statistic was evaluated. We observed that the analyzed samples from vineyards in Baja California do not show any relationship with wild species in North America. Additionally, sample E12, whose closest identity is the Rosa del Peru variety, had evidence of introgression from Muscat of Alexandria and Listan Prieto varieties, confirming the hybridization antecedent between these two varieties. Furthermore, we identify sample E11 as a posible Rosa del Peru variety, because we find high similarity with sample E12. The identification of grapevines in historical vineyards is important for the maintainement of gene

pool, because there can be found important agronomic traits for the adaptation.

582 - Population genetic analysis reveals the domestication and dispersal history of common buckwheat

Jeffrey Fawcett

Jeffrey A Fawcett, Euki Yazaki, Yang Liu, Harriet V Hunt, Ryoma Takeshma, Mariko Ueno, Takanori Ohsako, Yumei Dong, Meifang Li, Hideki Hirakawa, Tatsuya Ota, Chengyun Li, Martin K Jones, Yasuo Yasui

RIKEN iTHEMS, RIKEN iTHEMS, University of Cambridge, University of Cambride, NARO, Kyoto University, Kyoto Prefectural University, Yunnan Agricultural University, Yunnan Agricultural University, Yunnan Agricultural University, Kazusa DNA Research Institute, SOKENDAI, Yunnan Agricultural University, University of Cambridge, Kyoto University

Common buckwheat (Fagopyrum esculentum) is a rare example of a crop domesticated in southwest China, outside of major domestication centres, and spread widely across Eurasia likely over a thousand years ago. Here, we first sequenced and constructed a chromosome-scale reference genome of a selfing accession of common buckwheat. We then resequenced a few hundred accessions including ~50 of the wild progenitor species and performed various population genetic analyses. We found that a wild population from southeast Tibet is most closely related to all cultivated accessions and observed genetic differentiation between cultivated accessions from different geographic regions. We also identified a number of selective sweeps involved in the domestication and dispersal processes, including one strong selective sweep shared by all cultivated accessions which likely originated from standing variation within the wild population from southeast Tibet. Our results suggest that all cultivated common buckwheat accessions originated from a wild population in southeast Tibet and allow us to discuss the domestication and dispersal history of common buckwheat.

760 - Composition of microbiota of floral nectar in wild and domesticated squash (Cucurbita)

Luis Alberto Villanueva-Espino

Luis Alberto Villanueva-Espino, Irais Avila Eulogio, Violeta PatiÃ±o Conde, Rafael Lira Saade, Eric Fuchs Castillo, Mauricio Quesada AvendaÃ±o

Escuela Nacional de Estudios Superiores, UNAM, Campus Morelia, Instituto de investigaciones en ecosistemas y sustentabilidad, Escuela Nacional de Estudios Superiores, UNAM, Campus Morelia, Facultad de Estudios Superiores Iztacala, Universidad de Costa Rica, Escuela Nacional de Estudios Superiores, UNAM, Campus Morelia

Floral nectar is the most common floral reward in plant-pollinator interactions. Due to its high energy and nutritional content, it is also a favorable habitat for the growth of many microorganisms, mainly yeasts and bacteria. These microbes can significantly modify the chemical composition of nectar, affecting pollinator behavior and plant reproductive success. The effects of this tripartite relationship (plant-pollinator-microbiota) in the Cucurbita genus are addressed to understand domestication's impact on this triangle. Using a metabarcoding approach, we determine the composition of bacteria and fungi in the floral nectar of six species of plants of the genus Cucurbita, three wild and three domesticated. We compared the differences between species, the floral sexes, and the effect of pollinator visitation on the diversity of the microbiota of these squash species.

In general, we observed a higher richness of bacteria than fungi in the floral nectar microbiota of the studied species of Cucurbita. For bacteria, alpha diversity shows differences between species but not in floral sexes, visited vs. not visited flowers, and wild and domesticated species. For fungi, however, this parameter does not show any differences. The beta diversity for both groups indicates differences between species, floral sexes, and visited vs. not visited flowers, but not by wild and domesticated species.

This study shows that domestication does not affect the microbial community in floral nectar. However, it provides evidence of how the

microbiota varies based on other factors, including species, floral sex, and pollinator visitation.

775 - The post-domestication history of rice

Ornob Alam

Ornob Alam, Michael Purugganan, Rafal Gutaker
New York University, New York University, Royal Botanic Gardens, Kew

The dispersal of rice (Oryza sativa) following domestication influenced massive social and cultural changes across South, East, and Southeast (SE) Asia. While recent resequencing of traditional rice landraces from across Asia have revealed possible routes of dispersal across Southeast Asia, the early history and spread of rice across and out of China remain largely unresolved. This is mainly because of sparse sampling of traditional landraces of japonica rice â€" the ancestral rice domesticate â€" from China as these have largely been replaced by modern improved or hybrid varieties and indica rice. We addressed this dearth of sampling by collecting historical specimens of domesticated rice from herbaria in France, the UK, and the US, originally sampled from China and other parts of Asia by botanists in the 19th and early 20th century. Here, we performed whole-genome resequencing of 96 herbarium rice specimens and analyzed them alongside a previously published set of 367 japonica rice landraces. We used ancestral recombination graph-based approaches to reconstruct the demography and routes of dispersal of japonica rice following domestication in China. We also utilized ancestral recombination graphs to investigate the evolutionary history of adaptive and functional variants â€" including those involved in flowering time, immunity, and tasteâ€" in populations of japonica rice, revealing widespread local adaptation as rice spread to different regions, climates, and latitudes.

1035 - Population genetics of year-long bean (Phaseolus dumosus Macfady, Fabaceae) in Southern Mexico and Guatemala

Katya Garduño Obrajero

Katya GarduÃ±o, Azalea Guerra Garcia
UNAM, Cinvestav Unidad Irapuato

One of the crucial processes that allowed humans to transition from nomadic to sedentary lifestyles was the development of agriculture. This process began simultaneously around 11,000 years ago during the Neolithic period in different parts of the world and Mesoamerica is one of the centers of domestication. Among the plant species domesticated in this region, are five bean species, which are an important component of the diet in Mexico and are considered one of the seven basic and strategic crops for the country. The yearâ€"long bean (Phaseolus dumosus) is a widespread crop in the Southeast regions of Mexico and Central America. Its wild populations only occur in the highlands of Guatemala, where domestication presumably started. In this study, we aim to understand the population structure, diversity, and genetic differentiation of 8 populations of P. dumosus (six cultivated and two wild populations) by using genomic information from 43 individuals. After filtering, 19,321 SNPs were kept and 4 genetic groups were identified: three cultivated genetic clusters (Veracruz, Chiapas and Guatemala), and one wild cluster (Guatemala highlands). The highest genetic diversity was found in the cultivated population from Guatemala, meanwhile, the lowest was detected in the cultivars from the Trans-Mexican Volcanic Belt. The genetic differentiation was unexpectedly high when comparing the cultivars from Guatemala to the rest of the populations, including the wild one. This constitutes the first study at a population level that assesses the genetic diversity of the orphan crop bean species.

1093 - Fruit color variation in date palms (Phoenix dactylifera L.) and their wild relatives is

determined by an ancient trans-specific polymorphism and parallel degradation of anthocyanin pathway genes

Jonathan M. Flowers

Jonathan Flowers
New York University Abu Dhabi

Trait variation in crop species is often controlled by complex genetic and developmental networks and reconstructing their evolutionary origins is often complicated. Pigmentation traits have offered some of the most insightful examples into the origins of trait diversity and continue to yield unprecedented insight into genetic basis of evolutionary parallelism and convergence in crops and other organisms. Fruit color in cultivated date palms (Phoenix dactylifera L.) varies from red to yellow within species and similarly differs among its crop wild relatives. Here we describe the genetic basis and evolutionary origin of this trait across the Phoenix genus. We report an ancient origin of a previously reported causal LTR-retrotransposon insertion in an R2R3-MYB activator of the anthocyanin pathway. We estimate the age of the insertion and find it constitutes an ancient trans-specific polymorphism shared by divergent species. Fruit color variation between species appears to be caused both by this ancient insertion and parallel loss-of-function mutations in anthocyanin pathway genes. Our results suggest that both ancestral polymorphism and parallel evolution at minimally pleiotropic loci explain fruit color diversity in this genus of palms.

1101 - Novel insights on soybean domestication (Glycine max)

Luis B Gómez Luciano

Luis B GÃ³mez Luciano
Donald Danforth Plant Science Center

The domestication of soybean (Glycine max) is estimated to occur in East Asia around 6000-9000 years ago, most likely from its wild relative species G. soja. Although most studies accept that soybean originated in China; archaeological and genotyping studies place other regions as potential center of origin. Besides a single event hypothesis, there is the line of thought that considers a more complex process involving multiple independent events in soybean domestication. To test the center of origin of soybean and subsequent spreading of the crop to current cultivated regions, in this study, we have analyzed genomic sequences of 3,829 accessions of soybean landraces and wild soja representing natural and cultivated habitats. Including sequences of the genus Glycine as outgroup species, we found novel insights on the population transition from wild G. soja to domesticated soybean and present a fresh understanding on the process of soybean domestication and spreading. In addition, we have tested for gene flow among populations of both, domesticated and wild species.

66 - The evolutionary correlates of viral host jumps

Cedric C.S. Tan

Cedric CS Tan, Lucy van Dorp, Francois Balloux
UCL Genetics Institute, University College London, UCL Genetics Institute, University College London, UCL Genetics Institute, University College London

Most emerging and re-emerging infectious diseases stem from viruses that naturally circulate in non-human vertebrates. When these viruses cross over into humans, they can cause disease outbreaks, epidemics and pandemics. While zoonotic host jumps have been extensively studied from an ecological perspective, little attention has gone into characterizing the evolutionary drivers and correlates

underlying these events. To address this gap, we harnessed the entirety of publicly available viral genomic data, employing a comprehensive suite of network and phylogenetic analyses to investigate the evolutionary mechanisms underpinning recent viral host jumps. Surprisingly, we find that humans are as much a source as a sink for viral spillover events, insofar as we infer more viral host jumps from humans to other animals than from animals to humans. Moreover, we demonstrate heightened evolution in viral lineages that involve putative host jumps. We further observe that the extent of adaptation associated with a host jump is lower for viruses with broader host ranges. Finally, we show that the genomic targets of natural selection associated with host jumps vary across different viral families, with either structural or auxiliary genes being the prime targets of selection. Collectively, our results illuminate some ofthe evolutionary correlates of viral host jumps that may contribute to mitigating viral threats across species boundaries.

69 - Characterizing Pareto fronts: Trade-offs in the yeast growth cycle constrain adaptation

Jason Tarkington

Jason Tarkington, Gavin Sherlock
Stanford, Stanford

Mapping trait space accessible by single mutations reveals the constraints imposed on organismal fitness. The fitness of an organism in an environment is often dependent on more than one phenotype, for yeast these include fermentation, respiration, and stationary phase performance. Such fitness-related phenotypes can be constrained due to the pleiotropic effect of other fitness related phenotypes, resulting in a trade-off. Previous work has shown that following evolution in a glucose containing carbon limited media a pareto front, which is indicative of an underlying trade-offs, emerges between stationary phase and respiration and between respiration and fermentation but, not between stationary phase and fermentation. I aim to understand why such trade-offs emerge among some fitness-related traits but not others by allowing barcoded yeast to evolve in a non-fermentable carbon source with varying amounts of time spent in stationary phase. This eliminates selection for fermentation entirely and creates a gradient of selection for respiration and stationary phase between the two-day transfer regime and the 10-day transfer. Under these conditions stationary phase performance may be free to increase unconstrained by fermentation performance resulting in the emergence of a Pareto front between these components of fitness. In addition to the phenotypic analysis, I have characterized the molecular basis of 480 underlying adaptive mutations that emerge to understand the mechanistic basis of trade-offs in the growth cycle and have identified distinct genomic signatures adaptation that depend on time spent in stationary phase.

96 - Functional consequences of reductive protein evolution in a minimal eukaryotic genome

Aaron W Reinke

Jason Jiang, Rui Qu, Maria Grigorescu, Winnie Zhao, Aaron W Reinke
University of Toronto, University of Toronto, University of Toronto, University of Toronto, University of Toronto, University of Toronto

Microsporidia are parasites with the smallest known eukaryotic genomes. The extent of protein loss in these organisms has been well documented, but much less is known about how compaction of microsporidia proteins affects their function. Taking a comparative genomic approach, we identified microsporidia orthologs of budding yeast proteins and show that these orthologs are enriched for essential yeast genes. We show that the median microsporidia protein is 21% shorter than its yeast counterpart and although extensive protein loss occurred after the divergence of microsporidia, reduced protein sizes were already present in microsporidian relatives. Microsporidia proteins are shorter through reduced domain lengths, diminished linker lengths, and domain loss, with 21% of microsporidia orthologs having lost domains present in yeast. On average, 34% of microsporidia orthologs have lost C-terminal residues essential for function in yeast, including 13 essential domains lost per genome. We also found that microsporidia display distinct phylogenetic patterns

of domain loss, with losses occurring in a clade-specific manner. To investigate conservation of function, we used yeast complementation assays to test orthologs from several microsporidia species and their relative Rozella allomycis. These experiments reveal that most microsporidia proteins cannot complement their yeast orthologs, the ability to complement is about three-fold less than observed for R. allomycis orthologs, and proteins that do not complement are more reduced in length than their yeast orthologs. Altogether, our results demonstrate the drastic reduction of microsporidia proteins and show that these reductions have resulted in functional divergence from their fungal ancestors.

189 - Molecular Mechanism of Vancomycin Resistance in Enterococci : Insights from Hospital Cases

Megha Suresh

Megha Suresh
Indian Institute of Science Education and Research, Berhampur, India

The emergence of pseudomembranous enterocolitis, and Methicillin-Resistant Staphylococcus aureus in the 1980s, popularized Vancomycin as a drug of choice. However, the increased use of vancomycin led to the development of Vancomycin-resistant Enterococci (VRE). Resistance is acquired through the action of multiple genes. The most studied mechanism is that of the vanA gene cluster which collectively modifies the target, reducing its affinity for vancomycin. In this study, we aimed to investigate the prevalence and molecular mechanisms of VRE in India.
To achieve this, I analyzed antibiogram data from two groups of hospitals. I also analyzed the whole genome sequence data of four VRE isolates to identify resistance genes and mutations in the genome. Our research uncovered a higher prevalence of VRE in males compared to females, with a noticeable increase in cases post-pandemic. Sequence analysis brought to light multiple PBP5 polymorphisms in the genomes. Documentation of horizontal gene transfer of PBP5 genes in E.faecium was noted, potentially facilitating the dissemination of resistance genes. Ongoing efforts involve the continued analysis of genomes and the assessment of phylogenetic relationships to unravel the evolving patterns. This study extends implications to diagnostics, interdisciplinary collaboration that bridges molecular and evolutionary perspectives, and progression of evolutionary medicine. Our findings contribute to the field of evolutionary medicine, emphasizing the need for targeted interventions in response to observed demographic and genomic trends.

244 - Dynamics and change in fitness effects of mutations through long-term bacterial evolution

Olivier Tenaillon

Olivier Tenaillon, Alaksh Choudhury, Zoya Dixit, MÃ©lanie Magnan
Institut Cochin, UniversitÃ© Paris CitÃ©, Inserm, CNRS, IAME, UniversitÃ© Paris CitÃ©, Inserm, Genoscope, CEA, Institut Cochin, UniversitÃ© Paris CitÃ©, Inserm, CNRS, Institut Cochin, UniversitÃ© Paris CitÃ©, Inserm, CNRS

The influence of beneficial mutations takes a center stage in the intricate dynamics of adaptation. Since the early days of population genetics, scientists have ascribed varying degrees of significance to these mutations and, by extension, their impact on populations genetic diversity. On the conservative end, proponents argue for the sporadic emergence of rare beneficial mutations exerting a transient influence on genetic diversity as they independently reach fixation. On the more assertive side, a contrasting viewpoint posits an abundance of highly impactful beneficial mutations that engage in competitive clonal interference, fostering the emergence of multiple

high-frequency polymorphisms.

To rigorously assess the quantitative contribution of beneficial mutations to population genetic diversity, we turned to the Long-Term Experimental Evolution. In this experiment 12 lineages of Escherichia coli evolved in the laboratory over a span exceeding 70,000 generations. Genetic barcodes were introduced to both the ancestral strain and clones from two populations sampled after 2,000 or 20,000 generations of evolution. The subsequent replay of evolution, spanning up to 600 generations with these barcoded libraries, enabled a comprehensive tracking of the emergence of beneficial mutations at low frequencies.

This methodology unveiled a diverse array of patterns, encapsulating the entire spectrum of models elucidated over a century of population genetics. Notably, the findings underscored the substantial impact of beneficial mutations on population genetic diversity throughout the long-term course of adaptive evolution, and pinpoints to the molecular determinants involved in the changes of the beneficial mutation rate in the course of evolution.

329 - Evolutionary dynamics and variation characteristics of SARS-CoV-2

Xiaolu Tang

Xiaolu Tang, Changcheng Wu, Ruochen Ying, Xinmin Yao, Ben Hu, Ran Liu, Yongsen Ruan, Mei Hou

Peking University, Peking University, Peking University, Peking University, Wuhan University, Wuhan University, Sun Yat-sen University, Sun Yat-sen University

In recent years, numerous viral diseases have erupted frequently worldwide, profoundly impacting human health, socioeconomic dynamics, and national security. Since the onset of the COVID-19 pandemic, the virus has persistently undergone mutations, diversification, and iterations, presenting formidable challenges to epidemic containment strategies and the advancement of vaccine research and development. In the early outbreak of this pandemic, based on two tightly linked SNPs, we previously divided SARS-CoV-2 into two major lineages: the ancestral â€œS lineageâ€• (U8,782 and C28,144) and the derived â€œL lineageâ€• (C8,782 and U28,144). Based on the severity of clinical symptoms in patients, it was found that patients infected with the S lineage exhibited more severe clinical symptoms. Building upon this, we further developed a hierarchical lineage classification system for the SARS-CoV-2, revealing its evolutionary trajectory. Through molecular evolutionary studies of the SARS-CoV-2 genome, it has been observed that the increasing number of infections has led to a significant acceleration in the evolutionary rate of the virus. Moreover, we found that the S gene, particularly the S1 region, has experienced positive selection in both SARS-CoV-2 and other coronaviruses. However, after the transmission of SARS-CoV-2 from animals to humans, the targets of positive selection shifted from the S1 N-terminal domain to the S1 C-terminal domain, which is the receptor-binding domain of SARS-CoV-2.

437 - The evolutionary origins of human ACE2 receptor binding among bat sarbecoviruses

Tyler N Starr

Tyler N Starr, Caroline J Craig, Ashley L Taylor

University of Utah, University of Utah, University of Utah

Many human viruses trace their origins to zoonoses from animal reservoirs, but we lack clear models for how and why animal viruses evolve molecular traits that facilitate human spillover. Such information would improve prediction of future zoonoses and development of therapeutic reagents with prospective breadth of coverage against the range of animal viruses that might seed future epidemics. SARS-related coronaviruses (also known as sarbecoviruses) circulate in Rhinolophus bats across Asia, Europe, and Africa, but two human spillovers in the past 20 yearsâ€”SARS-CoV-1 in 2003 and SARS-CoV-2 in 2019â€”have caused devastating outbreaks. We combined phylogenetic analysis with high-throughput biochemical binding assays to chart the evolution of ACE2 binding specificities among bat

sarbecoviruses, revealing that high-affinity binding to human ACE2 evolved as a latent trait in an ancestral bat sarbecovirus long preceding human zoonosis. Here, we seek to understand how and why these bat sarbecoviruses fortuitously acquired this trait that facilitates zoonosis. We characterized paths of evolution within ancestrally reconstructed spike receptor-binding domains via combinatorial mutagenesis. We identify the specific historical substitutions that enabled acquisition of human receptor binding and highlight the divergent genetic and biophysical mechanisms by which human ACE2 binding evolved convergently in bat sarbecoviruses from Asia and Europe. We present preliminary data supporting a model of bat:virus coevolution that explains why evolution of viruses in bats drives fortuitous binding to the human receptor. Together, this work identifies ecological and molecular evolutionary forces driving the acquisition of latent traits that enable zoonosis.

498 - Metagenomic characterization of the oral microbial diversity of Aboriginal Australians with whole genome resolution

Gabriel Yaxal Ponce Soto

Samuel Neuenschwander, Gabriel Yaxal Ponce Soto, Davide Bozzi, David M Lambert, Eske Willerslev, Anna-Sapfo Malaspinas, Nicolas Rascovan

Department of Computational Biology, University of Lausanne, Vital-IT, Swiss Institute of Bioinformatics, Microbial Paleogenomics Unit, Institut Pasteur, , Department of Computational Biology, University of Lausanne, Australian Research Centre for Human Evolution, Environmental Futures Research In-stitute, Griffith University, Lundbeck Foundation GeoGenetics Centre, Globe Institute, University of Copenhagen, GeoGenetics Group, Department of Zoology, University of Cambridge, Department of Computational Biology, University of Lausanne, Swiss Institute of Bioinformatics, Microbial Paleogenomics Unit, Institut Pasteur, CNRS UMR 2000

Most human microbiome studies have focused on urbanized populations with high European ancestry, leaving a gap of knowledge concerning non-Westernized or underrepresented populations, despite the growing evidence suggesting that they may hold a diversified and undescribed microbial diversity. In this study, we analyzed the salivary microbiome of 84 Aboriginal Australian individuals, a human population that is among the most deeply diverged from other human groups, and therefore a perfect case to explore variations in the oral microbial diversity. Using common read-based taxonomic analyses, we found that all individuals exhibited a similar oral microbial structure (i.e., taxonomic composition and alpha and beta diversity metrics), thus confirming observations from previous works on the homogeneity of the human oral microbiome, even in divergent human populations. We reconstructed 815 near complete metagenome-assembled genomes (MAGs), which were analyzed in combination with >6000 reference genomes and publicly available MAGs of oral origin from other worldwide locations. We clustered these genomes by 95% ANI (an approximation to differentiate species), obtaining 2505 genome clusters, with 38 clusters being exclusive to the Aboriginal Australian population. We used phylogenomic analyses to explore the diversity at strain level of 61 oral bacterial species that are ubiquitous and abundant in human populations and found 14 species in which strains from Aboriginal Australians were clustering together when compared to other worldwide strains, potentially reflecting a codivergent history. These findings highlight the relevance of studying underrepresented human populations and contribute a new set of model species to study host-microbial coevolution.

516 - The Influence of Lateral Transduction on Bacterial Genome Content and Structure.

Rebecca Man

Rebecca Man, Jamie Gorzynski, Natalie Ring, JosÃ© R PenadÃ©s, J Ross Fitzgerald

The Roslin Institute, University of Edinburgh, The Roslin Institute, University of Edinburgh, The Roslin Institute, University of Edinburgh, Imperial College London, The Roslin Institute, University of Edinburgh

Lateral transduction (LT) is a powerful form of horizontal gene transfer involving prophage-mediated transfer of large regions of the bacterial chromosomal downstream of integration sites via delayed prophage excision. Staphylococcus aureus, a major bacterial pathogen, has a strongly conserved core genome with phage integration (attB) sites that are distributed in a way that could enable around 60% of the S. aureus chromosome to undergo LT-mediated mobilisation. LT can both introduce new genes while also removing mildly deleterious mutations and parasitic mobile genetic elements via homologous recombination. Based on the distribution of attB sites and packaging directionality, the S. aureus genome can be differentiated into two regions: the attB-containing R1 region where LT activity occurs, and the attB-free R2 region predicted to be unaffected by LT. In order to examine if R1 and R2 have distinct sequence or chromosomal characteristics reflecting LT v no LT activity, we examined the distribution of pseudogenes, insertion sequences, core and accessory genes, and the genes under positive selection in R1 v R2 regions. A higher frequency of core genes and core genes under positive selection were found in R1. Conversely, pseudogenes, insertion sequences, accessory genes and accessory genes under positive selection were more frequently located in R2. These findings are consistent with the role of LT in removing deleterious mutations, parasitic elements, and maintaining gene conservation, with important impacts on S. aureus genome organisation and diversity.

554 - Epistasis decreases with increasing antibiotic pressure but not temperature

Ana-Hermina Ghenu

Ana-Hermina Ghenu, AndrÃ© Amado, Isabel Gordo, Claudia Bank

Instituto Gulbenkian de CiÃªncia, UniversitÃ¤t Bern, Swiss Institute of Bioinformatics, UniversitÃ¤t Bern, Swiss Institute of Bioinformatics, Instituto Gulbenkian de CiÃªncia, UniversitÃ¤t Bern, Swiss Institute of Bioinformatics, Instituto Gulbenkian de CiÃªncia

One Health challenges us to adopt a holistic approach in antimicrobial resistance (AMR) evolution by predicting the fitness effects of resistant mutants across various environmental and genetic contexts. However, accurate fitness predictions are hindered by strong genetic interactions, whether among genes (GxG, or epistatic) or between genes and the environment (GxE and epistasis-by-environment, or GxGxE). Rifampicin resistance evolution by Escherichia coli (B strain) depends on a ubiquitous environmental variable: high temperature can drive the evolution of resistance in the absence of antibiotic. To better understand how environment impacts genetic interactions, we quantified an empirical fitness landscape combining rifampicin resistance amino-acid substitutions with whole-gene deletion mutations, in E. coli K-12, across a complete combinatorial set of temperature and antibiotic dosage gradients. We found GxG interactions and rugged fitness landscapes in the absence of antibiotic, but as antibiotic concentration increased across all temperature environments, the fitness effects of AMR genotypes quickly overshadowed those of gene knock-outs, and the landscapes became smoother. Futhermore, we found that overall environmental quality was less important than antibiotic concentration alone in explaining the effect of environment on epistasis. Our work reiterates that some single mutants, like those conferring resistance or susceptibility to antibiotics, and some types of environments, like the concentration of an antibiotic, can have consistent effects across genetic mutations and environments. Nevertheless, the different effects of temperature for B versus K-12 backgrounds underscore the persistence of (Gx)GxE interactions and the importance of a holistic perspective in AMR evolution.

661 - Multi-omics reveals host-microbiome interaction in the oral cavity underlying obesity

Aashish R Jha

Ahmed A Shibl, Tsedenia W Denekew, Anique R Ahmad, Salah Abdelrazig, Christopher E Leonor, Lina Utenova, Guihao Zhang, Mamon AbdelBaqi, Yashaswi Malla, Mohammed Arshad, Marc Arnoux, Nizar Drou, Raghib Ali, Shady A Amin, The UAE Healthy Future Study Investigators Group Idaghdour, Youssef Idaghdour, Aashish R Jha

Genetic Heritage Group, Program in Biology, New York University Abu Dhabi, Genetic Heritage Group, Program in Biology, New York

University Abu Dhabi, Genetic Heritage Group, Program in Biology, New York University Abu Dhabi, Marine Microbiomics Lab, Program in Biology, New York University Abu Dhabi, Public Health Research Center, New York University Abu Dhabi, Genetic Heritage Group, Program in Biology, New York University Abu Dhabi, 1Genetic Heritage Group, Program in Biology, New York University Abu Dhabi, Public Health Research Center, New York University Abu Dhabi, Genetic Heritage Group, Program in Biology, New York University Abu Dhabi, Center for Genomics and Systems Biology, Bioinformatics Core, New York University Abu Dhabi, Core Technology Platform, New York University Abu Dhabi, 5Center for Genomics and Systems Biology, Bioinformatics Core, New York University Abu Dhabi, Public Health Research Center, New York University Abu Dhabi, Marine Microbiomics Lab, Program in Biology, New York University Abu Dhabi, Public Health Research Center, New York University Abu Dhabi, Center for Genomics and Systems Biology, Bioinformatics Core, New York University Abu Dhabi, Public Health Research Center, New York University Abu Dhabi, Center for Genomics and Systems Biology, Bioinformatics Core, New York University Abu Dhabi, Genetic Heritage Group, Program in Biology, New York University Abu Dhabi, Center for Genomics and Systems Biology, New York University Abu Dhabi, Public Health Research Center, New York University Abu Dhabi

Gut microbiome alterations have been causally linked to several human diseases, but whether perturbations of the oral microbiome â€“the second largest microbial ecosystem in the human bodyâ€“ contributes to human health remains markedly underexplored. Here, we first applied 16S rRNA sequencing to find associations oral bacteria and 14 cardiometabolic traits in a cohort of ~700 Emirati nationals from the United Arab Emirates. We found several features of the oral microbiota was strongly linked with obesity in this population. To elucidate the mechanism by which oral microbiota contributes to obesity, we implemented a multi-omics analysis framework in a case-control subset of ~200 individuals. Integrated analysis of metagenomics and metabolomics revealed functional alterations in the oral microbiome in obesity contributes to several hallmarks of obesity. Specifically, oral microbes in obese individuals lead to increased uridine production, which is associated with feeding behavior. They also metabolize dietary carbohydrates into lactate, which leads to increase in visceral fat deposit. Finally, they produce obesogenic amino acids, such as methionine, cystine, and, threonine. Conversely, several pathways associated with Vitamin-B and heme production is depleted in obesity. Furthermore, several of the bacterial elevated salivary metabolites were positively associated with several obesity associated clinical markers in blood, indicating these metabolites can transcend the well vascularized oral epithelia and interfere with host physiological process. This study robustly links oral microbiome alterations with host metabolic changes, indicating that mechanistic understanding of host-microbiome interactions in the oral cavity may provide a promising avenue for obesity intervention.

674 - High-throughput assays for testing compounds active against sequenced, multidrug-resistant Pseudomonas aeruginosa isolates from Mexico

Katia Pamela Villavicencio Sánchez

Katia Pamela Villavicencio SÃ¡nchez , Jonathan E GÃ³mez GonzÃ¡lez , MoisÃ©s A Alejo HernÃ¡ndez , Diana Corina CeapÄƒ
UNAM, UNAM, UNAM, UNAM

Antimicrobial resistance (AMR) has emerged as a major threat to global public health, with Pseudomonas aeruginosa being one of the most concerning multidrug-resistant (MDR) bacteria. The lack of surveillance legislation in Mexico has made it even more challenging to combat this issue. Therefore, the identification and characterization of new compounds with antimicrobial activity against MDR bacteria, particularly Pseudomonas, is crucial to developing effective treatments against such infections.

High-throughput assays, such as the microbial dilution assay established and validated by the Clinical and Laboratory Standards Institute, using ATCC strains as standards, are used to evaluate novel antimicrobial compounds and combinations. In this study, we have created a

library of MDR strains from clinical isolates obtained from health centers in Mexico City (CDMX) and Aguascalientes. Antimicrobial susceptibility profiles and Minimum Inhibitory Concentration (MIC) of the library strains were obtained through detailed analysis of antibiograms and phylogenetic trees. Using selected standards as well as clinical strains, we have tested lantipeptides synthesized at the Institute of Chemistry, UNAM, to identify promising antimicrobials against MDR strains.

Our work aims to contribute to the fight against AMR by discovering new antimicrobial agents. By using a genetically and phenotypically characterized MDR strain library, we can gain a better understanding of the epidemiology and evolution of antimicrobial resistance in the pathogens studied. These findings will pave the way for the development of effective treatments to combat MDR bacteria and reduce the threat of AMR to public health globally.

677 - Exploring Uncharted Biosynthetic Territories: A Genomic Data Mining Approach for Antimicrobial Discovery

Corina Diana Ceapa

Corina Diana Ceapa
Microbiology Laboratory (MicroIQ), Chemistry Institute, UNAM

Genomic data mining represents an exciting frontier with immense potential for rapid screening and uncovering leads in modern antimicrobial research. In this study, we leverage data mining techniques to delve into unexplored biosynthetic gene clusters (BGCs) and identify novel sources of bioactive molecules. Our rigorous selection criteria ensure the identification of potential game-changers in the fight against antimicrobial resistance.
We focus on discovering new chemical classes produced by non-pathogenic microorganisms devoid of immunity or virulence genes. This meticulous approach ensures the discovery of compounds with high therapeutic potential. Through cutting-edge in-silico studies of protein evolution, we have pinpointed clusters producing lantipeptides, siderophores, and non-ribosomal peptides across the Tree of Life.
Our collaborative efforts have also led to the use of synthetic peptides, establishing a robust baseline and pioneering methodologies for predicting mechanisms of action. To validate the antimicrobial potential of these novel molecules, we conducted high-throughput assays against reference and clinically relevant ESKAPEE pathogens from Mexico.
In the spirit of open science and collaboration, we openly share a fully sequenced and characterized collection of recently isolated ESKAPE pathogens from Mexico. Researchers worldwide can access this valuable resource for antimicrobial and antimicrobial resistance (AMR) research.
Our ongoing laboratory work aims to confirm these predictions and publish the structure-function correlations obtained. By doing so, we lay a solid foundation for identifying novel BGCs and antimicrobial molecules. Ultimately, our research contributes to a sustainable strategy for addressing the pressing global health challenge of antimicrobial resistance.

750 - HGFiT: Assessing the impact of HGT on metabolic fitness through the simulation of random transfer events in Escherichia coli’s

Tania Alonso-Vásquez

Tania Alonso-VÃ¡squez, Christopher Riccardi, Marco Fondi
Department of Biology, University of Florence, Department of Biology, University of Florence, Department of Biology, University of Florence

Horizontal gene transfer (HGT) is the acquisition of genetic material outside of a parent-offspring relationship. While it has been detected

across all domains of life, its relevance in bacteria is of critical concern if we think of the role it plays in, for example, antimicrobial resistance. From a basic perspective, a horizontally acquired gene may have a positive or negative impact on the bacteriumâ€™s fitness, depending on its position in the recipientâ€™s molecular circuitry, and on the energetic cost on a cell via the demands at the DNA/RNA/protein levels that it imposes. Despite being one of the drivers of evolution, to this day, a large-scale computational model of the effect of HGT on bacterial metabolic fitness is lacking.

How does an HGT event affect metabolic fitness in the light of this demand/benefit trade-off? Here we developed a computational pipeline to simulate the random transfer of DNA fragments among genomes and estimate their (positive or negative) impact on microbial metabolic phenotypes using constraint-based modelling. We used a set of randomly depleted genome-scale metabolic models (GSMM) to increase its evolvability space. We then simulated random transfers of DNA fragments from 21 species of Bacteria (donors) to this set of GSMMs of E. coli (recipients) and evaluated the impact of such new acquisitions on the recipientsâ€™ metabolic fitness. By doing so, we identified trends (involving donor-recipient phylogenetic distance, fragment length, etc.) that illuminate on the role of HGT in microbial evolution and on the factors that influence it.

801 - Five Centuries of Genome Evolution and Multi-Host Adaptation of Campylobacter Jejuni in Brazil

Stephanie Souza

Stephanie S.R. Souza, Ana Beatriz G. Buiatte, Leticia R.M. Costa, Phelipe A.B.M. Peres, Roberta T. Melo, Simone Sommerfeld, Belchiolina B. Fonseca, Nicole I. Zac Soligno, Odion O. Ikhimiukor, Paulo M. Armendaris, Cheryl P. Andam, Daise A. Rossi

University at Albany, Federal University of Uberlandia, University at Albany, Federal University of Uberlandia, Federal University of Uberlandia, Federal University of Uberlandia, Federal University of Uberlandia, University at Albany, University at Albany, Federal Agriculture Defense Laboratory/RS, University at Albany, Federal University of Uberlandia

Consumption of raw, undercooked, or contaminated animal food products is a frequent cause of Campylobacter jejuni infection. Brazil is the worldâ€™s third largest producer and major exporter of chicken meat, yet population-level genomic investigations of C. jejuni in the country remain scarce. Analysis of 221 C. jejuni genomes from Brazil shows that the overall core and accessory genomic features of C. jejuni are influenced by the identity of the host source. Of the 60 sequence types detected, ST353 is the most prevalent and consisted of samples from chicken and human sources. ST353 harbors the highest number of antimicrobial resistance genes per genome and virulence genes associated with the autoimmune neurological disorder Guillain-BarrÃ© syndrome. We traced the most recent common ancestor of ST353 to the late 1500s, coinciding with the time the Portuguese first arrived in Brazil and introduced domesticated chickens into the country. Further population growth of ST353 was likely driven by the large-scale expansion and commercialization of poultry farming. We identified at least two chicken-to-human jumps in ST353. Our study highlights the need for continued surveillance of C. jejuni in clinical and agricultural settings to control its spread and limit its impact on food security and public health.

813 - Creating Online Phylogenetic Resources for the Mycobacterium Tuberculosis Complex

Lily M Karim

Lily M Karim, Francisco J Martinez-Martinez, Ash O'Farrell, Angie Hinrichs, IÃ±aki Comas, Russell Corbett-Detig

UC Santa Cruz Department of Biomolecular Engineering and Bioinformatics, UC Santa Cruz Genomics Institute, Tuberculosis Genomics Unit, Instituto de Biomedicina de Valencia (IBV), CSIC, UC Santa Cruz Genomics Institute, UC Santa Cruz Genomics Institute, Tuberculosis Genomics Unit, Instituto de Biomedicina de Valencia (IBV), CSIC, UC Santa Cruz Genomics Institute, UC Santa Cruz Department of Biomolecular Engineering and Bioinformatics

During the SARS-CoV-2 pandemic, online phylogenomics played a crucial role in genomic surveillance, contact tracing, and identifying new variants. This method involves continuously updating a phylogeny of all available samples and integrating new samples into the existing tree on demand, enabling rapid research and response despite the large volume of genomic data generated. Extending online phylogenomics to other pathogens holds the potential to significantly enhance research speed and capabilities for various illnesses. The Mycobacterium tuberculosis complex (MTBC) is a group of closely related bacteria responsible for 10 million tuberculosis infections worldwide. Current tuberculosis research typically involves subsampling from available genomes, which limits insights into the deep-time evolutionary histories and nearest neighbors of samples. Previous limitations in storage and computational capacity hindered the creation and maintenance of a global phylogenetic tree for the MTBC. With the emergence of pandemic-ready phylogenetic software tools like UShER, we are now hosting and maintaining an up-to-date global phylogeny of the MTBC. This comprehensive phylogeny facilitates tracing new outbreaks, identifying pathogen lineages and antibiotic resistance, and analyzing genomic divergence within the complex. Incorporating all available sequences and associated metadata enables a detailed examination of mutation and antibiotic resistance variation among lineages. Rather than relying on sub-sampled datasets with approximate neighboring sequences, the closest related samples for each new sequence can be used for transmission information and outbreak analysis within local communities. Thus, the up-to-date, comprehensive MTBC phylogeny enables the most effective response to new tuberculosis infections worldwide.

816 - Developing Robust Machine Learning Models for AMR Prediction in Pseudomonas aeruginosa

Lucia Grana Miraglia

Lucia Grana-Miraglia, Pauline W Wang, Nadia Morales-Lizcano, Allison McGeer, David S Guttman
University of Toronto, University of Toronto, Center for the Analysis of Genome Evolution and Function, University of Toronto, University of Toronto, University of Toronto, Center for the analysis of Genome Evolution and Function

The progression of antimicrobial resistance (AMR) continues to be a significant challenge for worldwide public health. Increasing AMR means fewer treatment options, which highlights the need for improved diagnostic methods to allow rapid detection of resistance patterns.
The availability of large-scale databases containing sequenced resistant microbial isolates represents a significant milestone in our study of AMR, underscoring the importance of advanced computational tools to thoroughly analyze the genetic basis of antimicrobial resistance.
Machine learning and associated statistical genetic approaches are widely applied in microbiology for predicting traits of interest, such as AMR, and unveiling the underlying genetic basis. Although both supervised and unsupervised machine learning methods have proven valuable in these area, there are significant inherent challenges in deploying machine learning within the field of genomics.
Among the most important, high dimensional data, sampling biases and the non-independent evolutionary backgrounds among samples (population structure). Pseudomonas aeruginosa (PA), responsible for causing both severe acute and chronic persistent infections, presents significant challenges. This opportunistic pathogen posses a high level of both intrinsic as well as acquired resistance. We used a large and representative dataset of 2311 PA isolates, collected from seven different facilities in the Great Toronto Area, to predict AMR from genomic data. Using different learning algorithms, feature extraction and selection techniques, we were able to handle high dimensionality and reduce the effect of population structure to obtain high performing and generalizable models. Obtaining accurate predictions is essential for forecasting treatment outcomes, designing effective interventions, and allocating societal resources appropriately.

991 - Host adaptation of Campylobacter jejuni and the sources of human infection in California

Evangelos Antonios Dimopoulos

Evangelos Antonios Dimopoulos, Caitlin Collins, Maliha Aziz, Daniel Park, Cindy Liu, Lance Price, Julian Parkhill
Department of Veterinary Medicine, University of Cambridge, Department of Veterinary Medicine, University of Cambridge, Antibiotic Resistance Action Center, Department of Environmental and Occupational Health, Milken Institute School of Public Health, George Washington University, Antibiotic Resistance Action Center, Department of Environmental and Occupational Health, Milken Institute School of Public Health, George Washington University, Antibiotic Resistance Action Center, Department of Environmental and Occupational Health, Milken Institute School of Public Health, George Washington University, Antibiotic Resistance Action Center, Department of Environmental and Occupational Health, Milken Institute School of Public Health, George Washington University, Department of Veterinary Medicine, University of Cambridge

Campylobacter jejuni is a commensal bacterium that colonises livestock animals, and is responsible for more than 80% of campylobacteriosis cases in humans. Its high genome plasticity and recombination rate allow it to adapt to multiple hosts and leads to the rapid emergence of lineages, some of which can be antibiotic resistant. To curb an increase in antimicrobial infections in humans, the State of California enacted Senate Bill 27, which prohibits the preventative use of antibiotics in livestock. Our aim in this study was to examine the population structure of Campylobacter in California, assess its transmission between animal hosts and humans, and to consider possible effects of the policy implemented by SB27.
We sequenced 69 human clinical isolates of C. jejuni from California and placed them in a concentric context of human and animal isolates. By comparing groups of isolates, we detected significant levels of differentiation between the human isolates post-SB27 and animal isolates from California. Through phylogenetic reconstruction we demonstrated that the human C. jejuni clinical isolates are largely derived from food animal sources. By identifying the genomic regions that were contributing to population differentiation among the different host groups, we were able to identify protein variants potentially responsible for host adaptation and propensity to cause infection in humans. By describing the population structure of Campylobacter in animal food sources, and the relationship of these populations to human-infecting isolates, this study attempts to identify what genes drive host adaptation and examine how policy-making shapes bacterial populations.

1075 - Host ACE2 Sequence Evolution Identifies SARS-Related Sarbecovirus Reservoirs

Eric X Gan

Eric X Gan, John A Frank, Viraj Jansari, Akiko Iwasaki
Yale University, Yale University, Stanford University, Yale University

SARS-CoV-2 (SCV2) is a zoonotic virus and the etiologic agent of the ongoing COVID-19 pandemic, which has severely impacted public health on a global scale. The documented spillover of SCV2 from humans into animals including felines, mink, and white-tailed deer highlights the critical need for improved surveillance and identification of virus reservoir species. Viruses that have persisted within host lineages for significant periods of evolutionary time are likely to have coevolved with host receptors, leaving behind a signature of selection-dependent sequence variation. We hypothesized that evolutionary sequence analysis of host virus receptors could be leveraged to identify reservoir animal lineages. Using SCV2 as a model, we performed a dN/dS analysis on 572 ACE2 sequences across all vertebrate orders to search for positive selection at residues known to be targeted by SCV2. Consistent with the literature, our ACE2 analyses identified the order Chiroptera as exhibiting positive selection consistent with persistent circulation of SARS-related sarbecoviruses. Because Chiroptera are the second most species-rich mammalian order, we subsequently performed bat lineage-specific dN/dS analyses of ACE2. We determined that the horseshoe/roundleaf (Rhinolophus spp./Hipposideros spp.) bat lineage displays a distinct pattern of positive selection at clusters of residues with known sarbecovirus spike interactions. These data suggest that SARS-related sarbecoviruses have circulated in and applied selective pressure to this lineage for approximately 64 million years. The

application of selection-based evolutionary analyses to pinpoint viral reservoir species could be vital in informing future viral surveillance efforts.

1115 - Assembly of the gut microbiota during early life in the Brown Booby

Misael Daniel Mancilla Morales

Misael Daniel Mancilla Morales, Mauricio GuillÃ©n-Parra, Yendi Navarro-Noya, Bibiana Montoya, Ana Escalante, Roxana Torres
Departamento de EcologÃ-a Evolutiva, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, EstaciÃ³n CientÃ-fica La Malinche, Centro Tlaxcala de BiologÃ-a de la Conducta, Universidad AutÃ³noma de Tlaxcala, Departamento de EcologÃ-a Evolutiva, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, Laboratorio de Interacciones BiÃ³ticas, Centro de InvestigaciÃ³n en Ciencias BiolÃ³gicas, Universidad AutÃ³noma de Tlaxcala, EstaciÃ³n CientÃ-fica La Malinche, Centro Tlaxcala de BiologÃ-a de la Conducta, Universidad AutÃ³noma de Tlaxcala, Departamento de EcologÃ-a Evolutiva, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, Departamento de EcologÃ-a Evolutiva, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico

Establishing gut microbiota during early life plays a crucial role in the hostâ€™s physiology and development, with potential long-term effects. In wild birds, the establishment and potential routes of transmission to colonize gut microbiota during early life are poorly known. In the Brown Booby (Sula leucogaster), a long-lived seabird that nests in Islas Marietas, Mexico we (1) characterized the bacterial microbiota diversity, composition and functionality of eggshells, chicks throughout early development (7- 56 days of age) and adults, and (2) evaluated the contribution of parental transmission of bacterial taxa to the chick microbiota. We analyzed exudates of cloaca and eggshells, and samples of feces to recover amplicon sequence variants (ASVs) of the 16S rRNA gene V3-V4 region. Gut microbiota of young and adults consisted mostly of the phyla Bacillota, Pseudomonadota, Actinobacteriota, Deferribacterota, Fusobacteriota, and Bacteroidota. For alfa diversity, the observed ASVs showed a slight decline over time during the period of chick development and were higher in eggshells than in any other stage. Frequent and dominant ASVs did not vary significantly through chick development. However, diversity values from cloacal exudates were significantly higher than those from feces samples. For beta diversity, the microbial communities differed throughout developmental stages and beta diversity values were higher in samples collected from feces than from cloacal exudates. Analyses of variation through the chick development of bacterial communitiesâ€™ functional traits and the importance of parental transmission to the chick gut microbiota are in progress and will be shown at the conference.

50 - Dissecting an ancient stress resistance trait syndrome in the compost yeast Kluyveromyces marxianus

Kaylee E Christensen

Kaylee E Christensen, Abel Duarte, Zhenzhen Ma, Judith L Edwards, Rachel B Brem
University of California, Berkeley, UC Berkeley, UC Berkeley, Stanford University, UC Berkeley, UC Berkeley

In the search to understand how evolution builds new traits, ancient events are often the hardest to dissect. Species-unique traits pose a particular challenge for geneticistsâ€”cases in which a character arose long ago and, in the modern day, is conserved within a species, distinguishing it from reproductively isolated relatives. In this work, we have developed the budding yeast genus Kluyveromyces as a model for mechanistic dissection of trait variation across species boundaries. Phenotypic profiling revealed robust heat and chemical-stress tolerance phenotypes that distinguished the compost yeast K. marxianus from the rest of the clade. We used culture-based, transcriptomic, and genetic approaches to characterize the metabolic requirements of the K. marxianus trait syndrome. We then generated a population-genomic resource for K. marxianus and harnessed it in molecular-evolution analyses, which found hundreds of housekeeping genes with evidence for adaptive protein variation unique to this species. Our data support a model in which, in the

distant past, K. marxianus underwent a vastly complex remodeling of its proteome to achieve stress resistance. Such a polygenic architecture, involving nucleotide-level allelic variation on a massive scale, is consistent with theoretical models of the mechanisms of long-term adaptation, and suggests principles of broad relevance for interspecies trait genetics.

51 - Molecular phylogenetic tree of a group of species with distant genetic distance using Orthopteran insects

Kazuhiro Satomura

Kazuhiro Satomura, Atsushi Ogura

Nagahama Institute of Bio-Science and Technology, Nagahama Institute of Bio-Science and Technology

Molecular phylogenetic trees are important tools for inferring phylogenetic relationships within and among species, taking into account their evolutionary history. Methods for estimating molecular phylogenetic trees have been established over the past 50 years of research. When inferring phylogenetic relationships, problems such as heterotachy and long branch attraction are known to arise due to homologous sequences being under different natural selection within a single data set or due to changing mutation bias during evolution. In addition, when estimating the age of species divergence from molecular phylogenetic trees, there is the problem of different clock rates for recent and older divergences, even when calibrated using the archaeological record. Orthoptera insects, the subject of this study, are more than 200 million years old. When constructing a molecular phylogenetic tree of this group of species, even using whole genome sequences, homologous genes common to all species cannot be found. Even when a common ortholog is obtained, it is not always derived from a common ancestor over the entire length of the sequence. Even if homologous sequences derived from a common ancestor can be defined, invariant sites and multiple substituted sites cannot be used to infer the tree. Even if we could define a one-time substitution site, we could not determine whether the mutation is fixed in the population without population genomic data. We investigated the methods and problems of molecular phylogenetic inference for a group of species with a long evolutionary history using DNA sequences of Orthoptera species.

53 - Archaeogenomics of maize evolution in the South American Andes

Shuya ZHANG

Shuya ZHANG, Robin Allaby

University of Warwick, Smithsonian Institution, University of Science and Technology of China, University of Warwick

Maize used as a staple crop evolved from a wild plant called teosinte due to human activities in Mexico approximately 9000 years ago. It then was brought through Central America by humans around 7500 years ago, and later, around 6500 years ago, it was spread into South America. While it is widely known when maize spread to various regions, there is still debate regarding whether maize had undergone complete domestication prior to its arrival in South America. Screening for the best-preserved representative samples across the collection of Andean maize specimens established 9 samples to take forward for genome analysis. Structure analysis placed the ancient genomes of the Andes, all pre-Inca, to a single group to the exclusion of modern Andean landraces. Genetic distances between ancient genomes established links between maize samples from distinct cultural groups and the occurrence of network hubs. Analysis of selection signatures showed two possible episodes of selection involved in the adaptation to altitude, the first in pre-Inca times and the second more recently. The relationships revealed between ancient genomes provide evidence for cultural interaction through maize in the Andes spreading from a hub in the Southeast, and later spreading across the continent to Brazil. The evidence also suggests the hypothesis of a second wave of maize into the Andes, possibly as late as the Inca period, which may have brought secondary adaptations to altitude.

58 - Evolution of Hox genes in Tetrapoda: an analysis of evolutionary rates, selection regimes and their correlation with phenotypic traits

Anieli Guirro Pereira

Anieli Guirro Pereira, Tiana Kohlsdorf

Universidade de SÃ£o Paulo, Universidade de SÃ£o Paulo, Universidade de SÃ£o Paulo

Homeotic genes that encode HOX transcription factors play a fundamental role in processes of embryo development. Hox genes can be found in all Bilateria; in tetrapods, these genes are organized into four chromosomal groups: HoxA, HoxB, HoxC and HoxD, with 11, 10, 9 and 9 genes, respectively, totaling 39 genes in most species of this lineage. However, some of these genes were apparently lost in specific groups of tetrapods: crocodilians have lost HoxB3, HoxB8 and HoxD1, amphibians do not exhibit HoxB13 and HoxD12, snakes lack HoxD12, and some groups of mammals do not present HoxB3 and HoxB8. Although the evolution of Hox genes in Bilateria is well documented, most studies have primarily used mouse and human genomes as representatives of vertebrates, so that patterns within Tetrapoda underexplored. The increasing number of new genomes available for non-mammalian lineages, especially from more ancient lineages such as Lissamphibia, offers a unique opportunity for a more comprehensive investigation. Therefore, the purpose of this study is to examine the origin and evolution of Hox genes in Tetrapoda and evaluate evolutionary rates and selection regimes of these genes. Furthermore, we intend to explore phenotypic characteristics that may be associated with changes in these rates within Tetrapoda.

79 - When branching and evolution are tightly coupled

Jordan Douglas

Jordan Douglas

Department of Physics, University of Auckland

Standard temporal phylogenetic methods make the assumption that lineages evolve independently, and their evolution is unaffected by speciation events. However, this assumption is often unmet. There are many known cases when branching drives the rate of evolution, or when evolution drives the rate of branching. At the molecular level, gene duplication events enable the emergence of novel function(s) (neofunctionalisation or subfuncationalisation). At the organism level, speciation is often accompanied by rapid evolutionary changes (punctuated equilibrium). And, at the cultural level, languages often undergo rapid changes when one population breaks off from another (schismogenesis), perhaps as a means to establish cultural identity. However, these granularities of biological and cultural evolution are generally studied in distinct fields of academia. This presentation will describe a unified Bayesian phylogenetic method that accounts for these processes. This model combines phylogenetic (strict and relaxed) clock models with the birth-death process. Unobserved extinction events, or â€œstubsâ€•, are accounted for.

80 - CRISPR is mutagenic, even in the absence of gRNA.

Stephan Baehr

Stephan Baehr, Samuel Perez, Adalyn Brown, Alyssa Cenzano, Katelyn Hancock, Lea Patrick, Michael Lynch

Arizona State University, Biodesign Institute, Arizona State University, The Ohio State University, Arizona State University, Arizona State University, Arizona State University, Arizona State University

CRISPR is a gene editing platform with seemingly endless applications, from engineering toscience and medicine. The excitement of CRISPRâ€™s potential has lead to hundreds of clinical trials aimed at treating human genetic diseases. However, the field is aware that CRISPRâ€™s promise is not entirely without risk; CRISPR by its nature modifies the genome, and may have off-target

effects. No enzyme or machine is perfect. Researchers have diligently sought to quantify mutation risk, by testing a specific locus, or set of loci. However, to date, we are unaware of any experiment using the most sensitive detection methods of evolutionary biology, which can detect error rates
that may occur one in a trillion bases per cell division. We demonstrate a mutational hazard of expressing Cas9 to an untargeted locus of yeast, S. cerevisiae, which demonstrates a dose-response relationship with increasing Cas9 expression. We find a modest 1.2-fold to 2.7-fold increase in
mutation rate, as quantified by fluctuation test. We demonstrate that gRNA appears to increase the mutagenicity of CRISPR, be it Cas9, Cas14, or an Adenine Base Editor (ABE). We note that
if indels are the primary means of mutation mechanism in Cas9, the indel rate may be 10-fold higher in our high-expression yeast lines relative to WT controls.

109 - The evolutionary history of the cycloamanide biosynthesis genes within Agaricomycetes species

Christian Quintero

CHRISTIAN QUINTERO-CORRALES, MELANIA VEGA, GRISELDA NALLELY HERNÃ•NDEZ-RICO, PABLO OCTAVIO-AGUILAR, ROBERTO GARIBAY-ORIJEL
Instituto de BiologÃ-a, UNAM, JardÃ-n BotÃ¡nico, Instituto de BiologÃ-a, UNAM, Centro de Investigaciones BiolÃ³gicas, Universidad AutÃ³noma del Estado de Hidalgo, Centro de Investigaciones BiolÃ³gicas, Universidad AutÃ³noma del Estado de Hidalgo, Instituto de BiologÃ-a, UNAM

Horizontal gene transfer (HGT) is one of Eukaryotic life's most interesting and complex evolutionary events. There is evidence that microorganisms share genes with other organisms, conferring potential evolutionary advantages. The cycloamanide, like amatoxins, are one of the most relevant fungal metabolites responsible for most worldwide deaths by mushroom poisoning. It has been suggested that involved genes passed to Galerina, Lepiota and Amanita sect. Phalloideae by HGT. Since several poisoningsâ€™ cases by the consumption of A. rubescens complex species in Mexico have been reported, we used de novo genome assembling and gene mining for cycloamanide homologous genes. The first record for Amanita bruenneolocularis and Amanita flavorubes in Mexico, together with the â€œA. cruentilemurumâ€• samples has structurally similar genes to the known cycloamanide synthesis pathway. Among all studied samples, species from one particular site showed most of the cycloamanide synthesis pathway genes. Also, high cryptic diversity was found for the prolyl-oligopeptidase gene family showing a more complex evolutionary history than previously reported. Our results suggested that several old horizontal transfer and duplication events among Agaricomycetes fungal lineages, such as Piloderma, Russula, Cortinarius, Hebeloma, and more Amanita happened.

126 - Gene and allele specific expression underlying the electric signal divergence in African weakly electric fish

Feng Cheng

Feng Cheng, Alice B Dennis, Otto Baumann, Frank Kirschbaum, Salim Abdelilah-Seyfried, Ralph Tiedemann
University of Potsdam, University of Namur, University of Potsdam, Humboldt University, University of Potsdam, University of Potsdam

In the African weakly electric fish genus Campylomormyrus, electric organ discharge (EOD) signals are strikingly different in shape and duration among closely related species, contribute to pre-zygotic isolation and may have triggered an adaptive radiation. We performed mRNA sequencing on electric organs (EOs) and skeletal muscles (SMs; from which the EOs derive) from three species with short (0.4

ms), medium (5 ms), and long (40 ms) EODs and two different cross-species hybrids. We identified 1,444 up-regulated genes in EO shared by all five species/hybrids cohorts, rendering them candidate genes for EO-specific properties in Campylomormyrus. We further identified several candidate genes, including KCNJ2 and KLF5, their up-regulation may contribute to increased EOD duration. Hybrids between a short (C. compressirostris) and a long (C. rhynchophorus) discharging species exhibit EODs of intermediate duration and showed imbalanced expression of KCNJ2 alleles, pointing towards a cis-regulatory difference at this locus, relative to EOD duration. KLF5 is a transcription factor potentially balancing potassium channel gene expression, a crucial process for the formation of an EOD. Unraveling the genetic basis of the species-specific modulation of the EOD in Campylomormyrus is crucial for understanding the adaptive radiation of this emerging model taxon of ecological (perhaps even sympatric) speciation.

128 - Evolution of discrete phenotypic plasticity in a gene regulatory network model

Takahiro Sakamoto

Takahiro Sakamoto, Hideki Innan

SOKENDAI (The Graduate University for Advanced Studies), SOKENDAI (The Graduate University for Advanced Studies)

Phenotypic plasticity is a common strategy for a species to adapt to temporally or spatially varying environments. Some cases of plasticity exhibit a discrete phenotype where the phenotype changes discontinuously along an environment gradient with no intermediate phenotypes arising in intermediate environments (discrete plasticity). It has been argued that discrete plasticity requires a developmental switch to control different phenotypes. However, since most previous theories of phenotypic plasticity are based on the quantitative genetic model and do not consider the developmental process, the evolutionary origin of the developmental switch underlying discrete plasticity remains unclear. In this study, we integrated a gene regulatory network into the population genetics framework and investigated the evolutionary dynamics of phenotypic plasticity in a temporally fluctuating environment. In our model, an individualâ€™s fitness is determined by the expression levels of n genes, which change throughout development due to gene regulation. Optimal expression levels differ between the environments so phenotypic plasticity is favored. An environmental sensor gene is also included to enable environment-dependent gene regulation. Using simulations, we investigated whether a population evolves adaptive plasticity and under what conditions the evolved plasticity tends to be discrete. We found that plasticity evolves in most cases, with the proportion of discrete plasticity varying across different parameter sets. Notably, the proportion of discrete plasticity is high when the fluctuation period of the environment is long. Our study highlights the importance of considering development in theoretical studies of phenotypic plasticity.

134 - Single-nucleus multi-omics analyses reveal cellular and molecular innovations in the anterior cingulate cortex during human evolution

Bing Su

Jiamiao Yuan, Kangning Dong, Haixu Wu, Xuerui Zeng, Xingyan Liu, Yan Liu, Jiapei Dai, Jichao Yin, Yongjie Chen, Yongbo Guo, Wenhao Luo, Na Liu, Yan Sun, Shihua Zhang, Bing Su

Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming Institute of Zoology, Kunming Institute of Zoology, Kunming Institute of Zoology, Kunming Institute of Zoology, School of Mathmatics, Renmin University of China, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming Institute of Zoology, Chinese Academy of Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Kunming Institute of Zoology, Chinese Academy of Sciences, Wuhan Institute for Neuroscience and Neuroengineering, South-Central Minzu University, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming Institute of Zoology, Chinese Academy of Sciences, Wuhan

Institute for Neuroscience and Neuroengineering, South-Central Minzu University, Wuhan Institute for Neuroscience and Neuroengineering, South-Central Minzu University, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Kunming Institute of Zoology, Chinese Academy of Sciences

The anterior cingulate cortex (ACC) of the brain involves higher-level cognitive functions such as emotion and self-awareness, and is closely related to human mental disorders. The ACC of Old World anthropoids contains von Economo neurons (VENs), one of the noticeable cellular innovations during primate brain evolution. Here, we present the profiles of human and macaque ACC gene expression and chromatin accessibility at single-nucleus resolution. We characterized the conserved gene expression, chromatin accessibility, and transcription factor binding patterns in different cell types and neuronal subtypes. Importantly, by combining the published mouse data, we discovered the molecular identities and the possible cell lineage origin of the primate VENs. In particular, with both in vitro and in vivo experimental validations, we reported a group of human-macaque-shared and human-specific VEN marker genes, such as PCSK6, ADAMTSL3 and CDHR3, potentially contributing to the morphogenesis and function of VENs. Furthermore, we established the connections between the genome-wide human-specific mutations and the specialized gene regulation in the human ACC, delineating the genetic basis of cellular and functional innovations in the ACC during primate evolution and human origin. These findings provide new insights into understanding the cellular composition and molecular regulation of ACC and its evolutionary role in shaping human-owned higher cognitive skills.

136 - Patterns of miRNA presence and absence in mammals have implications for placental phenotypes

Jonathan Fenn

Jonathan Fenn, Nkemdilim Nwaobi, Vladimir Ovchinnikov, Haidee Tinning, Zenab Butt, Olga Amelkina, Niamh Forde, Mary O'Connell
University of Nottingham, University of Nottingham, Queen Mary University, University of Leeds, University of Leeds, University of Leeds, University of Leeds, University of Nottingham

microRNAs (miRNAs) are small regulatory RNAs which inhibit translation of mRNA - but how does the presence or absence of different miRNAs in the genome affect phenotype? The evolution of placental implantation in Eutheria was accompanied by massive diversification in miRNAs, and there are miRNA families that are entirely unique to â€“ and conserved across- mammals. We revise previous estimates of Eutheria-synapomorphic miRNA families, using a dataset of 300 mammal genomes, homology searches and machine learning-based miRNA identification methods. We identify 10 core toolkit miRNA families whose functions are likely essential for successful pregnancy in Eutheria, and 314 miRNA families unique to mammals but with phylogenetically heterogeneous distributions. We hypothesised that the diversity of placental phenotypes could be underpinned by miRNA family presence/absence across the mammal phylogeny. We used a combination of phylogeny-aware phenotype diversity metrics (Evolink) and machine-learning approaches (Random Forest) using our purpose built placental phenotype dataset and our miRNA family presence/absence matrix. We identify 34 miRNA families strongly associated with placental phenotype. Many of these miRNAs are associated with aspects of mammalian reproduction and development. Amongst the strongest associations includes Mir-11968, present in all species examined within Ruminantia (distinctive in its cotyledonary placental shape), but completely absent outside this clade. We present our findings for the functional analysis of Mir-11968 here. In summary, our findings provide insight into the evolutionary relevance of miRNAs associated with placental phenotype, and provide targets for future work that are relevant to developmental pathologies.

140 - Creating and using open educational resources to teach biology to improve accessibility, inclusion, and social justice

Nadia Aubin-Horth

Nadia Aubin-Horth, Ariane Barrette, Alicia Pageau, Jeanne Picher-Labrie, Roxanne Turgeon
Universite Laval, Universite Laval, Universite Laval, Universite Laval, Universite Laval

Open educational resources (OER) are teaching and learning materials belonging to the public domain or published with an intellectual property license allowing their use, adaptation, and distribution free of charge. We will present how creating and using OERs supports the process of democratization of knowledge and can improve accessibility, inclusion, and social justice. We created an OER, based on the principles of Universal Design for Learning, dedicated to teaching evolution and physiology at the university undergraduate level. This resource is in the form of a digital text interspersed with videos, images, and interactive activities. This resource has a creative commons license that allows anyone to use, copy, redistribute, transform, remix and improve it for free (CC-BY-NC-SA). It was co-created in a team including a university professor and 4 graduate students. We will present approaches used to achieve our objective of accessibility and inclusion: various modes of presentation, applied cases using published data, subtitling of videos and transcription, alternative text for all images, compatibility with screen reading software, and PDF version to work offline or in print for students with reduced access to the internet. We will present our aim to diversify the representation of scientists as they are usually portrayed in textbooks, using the creation of video interviews with colleagues from underrepresented groups in science and at different stages of their careers, in which they present their research and their scientific journey in the field. Finally, we will discuss the co-creation process and its benefits.

144 - Development of a bioinformatics Pipeline for the Analysis of Complete Genomic Sequencing Data of Mycobacterium tuberculosis

Ikuri Alvarez-Maya

Ikuri Alvarez-Maya, Manuel GarcÃ-a-Ulloa, Jaime Martinez-Urtaza
Medical and Pharmaceutical Biotechnology, Center for Research and Assistance in Technology and Design of the State of Jalisco (CIATEJ) , Department of Genetics and Microbiology, Faculty of Biosciences, Autonomous University of Barcelona, Department of Genetics and Microbiology, Faculty of Biosciences, Autonomous University of Barcelona

The rapidly growing area of sequencing technologies, and more specifically bacterial whole genome sequencing, could offer new applications in clinical microbiology, such as the identification of bacterial species, prediction of genetic susceptibility to antibiotics, and virulence simultaneously. We developed a bioinformatics tool for the analysis of data obtained by Complete Genomic Sequencing that allows us to know the phylogenomic behavior.
Mycobacterium tuberculosis (MTB) raw Illumina short reads were collected from projects PRJEB30933, PRJNA824124, and PRJEB44165 on the website www. ebi.ac.uk/ena repository. Additionally, the scaffolds from project PRJNA751891 at ww.ncbi.nlm.nih.gov. were also collected. Raw reads were trimmed and quality filtered with fastp v0.12.4 using default parameters and a minimum read length of 50 bp. Genome assemblies were performed with SPAdes v3.15.5 using the --careful --cov-cutoff automatic options.
Validation of the assemblies was carried out with QUAST v5.0.2 and BUSCO v5.4.2 using the actinobacteria_class_odb10 single-copy ortholog database as a reference. Sequencing depth was calculated by mapping the filtered reads to their respective assemblies using Bowtie2 v2.2.9 and samtools v1.7.
Assemblies were scanned for contaminants by blasting their scaffolds against the NCBI nr database. Variable sites from all assemblies were extracted and aligned using the M. tuberculosis H37Rv strain (GCF_000195955.2) Multidrug resistance profiles, lineages, and sublineages were assigned to strains in silico with tb-profiler v4.3.0.
The bioinformatic pipeline developed in the study was able to adequately analyze and identify not only the phylogenetic behavior of M. tuberculosis strains circulating in Mexico but also mutations related to resistance to drug treatment.

150 - Transcriptome reference-based SNP calling as an alternative for SNP annotation in the absence of a reference genome for invasion genomics studies.

Gabriela Castellanos-Morales

JosÃ© Francisco Miranda-Vidal, Yocelyn T. GutiÃ©rrez-Guerrero, AnahÃ- Canedo-TÃ©xon, Maricela GarcÃ-a-Bautista, Everardo Barba, Gabriela Castellanos-Morales

El Colegio de la Frontera Sur, Unidad Villahermosa, UC Berkeley, El Colegio de la Frontera Sur, Unidad Villahermosa, El Colegio de la Frontera Sur, Unidad San CristÃ³bal de las Casas, El Colegio de la Frontera Sur, Unidad Villahermosa, El Colegio de la Frontera Sur, Unidad Villahermosa

For many invasive species, as in other non-model species, reference genomes for SNP annotation are lacking. This limits our ability to identify candidate loci to advance our understanding about how these species adapt to new environmental conditions. In such cases, available transcriptomic data can be used to annotate candidate SNPs. We used this approach to study the invasive genomics of the suckermouth armored catfish (Pterygoplychthys sp.) in the Grijalva-Usumacinta River Basins. These South American fish have been introduced and become invasive in many areas of Asia and North America. In invaded areas, they threaten biodiversity and cause economic loss for freshwater fisheries. Previous studies based on morphological and genetic data suggested that hybridization could be playing a role in invasion success for these taxa. In the present study, we conducted genotyping-by-sequencing for 103 samples from 5 localities representing an environmental gradient in the Grijalva and Usumacinta River Basins in Mexico. We assembled and annotated the transcriptome for Pterygoplichthys pardalis from available data at the NCBI database. We conducted de novo and reference SNP calling, using P. pardalis reference transcriptome to map reads. We obtained 22,595 and 3,083 SNPs, respectively. Patterns of genetic variation and genetic structure were similar between datasets, and our results support the hypothesis of a hybrid origin for this invasive population, and rapid population expansion. Moreover, the transcriptome referenced database allowed identifying candidate loci associated to stress response. Transcriptome reference-based SNP calling is an alternative in the absence of a reference genome.

163 - Mitochondrial whistleblower: Unveiling a Hidden Third Partner in Sponge Symbiosis

Momin Ahmed

Momin Ahmed, Jan Vicente, Dennis V Lavrov

Iowa State University, Hawaiâ€™i Institute of Marine Biology, University of Hawaiâ€™i at Manoa, Iowa State University

Sponges (Phylum Porifera) form symbiotic relationships â€“ ranging from mutualism to parasitism â€“ with a variety of other organisms, including other sponges. Recently, Vicente et al. described a symbiotic association between two sponge species, a demosponge Haliclona plakophila and a homoscleromorph Plakortis symbiotia. Surprisingly, our genome skimming of the H. plakophila â€“ P. symbiotica samples revealed the presence of, not two, but three complete sponge mitochondrial genomes (mitogenomes), all with high sequence coverage. Here we report phylogenetic analysis of mitochondrial as well as 18S and 28S sequences assembled from H. plakophila â€“ P. symbiotica association aiming to resolve the phylogenetic position of the third â€œMysteryâ€• sponge. Our findings indicate that mitochondrial sequences from this species were phylogenetically uninformative, likely due to their highly unusual GC-rich nucleotide composition. By contrast, 18S and 28S genes have placed the Mystery sponge within Haplosclerida, but not in any of the recognized clades within this group. We suggest that the Mystery sponge's atypical nucleotide composition in its mitogenome reflects its highly unusual, potentially parasitic, lifestyle. This work opens new avenues for exploring the biology of this enigmatic sponge.

178 - Disentangling the genomic basis of terrestrialization across animals through a transcriptomics spyglass

Klara Eleftheriadi

Klara Eleftheriadi, Judit Salces-Ortiz, Nuria Escudero, Carlos Vargas-Chavez, Fernando-Ã•ngel FernÃ¡ndez-Ã•lvarez, Rosa FernÃ¡ndez
Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Institut de CiÃ¨ncies del Mar (ICM), Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra)

Animal terrestrialization, the colonization of the land from aquatic species, is one of the most remarkable events in the history of life on Earth. From a relatively stable marine environment, animals adapted the way they breathe, excrete, reproduce etc. to such an extreme change of environment. From all animal phyla, 9 undoubtedly, comprise terrestrial forms. The genomic basis of this exceptional phenomenon remains largely unexplored. Here, aiming to unveil the genes that facilitated the adaptation to such an extreme environment, we subjected terrestrial animals and their aquatic relatives in each phylum to 12 experiments mimicking some of the abiotic stressors related to habitat shift and land colonization (visible light and UV exposure, responses to chemosensory cues and variations in oxygen concentration, salinity, and humidity). These involved 23 species from the 9 animal phyla, yielding around 1,000 transcriptomes sequenced using a combination of Illumina and PacBio Iso-Seq. Through comparative transcriptomics and phylogenetic comparative methods, we discuss the genomic toolkit involved in terrestrialization and pinpoint genes in both lineage and phylum specific. Our results indicate that some animal phyla seem to behave similarly under stress conditions (e.g., with marine species having a higher response in terms of differentially expressed genes compared to terrestrial ones), whilst others seem to have different mechanisms to cope with stressors related to this adaptation. The results of this project will shed light on how animals reshaped their genomes to cope with the dramatic physiological and morphological changes associated with life on land.

190 - The probability of a unique gene occurrence at the tips of a phylogenetic tree in the absence of horizontal gene transfer (the last-one-out)

Nico Bremer

Nico Bremer, William F. Martin, Mike Steel
Institute for Molecular Evolution, Heinrich-Heine-University DÃ¼sseldorf, Institute for Molecular Evolution, Heinrich-Heine-University DÃ¼sseldorf, Biomathematics Research Centre, School of Mathematics and StatisticsUniversity of Canterbury

Gene loss is an important process in gene and genome evolution. If a gene is present at the root of a rooted binary phylogenetic tree and can be lost in one descendant lineage, it can be lost in other descendant lineages as well, and potentially can be lost in all of them, leading to extinction of the gene on the tree. In that case, just before the gene goes extinct in the rooted phylogeny, there will be one lineage that still retains the gene for some period of time, representing a â€˜last-one-outâ€™ distribution. If there are many (hundreds) of leaves in one clade of a phylogenetic tree, yet only one leaf possesses the gene, it will look like the result of a recent gene acquisition, even though the distribution at the tips was generated by loss. Here we derive the probability of observing last-one-out distributions under a Markovian loss model and a given gene loss rate Î¼. We find that the probability of observing such cases can be calculated mathematically, and can be surprisingly high, depending upon the tree and the rate of gene loss. Examples from real data show that gene loss can readily account for the observed frequency of last-one-out gene distribution patterns that might otherwise be attributed to lateral gene transfer.

209 - Historical biogeography and climate niche differentiation in extant species of Ceratophryidae (Anura: Hyloidea) frogs in South America.

H. Mauricio Ortega Andrade

H. Mauricio Ortega Andrade, Arjen A.W. van der Kamp

Universidad Regional AmazÃ³nica Ikiam, Biogeography and Spatial Ecology Research Group, Instituto Nacional de Biodiversidad, 3Wildlife Ecology & Conservation Group, Wageningen University

Ceratophryidae is a family of frogs containing twelve extant species distributed in South America. Several studies have been done about the systematic, morphology, karyotype, and behaviour of this monophyletic family. However, little is known about the historical biogeography and the divergence in time of this group. Here, we present a calibrated chronogram and the historical biogeography area reconstruction to explore the radiation and environmental niche differentiation of eleven extant species of Ceratophryidae in South America. A phylogeny based on morphological and genetic data was obtained from 256 morphological characters, six mitochondrial and eight nuclear genes for up to a total of 8428 characters in the homological matrix. Our results indicate that the genus Ceratophrys is sister to the clade Chacophrys + Lepidobatrachus. The divergence of the crown group of Ceratophryidae is estimated to have occurred 19.2 Mya at the beginning of the Miocene, with recent cladogenetic events related to the late Miocene (10.18â€“13.70 Mya) and the Pliocene (~5.3-2.6 Mya). Moreover, the Cerrado region is estimated to be the ancestral area of the family, as well for the genera Chacophrys and Lepidobatrachus. In addition, temperature seasonality and annual precipitation play a major role in the niche differentiation of extant species within Ceratophryidae. In conclusion, our data suggests multiple dispersal and vicariant events which originated from the Cerrado region in the early Miocene, and recognize the role of environment in the differentiation of the ecological niche among extant species of this family in South America

228 - Incipient genetic differentiation of A. ludens (Diptera: Tephritidae) as a result of its recent geographic and host expansion leading to the attack of commercially grown apples in Mexico.

Helena Socorro Hernández-Rosales

Helena S. Hernandez-Rosales, Daniel Helena S. PiÃ±ero, Nancy Helena S. Nancy GÃ¡lvez-Reyes, Alma AltÃºzar-Molina, MartÃ-n Aluja

Universidad Nacional AutÃ³noma de MÃ©xico, Universidad Nacional AutÃ³noma de MÃ©xico, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de EcologÃ-a, ACâ€“INECOL, Cluì•ster Cientiì•fico y Tecnoloì•gico BioMimic, Instituto de Ecologiì•a, ACâ€“INECOL, Cluì•ster Cientiì•fico y Tecnoloì•gico BioMimic

The fruit fly Anastrepha ludens is a significant citrus pest in Mexico, primarily feeding on citrus pulp and other cultivated and wild species. Typically found in humid tropical regions below 500 meters altitude, its range has expanded to higher elevations due to warming, even reaching apple crops. Through population genetics, focusing on single nucleotide polymorphisms (SNPs), we discovered A. ludens exhibits greater genetic diversity (Ï€ = 0.176) than its sister species, A. obliqua (Ï€ ~ 0.0105). Flies infesting apples displayed extensive polymorphism (99%), contrasting with grapefruit-infesting flies, which were more conserved (50%). Principal Component Analysis identified genetically distinct apple parasites. Admixture analysis revealed genetic differentiation in apple-collected flies. Despite low differentiation (FST = 0.0125) among samples, a distinct group of apple-infesting flies closely related to pear parasites emerged. Evidence suggests recent diversification of apple-infesting fly lineages. Our study supports A. ludens possibly undergoing early population differentiation driven by host plants. Focusing on apple varieties, our findings indicate ongoing ecological divergence.

235 - Dinotoms possess two evolutionary distinct autophagy-related ubiquitin-like conjugation systems

Euki Yazaki

Euki Yazaki, Tadaaki Uehara, Hirokazu Sakamoto, Yuji Inagaki

Research Center for Advanced Analysis, National Agriculture and Food Research Organization, RIKEN iTHEMS, Graduate School of Life

and Environmental Sciences, University of Tsukuba, Department of Infection and Host Defense, Graduate School of Medicine, Chiba University, Department of Pathology, Stanford School of Medicine, Graduate School of Life and Environmental Sciences, University of Tsukuba, Center for Computational Sciences, University of Tsukuba

Autophagy is an intracellular degradation mechanism by which cytoplasmic materials are delivered to and degraded in the lysosome-fused autophagosome (autolysosome) and proposed to have been established at an early stage of eukaryotic evolution. Dinoflagellates harboring obligate endosymbiotic diatoms (so-called â€œdinotomsâ€•), which retain their own nuclei and mitochondria in addition to plastids, have been investigated as an intermediate toward the full integration of a eukaryotic alga into the host-controlled organelle (i.e., plastid). Pioneering studies systematically evaluated the degree of host governance on several metabolic pathways in the endosymbiotic diatoms (ESDs). However, little attention has been paid to the impact of the endosymbiotic lifestyle on the autophagy operated in the ESDs. In this study, we searched for ATG3, ATG4, ATG5, ATG7, ATG8, ATG10, and ATG12, which are required for autophagosome formation, in the RNA-seq data from dinotoms Durinskia baltica and Kryptoperidinium foliaceium. We detected two evolutionally distinct sets of the ATG proteins in the dinotom species, one affiliated with the dinoflagellate homologs and the other with the diatom homologs in phylogenetic analyses. The results suggest that the ATG proteins descended from the diatom taken up by the dinoflagellate host persist for autophagosome formation and, most likely, autophagy.

247 - Evolutionary and functional characterization of a new TRPV1 splicing variant originated in the ancestor of catarrhine primates

Juan C Opazo

Sofia Cacino, Deny Cabezas-Bratesco, Julia York, Kattina Zavala, Michael Vandewege, Sebastian Brauchi, Juan C. Opazo
Universidad Austral de Chile, Universidad Austral de Chile, University of Illinois, Universidad San Sebastian, North Caroline State University, Universidad Austral de Chile, Universidad San Sebastian

The evolutionary origin of new exons provides an opportunity for a more diverse repertoire of protein isoforms from a single gene. This process is a fundamental mechanism in evolution, increasing the functional complexity of genomes. The TRPV1 (Transient Receptor Potential Vanilloid 1) channel is a non-selective cation channel, a member of the TRPV gene family. It receives its name from being the receptor for capsaicin. It is also activated by temperature (>43Â°C), protons, and endogenous ligands. Physiologically, TRPV1 plays a critical role as a nociceptor, part of the mammalian pain pathways. Most studies have been conducted for a canonical splice variant of 839 amino acids; much less is known about the other variants. In this work, we studied the evolution of TRPV1 gene structure in primates. Our results show the existence of an exon (encoding for 11 amino acids) in the sixth position, which is only found in catarrhine primates. The splice variant incorporating the sixth exon is preferentially located in the endoplasmic reticulum; however, it is localized in the plasma membrane when co-expressed with the canonical variant. From an electrophysiological perspective, it presents lower currents than the canonical isoform, and it is not activated in the presence of capsaicin. Cells that co-expressed both variants showed currents similar to the canonical variant, although the activation in the presence of capsaicin was lower. Thus, our results suggest that the splice variant incorporating the sixth exon would be a negative dominant of the canonical isoform.

254 - Evolution of canonical circadian clock genes underlies unique sleep strategies of marine mammals for secondary aquatic adaptation

Daiqing Yin

Daiqing Yin, Zhaomin Zhong, Fan Zeng, Zhikang Xu, Wenhua Ren, Yang Guang, Han Wang, Shixia Xu
Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Jiangsu Key Laboratory for Biodiversity and

Biotechnology, College of Life Sciences, Nanjing Normal University, Center for Circadian Clocks, Soochow University, School of Biology & Basic Medical Sciences, Suzhou Medical College, Soochow University, School of Biology & Basic Medical Sciences, Suzhou Medical College, Soochow University, Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Center for Circadian Clocks, Soochow University, School of Biology & Basic Medical Sciences, Suzhou Medical College, Soochow University, Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University

To satisfy the needs of sleeping underwater, marine mammals, including cetaceans, sirenians, and pinnipeds, have evolved an unusual form of sleep, known as unihemispheric slow-wave sleep (USWS), in which one brain hemisphere is asleep while the other is awake. All aquatic cetaceans have only evolved USWS without rapid eye movement (REM) sleep, whereas aquatic sirenians and amphibious pinnipeds display both bihemispheric slow-wave sleep (BSWS) and USWS, as well as REM sleep. However, the molecular genetic changes underlying USWS remain unknown. The present study investigated the evolution of eight canonical circadian genes and found that positive selection occurred mainly within cetacean lineages. Furthermore, convergent evolution was observed in lineages with USWS at three circadian clock genes. Remarkably, in vitro assays showed that cetacean-specific mutations increased the nuclear localization of zebrafish clocka, and enhanced the transcriptional activation activity of Clocka and Bmal1a. In vivo, transcriptome analysis showed that the overexpression of the cetacean-specific mutant clocka (clocka-mut) caused the upregulation of the wakefulness-promoting glutamatergic genes. In contrast, the GABAergic and cholinergic pathways, which play important roles in promoting sleep, were downregulated in the bmal1a-mut-overexpressing zebrafish. Concordantly, sleep time of zebrafish overexpressing clocka-mut and bmal1a-mut were significantly less than the zebrafish overexpressing the wild-type genes, respectively. Together, these findings suggest that canonical circadian clock genes have evolved adaptively to enhance circadian regulation ability relating to sleep in cetaceans and, in turn, contribute to the formation of USWS.

256 - Comparative demographic analysis of Cardamine hirsuta and Arabidopsis thaliana

Rachita Srivastava

Rachita Srivastava, Bjorn Pieper, Sileshi Nemomissa, Donovan Bailey, Angela Hancock, Carlos-Alonso Blanco, Stefan Laurent, Miltos Tsiantis

Max Planck Institute for Plant Breeding Research, Max Planck Institute for Plant Breeding Research, Addis Ababa University, New Mexico State University, Angela, Centro Nacional de BiotecnologÃ-a, Max Planck Institute of Plant Breeding Research, BioNTech, Max Planck Institute of Plant Breeding Research

Understanding the dynamics of populations and their genetic divergence in relation to historical ecological and climatic events is a major goal of evolutionary biology. Comparing the population dynamics and distribution of closely related species can illuminate the roles of common versus divergent ecological parameters in shaping trait and sequence diversity patterns, identifying convergent genetic responses to global ecological changes. Cardamine hirsuta (Cardamine) is a small crucifer related to Arabidopsis thaliana (Arabidopsis), and shares features with Arabidopsis, making it an attractive experimental system. Population genetics analyses of Pan-European Cardamine samples comprising 746 natural accessions revealed that Cardamine from the Iberian Peninsula and the Macaronesian Islands retain considerable genetic diversity, thus resembling Iberian relict populations of Arabidopsis. However, Cardamine strains extend further into mainland Europe than Arabidopsis relicts, suggesting a greater ability to establish beyond glacial refugia. In

Arabidopsis, the deepest population splits have been reported to be those between the African populations. The absence of information on Cardamine African diversity leaves many open questions on the comparative understanding of the population dynamics of the two species. In this study, we developed a standardized bioinformatics workflow to generate variation data for two related Brassicaceae species sampled from Europe, North Africa, and East Africa. Next, we joint-analyzed the demographic history of Cardamine from Europe and Africa and compared it to that of Arabidopsis. Our findings shed light on the complex interplay of historical events, geographic influences, and ecological factors in shaping these closely related plant species' population history and genetic divergence.

262 - Comparing Ancient DNA Yields from Skeletal Elements Recovered in Subtropical Environments.

Margaret (Maggie) L Hassler

Margaret L Hassler, Jessica H Stone, Jaime Zolik, Laura N Pott, Carrie M Miller, Meredith A.B. Ellis, Maria Nieves-ColÃ³n
University of Minnesota, University of Minnesota, University of Minnesota, University of Minnesota, University of Minnesota, Florida Atlantic University, University of Minnesota

Sampling ancient DNA (aDNA) from skeletal elements beyond the dentition in the remains of human ancestors is a relatively new area of research. Historically, aDNA sampling has been focused on dentin. Scholars have recently begun examining other parts of the human skeleton that may be optimal for aDNA preservation, like the petrous portion and auditory ossicles of the cranium. However, existing literature on this subject mostly provides reports about endogenous DNA preservation in the cranium in samples from colder climates. This study addresses this gap by comparing endogenous DNA preservation in three different skeletal elementsâ€”the petrous portion, the auditory ossicle, and dentinâ€”from a historic individual whose remains were preserved in subtropical Okeechobee County, Florida. We sampled the skeletal elements in a clean lab, extracted DNA via spin columns, and prepared double-stranded libraries for shotgun sequencing on the Illumina HiSeq X platform. Our results indicate that the auditory ossicle yielded nearly 1.5 times more reads mapping to the human genome than the petrous portion and nearly 80 times more mapped reads than the dentin. However, the dentin yielded over 5 times greater mapped reads of mitochondrial DNA than the cranial elements. Ongoing experiments will investigate how libraries prepared from these three skeletal elements perform after target enrichment for the human mitochondrial genome. Understanding the variation that exists between skeletal elements regarding DNA yield from remains that have been preserved in extreme environments is a necessary step in optimizing paleogenomics protocols in order to minimize the destruction of human remains.

269 - FLY HIGH & VAPE LOW: GENETIC MECHANISM OF NICOTINE ADDICTION

TITILAYOMI AYOMIDE OTENAIKE

TITILAYOMI AYOMIDE OTENAIKE, Oluwabukola Mary Farodoye, Elgion da Silva Loreto
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL, PORTO ALEGRE, BR, Universidade Federal do Rio Grande do Sul, Universidade Federal do Rio Grande do Sul, Universidade Federal do Rio Grande do Sul, Universidade Federal do Santa Maria

Nicotine, an addictive component of tobacco products, linked to neuronal stress, addiction, and memory loss, has been associated with cancer, neuroinflammation, dementia, Alzheimer's disease, and Parkinson's disease. Nicotine addiction is a multifactorial trait influenced by genetic predisposition, environmental factors, and gene-environment interactions. Despite the prevalence of nicotine and e-cigarette use, the genetic mechanisms of nicotine toxicity, especially in vape usage, are poorly understood. Using Drosophila melanogaster as a model due to its neurological and genomic similarity to humans, we analyzed varying concentrations of nicotine and vape on cognition and examined the genetic effects of nicotine and e-cigarettes on the physiological condition of the flies.
In this study, using aversive conditioning, flies were trained to associate nicotine with shock conditioning. After repeated pairings with

fruit odours, the flies learned fear conditioning. We carried out memory retention analysis and found that nicotine triggers addictive behaviour and induces cognitive changes, a form of behavioural plasticity in nicotine smoking and vaping flies. Also, flies exposed to nicotine showed impaired locomotor activity, elevated levels of reactive oxygen species, and increased gene expression of addiction agonist genes, oxidative stress, apoptosis, and dopamine signalling genes.
Overall, these data indicate that oxidative stress mediates genetic alteration through the alteration of key biochemical and molecular biomarkers in nicotine-treated flies.

271 - A Narrow Range of Transcript-error Rates Across the Tree of Life

Weiyi Li

Weiyi Li, Stephan Baehr, Michael Lynch
Stanford University, Biodesign Center for Mechanisms of Evolution, Arizona State University, Biodesign Center for Mechanisms of Evolution, Arizona State University

The expression of genomically-encoded information is not error-free. Transcript-error rates are dramatically higher than DNA-level mutation rates, and despite their transient nature, the steady-state load of such errors must impose some burden on cellular performance. However, a broad perspective on the degree to which transcript-error rates are constrained by natural selection and diverge among lineages remains to be developed. Here, we present a genome-wide analysis of transcript-error rates from 21 species across three domains of life, showing that the range in error rates is remarkably narrow across diverse species. A model based on cellular and population genetics, incorporating information on cell volume, proteome size, average degree of exposure of individual errors, and effective population size, explains only a fraction of the variation in the error-rate data and suggests that other unobserved parameters, likely negatively correlated with the former, play a role in the evolution of transcription fidelity and mRNA surveillance. The idea that transcript-error rates are adaptively reduced in genes with high expression finds little support in the data.

277 - Unraveling natural Variants in the Gene Regulatory Network underlying Root Hair Formation: A Comprehensive Analysis of 855 Arabidopsis thaliana Accessions

Elsa Herminia Quezada Rodríguez

Elsa H Quezada RodrÃ-guez, Alma PiÃ±eyro Nelson, Adriana Garay Arroyo, Ana M Almeida, Elena Alvarez Buylla
Departamento de ProducciÃ³n AgrÃ-cola y Animal, Universidad AutÃ³noma Metropolitana Xochimilco, Departamento de ProducciÃ³n AgrÃ-cola y Animal, Universidad AutÃ³noma Metropolitana Xochimilco, Departamento de EcologÃ-a Funcional, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico. 3er Circuito Exterior, Ciudad Universitaria, CoyoacÃ¡n, MÃ©xico D.F. 04510., Green Biome Institute. California State University, East Bay. 25800 Carlos Bee Blvd. Hayward/CA. US. 94542, Departamento de EcologÃ-a Funcional, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico. 3er Circuito Exterior, Ciudad Universitaria, CoyoacÃ¡n, MÃ©xico D.F. 04510., Centro de Ciencias de la Complejidad, Universidad Nacional AutÃ³noma de MÃ©xico. 3er Circuito Exterior, Ciudad Universitaria, CoyoacÃ¡n, MÃ©xico D.F. 04510.

A plant's capacity to adapt and respond to diverse challenges is enhanced by the plasticity of the root. Root hairs, thin tubular structures elongating in the root's differential zone, play a pivotal role in water absorption, nutrient uptake, and microorganism interaction. In this research, insights into the genetic diversity of genes that govern root hair development in diverse Arabidopsis accessions was analyzed. We employed computational methods to assess genetic variations in genes that comprise the Gene Regulatory Network (GRN) underlying the development of the hair/no-hair pattern in primary root, analyzing 855 accessions distributed around the world. At the protein level, we conducted an exhaustive analysis that included constructing phylogenetic trees, identification of domains, motifs, families,

Posttranslational modifications (PTMs), protein features, secondary and tertiary protein structure. Through a Single Nucleotide Polymorphism (SNP) search in the GRN, SABRE, EGL3, GL3 and GL2 were found to have a higher number of SNPs. In these four genes, we identified 22 accessions with greater variability in SABRE, 44 for EGL3, 48 for GL3, and 28 for GL2. Moreover, our study highlights five accessions with at least two significant variabilities in one of the four analyzed genes. Finally, Chat-1, Lov-1, Sq-1, and Old-1 accessions were utilized for phenotyping, revealing variations in the trichoblast and atrichoblast patterns, as well as in the quantity of root hairs. These findings provide valuable insights into the intricate molecular processes contributing to root hair developmental plasticity in naturally occurring Arabidopsis accessions.

279 - Modeling site- and branch-level compositional heterogeneity with GFmix

Charley G. P. McCarthy

Charley G. P. McCarthy, Edward Susko, Andrew J. Roger
Institute for Comparative Genomics, Dalhousie University, Department of Biochemistry and Molecular Biology, Dalhousie University, Institute for Comparative Genomics, Dalhousie University, Department of Mathematics and Statistics, Dalhousie University, Institute for Comparative Genomics, Dalhousie University, Department of Biochemistry and Molecular Biology, Dalhousie University

Phylogenetic trees are typically inferred from protein sequence data sampled from diverse taxa. The amino acid composition of this data may be heterogeneous across sites and branches. Mismodeling compositional heterogeneity may produce artifactual trees, but accurate modeling often requires inordinate amounts of time and computational resources.

The novel site-and-branch-heterogeneous model GFmix has seen application in trees with heterogeneous composition of GARP/FYMINK amino acids, as is expected when GC content varies. GFmix modifies a fitted site-heterogeneous profile mixture model with parameters reflecting branch-specific amino acid composition. This approach has been shown to improve likelihood and reduce compositional heterogeneity artifacts. However, the current implementation of the model includes constraints which may not reflect biological reality and is limited to modeling GARP/FYMINK composition.

Here we present a formal analysis of the GFmix model and several improvements to the model. The improved GFmix model loosens previous constraints on branch-specific composition parameters, allows more general groups than GARP/FYMINK, and provides maximum-likelihood optimization of model and tree parameters. We have also developed new methods of detecting compositional heterogeneity directly from sequence data using GFmix. Analysis of simulated data indicates that the improved model better estimates branch-specific composition, branch lengths, and overall compositional heterogeneity.

We then applied GFmix to two real datasets with known compositional heterogeneity artefacts: an 8-taxon Opisthokonta dataset, and a 120-taxon Alphaproteobacteria dataset. We find GFmix reduces support for artifactual trees over non-branch-heterogeneous models.

286 - A Molecular Timetree of Mammals in Zoological Parks

Grace Lebenean Bamba

Grace L. Bamba, Jack M. Craig, Sudhir Kumar
Institute for Genomic and Evolutionary Medicine at Temple University, Institute for Genomics and Evolutionary Medicine at Temple University, Institute for Genomic and Evolutionary Medicine at Temple University

Hundreds of millions of visitors go to zoological parks to see charismatic mammals like lions, elephants, and great apes. During their visit,

they are rarely informed about the divergence times and phylogeny of their closest living taxa, creating a missed opportunity to foster passion for evolution, conservation, and STEM education. There are no global timetrees of mammals at subspecies resolution, which is needed because zoos tend to house local subspecies of mammals. To fill this gap, we have assembled the first timetree of 644 mammals, including 448 subspecies and 13 endangered taxa, commonly found in zoos worldwide. The Zoo-mammals timetree, ZooMaT, is based on manual curation of timetrees drawn from more than 50 published studies, contributing divergence times for 120 new mammals. Additionally, we generated novel timetrees by assembling alignments and building timetrees to address the gaps in the divergence times for more than 60 subspecies. We combined these individual timetrees with the timetree of other common zoo taxa available from the TimeTree 5 database to build ZooMaT. The number of subspecies varied widely across species, none to nearly 20. On average, subspecies divergences were about one quarter of the timing of species origins, with some diverging close to the species origins and others very recent. We will provide an online resource for visitors and zoo-keepers to view ZooMaT online and produce timetrees restricted to mammals found in individual Zoos (www.TimeTree.org/ZooMaT).

300 - Diverse origins of near-identical antifreeze proteins in unrelated fish lineages provide insights into evolutionary mechanisms of new gene birth and protein sequence convergence

Xuan Zhuang

Nathan Rives, Vinita Lamba, C.-H. Christina Cheng, Xuan Zhuang

University of Arkansas, University of Arkansas, University of Illinois, University of Arkansas

Determining the origins of novel genes and the genetic mechanisms underlying the emergence of new functions is challenging yet crucial for understanding evolutionary innovations. The novel fish antifreeze proteins, exemplifying convergent evolution, represent excellent opportunities to investigate the evolutionary origins and pathways of new genes. Particularly notable is the near-identical type I antifreeze proteins (AFPI) in four phylogenetically divergent fish taxa. This study tested the hypothesis of protein sequence convergence beyond functional convergence in three unrelated AFPI-bearing fish lineages, revealing different paths by which a similar protein arose from diverse genomic resources. Comprehensive comparative analyses of de novo sequenced genome of the winter flounder and grubby sculpin, available high-quality genome of the cunner, and those of 14 other relevant species found that the near-identical AFPI originated from a distinct genetic precursor in each lineage, and independently evolved coding regions for the novel ice-binding protein while retaining sequence identity in the regulatory regions with their respective ancestor. The deduced evolutionary processes and molecular mechanisms are consistent with the Innovation-Amplification-Divergence (IAD) model applicable to AFPI formation in all three lineages, a new Duplication-Degeneration-Divergence (DDD) model we propose for the sculpin lineage, and a DDD model with gene fission for the cunner lineage. This investigation illustrates the multiple ways by which a novel functional gene with sequence convergence at the protein level could evolve across divergent species, advancing our understanding of the mechanistic intricacies in new gene formation.

307 - Exploring fitness landscapes in laboratory yeast crosses

Misha Gupta

Misha Gupta, Artur Rego-Costa, Evgeniia Belousova, Michael M Desai

Harvard University, Harvard University, Regeneron, Harvard University, Harvard University

Fitness landscapes are a mathematical tool used to visualize genotype-phenotype maps and have important implications for studying evolution. Populations can be thought of as â€˜navigatingâ€™ these landscapes during adaptive evolution, with evolutionary paths and outcomes constrained by the structure and ruggedness of a landscape. The complexity of these landscapes can also be influenced by interactions across the genome. Studying these fitness landscapes can help predict potential evolutionary trajectories and understand the

constraints and possibilities within evolutionary processes. However, these landscapes are extremely high dimensional and often challenging to explore.

Previous evolution experiments have observed unexpectedly flat fitness landscapes (1). Other quantitative genetic experiments have built a large yeast cross and tried to sample a local neighborhood of the landscape densely (2), showing similarly flat structures. These flat landscapes contradict intuition, especially given complex genetic backgrounds and environmental effects. One hypothesis suggests that this results from sampling small enough local neighborhoods that they appear effectively flat. Our study aims to explore different local neighborhoods of the fitness landscape by generating individuals close to the large yeast cross, yet differing in genetic relatedness. By exploring these disjoint yet nearby neighborhoods in the fitness landscape, we might find some local curvature and complexity in the landscape. On the other hand, flat landscapes may suggest that fitness in laboratory crosses is robust to genetic changes, and such populations may evolve through drift.

(1): https://doi.org/10.1111/j.1558-5646.2011.01569.x
(2): https://doi.org/10.7554/eLife.73983

309 - Low complexity regions across the Tree of Life: sources of diversity or just noise?

Fabia Ursula Battistuzzi

Fabia Ursula Battistuzzi, Vineet Saravanan, Krysta Krawzcyk
Oakland University, Cranbrook Upper school, Oakland University

Low complexity regions (LCRs) are loosely defined as regions of a genome with lower than expected nucleotide or amino acid diversity. Traditionally considered non-functional parts of a genome, they have also been suggested as potential sources of genomic variability. Most of our knowledge regarding LCRs comes from eukaryotic genomes that, with their large amount of non-coding DNA host relatively large numbers of LCRs. However, despite their more compact nature, prokaryotic genomes also have a consistent presence of LCRs, suggesting that similar evolutionary forces driving their origin and evolution may be shared by the three domains of life. The availability of thousands of prokaryotic genomes provides a unique opportunity to evaluate the evolutionary history of LCRs at the population and species levels, thus shedding light on their potential role in genomic variability. In this study, we will present a computational analysis of thousands of prokaryotic species and some example populations to reconstruct patterns of presence/absence and composition of LCRs within these genomes.

315 - Finding LUCA: Bioinformatic tool and database to infer the protein content in the last universal common ancestor

José Emilio Ramírez-Piña

JosÃ© Emilio RamÃ-rez-PiÃ±a, Luis Delaye
Departamento de IngenierÃ-a GenÃ©tica, Cinvestav Irapuato, Departamento de IngenierÃ-a GenÃ©tica, Cinvestav Irapuato

The genetic content of the last universal common ancestor (LUCA) is one of the most intriguing questions in evolutionary biology. Proteins inferred to be present in LUCA participate in transcription, translation, and some metabolic processes. These proteins have been inferred by using different methodologies. Some of them attempt at identifying most likely proteins inherited from LUCA while others correcto for secondary gene losses and others take advantage of the information provided by 3D structure of proteins. As a result of the different methodologies, the sets of proteins inferred to by coded by LUCA do not fully coincide. Here, we propose to build a pipeline for

the inference of the proteins present in LUCA and a database containing such proteins. The novelty of our approach is that the pipeline will include manual curation to identify proteins that were missed by the bioinformatic approach which are most likely false negatives. The inferred proteome will be made accessible to the scientific community through a dedicated Internet database.

317 - Challenging the Gram-Positive/Gram-Negative Dichotomy: Discovery of Gram-Negative Monoderm Bacteria

José Norberto García Miranda

JosÃ© Norberto GarcÃ-a Miranda, Luis JosÃ© Delaye Arredondo, Gabriela Olmedo Ã•lvarez
Centro de InvestigaciÃ³n y de Estudios Avanzados de Instituto PolitÃ©cnico Nacional Unidad Irapuato, Centro de InvestigaciÃ³n y de Estudios Avanzados de Instituto PolitÃ©cnico Nacional Unidad Irapuato, Centro de InvestigaciÃ³n y de Estudios Avanzados de Instituto PolitÃ©cnico Nacional Unidad Irapuato

The classification of bacteria into Gram-positive and Gram-negative groups has long been considered a fundamental divide in microbiology, reflecting differences in cell envelope structure. This dichotomy is closely tied to the presence of one or two cell membranes, leading to the classification of bacteria as monoderms or diderms, respectively. Historically, all diderm bacteria have been characterized as Gram-negative, while monoderm bacteria were presumed to be Gram-positive.

Here, we present groundbreaking findings that challenge this paradigm. We have identified and characterized a previously unrecognized group of Gram-negative monoderm bacteria through comprehensive genomic and phenotypic analyses. What is more interesting is that these gram-negative bacteria occur as three intercalated clades within the Bacillacea clade, where the well-known Bacillus subtilis and Bacillus cereus clades occur.

Our findings also have revealed that, Teichoic Acid (TA), once deemed a hallmark feature of Gram-positive cell walls, is absent in many Gram-positive bacteria and present in the â€œgram negativeâ€• bacillacea. TA is thought to be indispensable for maintaining cell structure, as well as playing roles in cell division, cell wall synthesis, and interactions with the environment.

This work not only contributes to the ongoing discourse on bacterial evolution but also lays the foundation for future investigations into the ecological significance, evolutionary origins, and physiological adaptations of these enigmatic Gram-negative monoderm bacteria. By elucidating the existence of Gram-negative monoderm bacteria, we provide a novel perspective on bacterial cell envelope diversity.

322 - Coevolution and adaptation of TNPs and PRMs for reproductive health in natural ascrotal mammals support the Black Queen Hypothesis

Simin CHAI

Simin CHAI, Jieqiong KANG, Tianzhen WU, Yu ZHENG, Xu ZHOU, Shixia XU, Wenhua REN, Guang YANG
Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Nanjing Normal University, Nanjing Normal University, Nanjing Normal University, Nanjing Normal University, Nanjing Normal University, Nanjing Normal University, Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Nanjing Normal University

Protamines (PRMs) and transition nuclear proteins (TNPs) are two major classes of sperm nuclear basic proteins functioning in chromatin reorganization and condensation in the sperm head, playing key roles in mammalian spermatogenesis and successful fertilization. Generally, cryptorchidism in scrotal mammals such as humans is associated with the elevated incidence of the defect sperm

quality, quantity and function. However, ascrotal mammals such as cetaceans with natural undescended testes, generate the same normal sperm as their scrotal counterparts. Here, we investigated the genomic landscape, evolutionary pattern and protein function change of PRMs and TNPs, aiming to explore the molecular mechanisms underlying the reproductive health in natural ascrotal mammals. We found that the absence of the intact PRM2 and TNP2 coevolved and was significantly associated with divergent testicular positions, despite the conserved genomic arrangement in most mammals. Importantly, greater thermostability and DNA binding of PRM1, and superior DNA repair capacity of TNP1 in cetaceans, the representative ascrotal mammals lacking the intact PRM2 and TNP2, were revealed. It was further shown that multi-point mutations induced from ascrotal cetaceans had significance of functional changes. Thus, we hypothesized and concluded that, in response to physiological challenges posed by natural â€œcryptorchidismâ€•, most ascrotal lineages evolved by adaptive genomic reduction of PRM2 and TNP2 while enhancing the protein function of PRM1 and TNP1. This evolutionary strategy, which enhances chromatin condensation efficiency and benefits DNA repair during spermatogenesis, likely represents a mechanism for reproductive health in naturally cryptorchid mammals, within the context of the Black Queen Hypothesis.

323 - If the door is locked, take another one: Evolution of alternative entrance points for a parasite

Eglantine MATHIEU-BEGNE

Eglantine MATHIEU-BEGNE, Sabrina GATTIS, Dieter EBERT

Department of Environmental Sciences, Zoology, University of Basel, School of Zoology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Department of Environmental Sciences, Zoology, University of Basel

The acquisition of alternative routes of infection by parasites is a way to circumvent host resistance. This strategy makes only sense, if host resistance to alternative routes is based on distinct loci. We investigated this hypothesis by examining whether the alternative infection routes of the parasite Pasteuria ramosa in its crustacean host, Daphnia magna, are associated with different resistance loci. We investigated the genomic basis of host resistance to a strain of P. ramosa, P15, able attach and penetrate into the host either from the hindgut or the foregut of its host. We used a worldwide panel of 174 D. magna genotypes with further refinement through breeding experiments. Our findings support our hypothesis: hindgut attachment is determined by the D locus, whereas foregut attachment is governed by a newly identified locus, named the G locus. We established a gene model for the G locus, which predicts Mendelian segregation and epistatic interaction with at least one other P. ramosa resistance locus. Through genomic Pool-sequencing data analysis, we localized the G locus within a know Pasteuria Resistance gene cluster on chromosome 4, whereas the D locus is located on chromosome 7. Two candidate genes for the G locus, both belonging to the Galactosyltransferase gene family, were identified. Our study sheds a new light on how host and parasite co-evolve and enhances our understanding of how parasites evolve infection strategies.

326 - Adaptive Genomic Signatures and Evolutionary Mechanisms in Anguillid Eels

Hyeongwoo Choi

Hyeongwoo Choi, Jiwon Nam, Seong-il Eyun

Chung-Ang University, Chonnam National University, Chung-Ang University

Anguillid eels undergo a catadromous life cycle, transitioning from deep-sea larvae to adults in estuaries and freshwater environments. Their unique life cycle makes them an ideal model for exploring evolutionary shifts between saltwater and freshwater habitats, as well as between deep-sea and shallow-water environments. Despite inhabiting diverse ecological niches for over 21 million years, the evolutionary mechanisms of anguillid eels remain ambiguous. Hence, we generated the chromosome-level genome sequence of the short-finned eel, Anguilla bicolor pacifica, aiming to identify positively selected genes (PSGs) and elucidating their functions. Using the branch model, we detected 82 genes exhibiting positive selection on the anguillid branch ($p < 0.05$). It is widely known that adult eels

cease feeding and experience gut degeneration during their migrating back to the deep-sea. Significantly, we observed PSGs associated with pathways involved in autophagy, longevity regulating, and peroxisome, suggesting their role in energy production during nutrient scarcity. Also, eels that migrate to the oxygen-minimum layer for spawning experience relaxed evolutionary constraint on the AMPK- and HIF-1 signaling pathways among the 82 identified PSGs, potentially contributing their adaptation to hypoxic environments. Comparative genomic analysis revealed that four expanded gene families in the four eel species, predominantly enriched in visual pathways such as NTN1, SIXs, and MAP1B ($p < 0.05$). In conclusion, this study provides valuable insights into the adaptability mechanisms in anguillid eels, shedding light on how they manage to thrive in diverse environment.

340 - Diving deeper: demystifying the distribution of the ‘rare’ deep-sea amphipod, Alicella gigantea.

Paige J. Maroni

Paige J Maroni, Yakup Niyazi, Alan J Jamieson

Minderoo-UWA Deep-Sea Research Centre, School of Biological Sciences and Oceans Institute, The University of Western Australia, Minderoo-UWA Deep-Sea Research Centre, School of Biological Sciences and Oceans Institute, The University of Western Australia, Minderoo-UWA Deep-Sea Research Centre, School of Biological Sciences and Oceans Institute, The University of Western Australia

Amphipods (Arthropoda: Crustacea) inhabit all aquatic environments worldwide and are amongst the most specious and ecologically diverse orders of crustaceans encompassing over 10,000 extant species. Their ubiquity extends to the deep-sea, particularly the abyssal (3000-6000 m) and hadal environments (6000-~11000 m), yet, due to sporadic and infrequent sampling, plus a lack of physical specimens retrieved from the deep-sea, most amphipod species are classified as â€œrareâ€•. One such example, Alicella gigantea Chevreux, 1899, known as the â€œsupergiant amphipodâ€• has historically been collected in small numbers, perhaps signifying low population densities, and providing a sense of rarity. Consequently, little is known about the demography, genetic variation, and population dynamics of A. gigantea with only six studies having published DNA sequence data. As more records emerge across the vastness of the deep-sea, and from depths that are rarely sampled, there is an ever-growing body of evidence to show that A. gigantea should be considered far from rare. In this study, we compile all records of A. gigantea and use two mitochondrial and one nuclear gene (COI, 16S and 28S) to explore the distribution patterns of A. gigantea across all oceans, further examine the phylogeographic structure across the species and discuss the species history throughout tectonic time. Specifically, we utilised time-calibrated phylogenies and habitat mapping software to estimate divergence times and discuss these speciation events in relation to global oceanographic and geological events, while also investigating haplotype and phylogenetic data to understand this species contemporary distribution.

345 - Identifying the Ancestral Origins of Genetic Variants in Pediatric Obese Asthma with Population Genetics

David J Yang

David Yang, Anthony Griffen, Mariko Isshiki, John Greally, Srilakshmi Raj, Deepa Rastogi

Albert Einstein College Of Medicine, Albert Einstein College Of Medicine, Albert Einstein College Of Medicine, Albert Einstein College Of Medicine, Albert Einstein College Of Medicine, Albert Einstein College Of Medicine, George Washington University School of Medicine

Obesity-related asthma disproportionately affects minority children and is associated with nonatopic T-helper type 1 (Th1) cell polarized inflammation. Prior analysis by Rastogi et al. (2020) of the transcriptome and DNA methylome of the CD4+ Th cells collected from Hispanic and African American children revealed upregulation and differential methylation of genes in the Rho-GTPase pathways in those with obesity-related asthma. eQTL-gene analysis also identified RPS27L among 157 differentially expressed genes which was downregulated in obese asthma Th cells compared to those from normal-weight children with asthma.

RPS27L gene expression did not differ between groups when comparing across alleles, although all children homozygous for the minor allele (rs69494395) were in the obese asthma group, suggesting that genotype rather than study group was associated with gene expression. The minor allele, more common in the Latino and Afro-Caribbean populations than in white populations was an independent predictor of lower gene expression.
Using local ancestry inference, we aimed to identify the genetic ancestry of rs69494395. Current approaches for treating pediatric asthma are insufficient for pediatric obese asthma. There are different obesity-related mechanisms and underlying genetic variants that influence its pathophysiology. Our follow-up analyses suggests that in most individuals with rs69494395, the allele occurs on an African ancestry haplotype, particularly in homozygous individuals. Our finding is consistent with other studies showing that several alleles that contribute to increased asthma risk are also derived from African ancestry.

346 - The effect of genetic inheritance impact on de novo mutation patterns in Macaca fascicularis.

Hye Ri Park

Hye-ri Park, Hyeon-Mu Cho, Se-Hee Choe, Min-Gyeong Ko, Yun-Jung Lee, Sang-Je Park, Young-Hyun Kim, Jae-Won Huh
Korea Research Institute of Bioscience and Biotechnology(KRIBB), Korea University of Science and Technology (UST), Korea Research Institute of Bioscience and Biotechnology(KRIBB), Korea University of Science and Technology (UST), Korea Research Institute of Bioscience and Biotechnology(KRIBB), Korea Research Institute of Bioscience and Biotechnology(KRIBB), Korea University of Science and Technology (UST), Korea Research Institute of Bioscience and Biotechnology(KRIBB), Korea University of Science and Technology (UST), Korea Research Institute of Bioscience and Biotechnology(KRIBB), Korea Research Institute of Bioscience and Biotechnology(KRIBB), Korea Research Institute of Bioscience and Biotechnology(KRIBB), Korea University of Science and Technology (UST)

Mutations arise from errors in the DNA replication process or exposure to environmental elements like radiation or chemicals. And occurring within germ cells possess the potential to be passed down to successive generations, perpetuating their genetic legacy. Whether the process of transferring unrepaired mutations to the next generation is repeated, it can proceed to microevolution that changes population gene pool at the species level. In the case of sexual reproduction, compared to female, male germ cells have more opportunities to mutate due to spermatogenesis during whole lifetime. In the present study, we perform whole genome sequencing (WGS) to analysis of somatic/germline mutations accumulated within individuals and patterns of mutations transmitted to offspring by parentsâ€™ age in crab-eating monkeys (Macaca fascicularis). We selected 4 families (total 16 monkeys; 10 males, 6 females) around various age to get data at all age point and considered that they can produce 3rd generation for detecting transmission to next generation. Also, we considered families with males as 2nd generation offspring in order to directly confirm the effect of transmitted mutations on the reproductive cells of the child. As a result of 3 years process, WGS data of monkey blood was built up, and based on that, diverse pipelines were devised to analyze mutations accumulated at each sampling time. From these analyses, we demonstrate the parental effect of de novo mutations accumulating with age in crab-eating monkeys throughout whole life on next generation.

349 - Telomere length estimation in the blood of the cynomolgus monkey using the sequencing data

YUN JUNG LEE

Yun-Jung Lee, Se-Hee Choe, Hyeon-Mu Cho, Hye-Ri Park, Min-Gyeong Ko, Sang-Je Park, Young-Hyun Kim, Jae-Won Huh
Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea University of Science & Technology (UST), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea University of

Science & Technology (UST), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea University of Science & Technology (UST), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea University of Science & Technology (UST), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Korea University of Science & Technology (UST)

Telomere is structure with TTAGGG repeat sequences located at the ends of chromosomes, maintaining chromosome stability and protecting chromosomes from degradation. Due to the end replication problem where the ends of linear DNA cannot be fully replicated, telomeres gradually shorten. It is believed that the telomere length (TL) is a representative hallmark of aging. Initially, experimental methods were used to estimate TL, but advances in whole genome next-generation sequencing (NGS) technology and programming tools now allow the length to be estimated using computational methods. Macaca fascicularis is an experimental animal model suitable for aging research because it is evolutionarily close to humans. This study was conducted because there was no paper measuring the TL of monkeys using a computational method. We selected a total of eight individuals, two of the same age each, and performed paired-end sequencing for two years. We used Computel and Telomerecat to measure TL using sequencing data. The sequencing result files were manipulated for using different tools: Computel is an R-based program that takes the fastq file as input, while Telomereat, a Python-based program, uses the bam file as input. The results have shown that both tools showed shorter telomere lengths compared to previous years. When calculate the correlation coefficient of TL by age, Telomereat's is more negative than Computel's. Due to small sample sizes, neither analysis tools revealed significance for telomere length changes with age; however, we identified a decline in TL with age using Macaca WGS data.

353 - Origin and evolution of specialised ribosomes across eukaryotes

Alan J. S. Beavan

Alan J. S. Beavan, Mary J O'Connell

University of Nottingham, University of Nottingham

Ribosomes are essential to all cellular life. Evidence is emerging that ribosomes are not the unprejudiced translational machines we once thought, thus challenging the idea that gene regulation is driven principally by transcriptional regulation. Instead, heterogeneous pools of ribosomes exist within and between cells and these structurally different ribosomes appear to have differential capacities for translation of different sets of mRNAs. It has been previously shown that ribosomes can be â€œspecialisedâ€• in this way through the inclusion of closely related ribosomal protein paralogs with distinct sequences. To identify all such duplication events, we use gene tree-species tree reconciliation methods to elucidate the relationships between all ribosomal protein coding homologs, mapping gene duplication events that facilitate ribosome specialisation onto a eukaryote phylogeny. Further, we identify sites that have undergone rapid evolution in cases of specialisation, and in conjunction with 3D models of ribosomal complexes, propose specific regions of the ribosome that are most likely to facilitate specialisation. Our high-throughput method features the integration of standard phylogenetic methods with 3D protein structural methods - putting the evolution of specialised ribosomal complexes into a 3D context. In the future, these methods will be used not only to elucidate how specialisation has emerged in known cases but also to identify potential new cases of specialisation.

357 - Seascape genomics and genetic basis of adaptation to low salinity of Harbour porpoises (Phocoena phocoena) across environmental gradients in the North Atlantic and adjacent waters

Enrique Celemin Amaro

Enrique Celemin Amaro, Ralph Tiedemann

Potsdam University, Potsdam University, Potsdam University

The Harbour porpoise (Phocoena phocoena) is a highly mobile cetacean species primarily occurring in coastal and shelf waters across the Northern hemisphere. It inhabits heterogeneous seascapes broadly varying in salinity and temperature. Here we produced 82 whole genomes at intermediate coverage to study Harbour porpoiseâ€™s evolutionary history and investigate the role of local adaptation in the diversification into subspecies and populations. We identified ~6 million high quality SNPs sampled at 8 localities across the North Atlantic and adjacent waters, which we used for population structure, demographic, and genotype-environment association analyses. Our results suggest a genetic differentiation between three subspecies (P.p. relicta, P.p. phocoena and the recently proposed P.p meridionalis), and three distinct P.p. phocoena populations occurring in the Baltic: Atlantic, Belt Sea and Proper Baltic Sea. Effective population size and Tajimaâ€™s D levels suggest a population contraction in Black Sea and Iberian porpoises, but a population expansion in the P.p. phocoena populations. Genotype-environment association analysis identified salinity as the major environmental driver in genomic variation and we identified candidate genes putatively underlying adaptation to different salinity levels in Baltic porpoises. Our study highlights the value of whole genome resequencing to unravel the genetic basis underlying adaptation to brackish waters and shows how strong environmental gradients may lead to population differentiation. The results have great conservation implications as we found inbreeding and low genetic diversity in the endangered Black Sea subspecies and identified the critically endangered Proper Baltic Sea porpoises as a separate population.

361 - Genetic architecture of adaptation varies with the strength of selection

Deepa Agashe

Shivansh Singhal, Deepa Agashe

National Centre for Biological Sciences (NCBS), National Centre for Biological Sciences (NCBS)

In new environments, populations face novel selection pressures, and adaptation to such conditions is of interest to evolutionary biologists, conservation biologists and ecologists alike. The genetic architecture of adaptation â€“ i.e., the number and effect sizes of loci involved in adaptation, varies between species and sometimes even between populations of the same species. Theoretical models predict that the genetic architecture of adaptation changes with the strength of selection, but empirical evidence for this phenomenon is scarce. Under weak selection, we expect to see many loci with small effects on adaptive traits; whereas with strong selection, a few highly beneficial loci are likely to drive adaptation. We tested this using laboratory populations of the red flour beetle Tribolium castaneum adapting to novel resources (ancestral resource: wheat, novel resources: corn, finger Millet and sorghum, in decreasing order of selection pressure). After 35-60 generations of adaptation, we sequenced pooled genomic DNA from ancestral and evolved lines and quantified the change in allele frequencies over time. Compared to the ancestor, evolved lines had reduced nucleotide diversity and many more alleles at extreme frequencies. Supporting theoretical predictions, the magnitude of allele frequency change (a measure of effect size) as well as the number of alleles with large effect was correlated with the strength of selection imposed in each resource. Allele frequency changes were also highly parallel across populations, suggesting widespread beneficial pleiotropy. Thus, the genetic architecture of adaptation is associated with the strength of selection, suggesting broad predictability of adaptation in different environments.

369 - A new method for estimating additive genetic variance for fitness using genome-wide allele frequency change data

Manas Geeta Arun

Manas Geeta Arun, Darren Obbard, Aidan Angus-Henry, Jarrod Hadfield

Institute of Ecology and Evolution, the University of Edinburgh, Institute of Ecology and Evolution, the University of Edinburgh, CharitÃ© â€“ UniversitÃ¤tsmedizin Berlin, Institute of Ecology and Evolution, the University of Edinburgh

Fisherâ€™s Fundamental Theorem of Natural Selection shows that the increase in the mean fitness of a population due to one generation of natural selection is equal to the additive genetic variance in relative fitness (Vw) in the absence of drift, migration and mutation. Traditional methods for empirical estimation of Vw typically involve measuring proxies of fitness on a large number of individuals from a known pedigree. Data are hard to collect and the results are often highly sensitive to the definition of fitness used. Here, we present a new method for estimating Vw that does not involve making measurements of fitness proxies on individuals, but instead tracks changes in the genetic composition of a population. First, we show that Vw can readily be expressed as a function of the genome-wide diversity/LD matrix and genome-wide expected allele frequency change due to selection. We then show how independent experimental replicates can be used to infer expected allele frequency change due to selection and then estimate Vw with the help of a linear mixed model. Finally, we test whether our approach gives precise and unbiased estimates of Vw using forward simulations performed using SLiM over a range of biologically plausible levels of Vw. Looking forward, we will apply this method to data from a large experiment on Drosophila melanogaster that is near completion.

371 - Population genetic structure of Himalayan Langurs

Sharwary M R

Sharwary M R, Kunal Arekar, Praveen Karanth

Indian Institute of Science, Indian Institute of Science, Indian Institute of Science

Himalayan Langurs (Semnopithecus schistaceus) are one of the most widely distributed colobine monkeys found in the Himalayas from Pakistan in the west to Bhutan in the east. Further, their distribution encompasses a wide range of altitudes (300-3000 meters) and is interspersed with numerous deep river valleys. In this study, we investigate the role of riverine barriers and altitudinal gradients in shaping the population genetic structure of these langurs. Previous approaches relied on mitochondrial markers, but this study employs nuclear microsatellite markers for enhanced genetic resolution. Fecal samples were non-invasively collected across 4 Indian Himalayan states: Jammu and Kashmir, Himachal Pradesh, Uttarakhand, and Sikkim based on distribution records from past studies. DNA extraction, PCR amplification of 7 microsatellite regions, and microsatellite genotyping were conducted followed by analysis using Arlequin, PCoA, STRUCTURE, Neighbor-joining tree, AMOVA, and paired Mantel test with pairwise F_ST and pairwise altitudes. The preliminary results show extensive gene flow across riverine barriers and altitudinal gradient indicating a lack of significant genetic differentiation. This implies that the Himalayan heterogeneous landscape features (altitude and rivers) might not be significantly affecting the genetic diversity of these langurs. These results do not support splitting the Himalayan langurs into multiple species/subspecies.

373 - Phylogenomics provides insights into the evolution of dragonflies and damselflies

Zifeng Li

Zifeng Li

Chinese Academy Of Science (Kunming Animal Research Institution)

Odonata is a group with high species diversity in the Insecta, but its phylogenetic relationship with Ephemeroptera and Neoptera has been controversial. Although previous studies have attempted to provide a strong argument for this dispute based on morphological classification and mitochondrial genes, powerful evidence from the genome-wide level is still lacking. Dragonflies are a group with a very high success rate of predation among flying insects which is related to the visual ability of dragonflies. How the dragonflyâ€™s visual ability evolved to cope with this specialized hunting strategy and various environmental clues in an aerial space remains unknown.

Here, for the first time, our study systematically analyzed the phylogeny and compound eye evolution of dragonflies by combining long-read genome sequencing, bulk transcriptome sequencing and single-cell transcriptome sequencing. Our phylogenetic analysis revealed that Ephemeroptera was sister to Odonat+Neoptera. Gene family analysis showed that the gene families of rhodopsin gene Rh6 had significant gene family expansion in the ancestor of the Odonata. Furthermore, we analyzed the transcriptome data of 26 tissues of the Chlorogomphus papilio and found that Rh6_Cp1, Rh6_Cp2, Rh6_Cp3 and Rh6_Cp4 were all tissue-specific highly expressed genes in compound eyes. The single cell transcriptome results showed that there were 16 cell types in the compound eye of Chlorogomphus papilio, among which Rh6_Cp1, Rh6_Cp2 and Rh6_Cp3 were highly expressed in the photoreceptor cell types. Our results provide key sources for understanding the association between gene duplication in Odonata genome and the evolution of visual ability.

389 - European amphioxus' enormous genetic diversity explained by asymmetric migration and long-standing high effective population size

Marina Brasó-Vives

Marina BrasÃ³-Vives, Diego A HartasÃ¡nchez, StÃ©phanie Bertrand, Hector Escriva, Marc Robinson-Rechavi
Department of Ecology and Evolution, University of Lausanne, Swiss Institute of Bioinformatics (SIB), Department of Computational Biology, University of Lausanne, Oceanological Observatory of Banyuls-sur-Mer (OOB), CNRS-Sorbonne University, Oceanological Observatory of Banyuls-sur-Mer (OOB), CNRS-Sorbonne University, Department of Ecology and Evolution, University of Lausanne, Swiss Institute of Bioinformatics (SIB)

Amphioxus are marine filter-feeding chordates that reproduce by external fertilization. The European amphioxus (Branchiostoma lanceolatum) has an ecological range extending at least from the northeastern Atlantic Ocean to the Mediterranean Sea. The amphioxus lineage has had a slow phenotypic evolution and a low substitution rate since their split with the other chordate lineages around 500 Mya. Curiously, previous estimates extracted from single individual genomes suggest amphioxus have one of the most diverse genomes ever measured. Here, we use genomic data from two European amphioxus wild populations (Mediterranean and Atlantic) to accurately quantify and understand their genomic diversity in the context of their geographic history. We confirm both the low substitution rate of the amphioxus lineage and the extremely high levels of genomic diversity in amphioxus wild populations. We confidently estimate that at least 3% of sites are heterozygous in each individual genome. Additionally, our results show two distinct populations of amphioxus in the Mediterranean and the Atlantic and suggest that the former one was originally founded by the latter probably during the refill of the Mediterranean Sea around 5.3 Mya. We also find evidence for maintained unidirectional gene flow from the Atlantic to the Mediterranean populations. These results point towards a long-standing high effective population size as responsible for the enormous genomic diversity of amphioxus. This study highlights the exceptionality of amphioxusâ€™ evolution and contributes to the understanding of the relation between historical geographical range distribution and genomic diversity.

391 - Evolution of Eukaryotic Transcription Factor Co-dependence Through Enhanced Activation and Reduced co-TF Interaction

Bin He

Lindsey F. Snyder, Emily M. O'Brien, Jia Zhao, Thomas Cassier, Yuning Zhang, Raluca M. Gordan, Nicholas Schnicker, Xu Zhou, Bin Z. He
University of Iowa, University of Iowa, University of Iowa, University of Iowa, Duke University, Duke University, University of Iowa, Boston Children's Hospital, University of Iowa

Combinatorial control by multiple transcription factors (TFs) is prevalent in eukaryotic gene regulation, offering increased specificity and

allowing for information integration. However, the genetic and biochemical basis for the dependence is poorly understood, and little is known about its evolutionary potential to cause regulatory divergence. We exploit divergence in co-TF dependence in the phosphate starvation response system in related yeasts to elucidate these questions. In Saccharomyces cerevisiae, the main TF, Pho4, relies on the co-TF Pho2 to regulate ~28 genes. By contrast, Pho4 in a related yeast pathogen Candida glabrata exhibits significantly reduced Pho2 dependence and expanded its targets to ~70 genes. Biochemical and genetic dissections showed that 1) the two Pho4 orthologs bind to indistinguishable motifs by ChIP-seq and Protein Binding Microarray; 2) chimeric TFs coupled with a single-cell reporter identified a region adjacent to the DNA binding domain in S. cerevisiae Pho4 as a double-edged sword: it both allows for enhanced TF activity with Pho2, but also constrains the activity when the co-TF is absent; 3) lacking the ability to use the co-TF, C. glabrata Pho4 binds the consensus motif with 3-4 times lower Kd, and also encodes two unique activation booster domains that enhance the main activation domain conserved between the two Pho4â€™s. Our results provide a detailed molecular picture of how co-TF dependence is mediated and how evolution in functional domains of the TF can lead to changes in codependency, resulting in an expansion of the targets.

398 - Phenotype bias in ancestral sequence space shaped the diversification of a protein family * This is a joint submission for the Graduate Student Excellence Award. Co-leading authors: Santiago Herrera Álvarez and Jaeda E. J. Patton

Jaeda E. J. Patton

Santiago Herrera-Ã•lvarez, Jaeda E. J. Patton, Joseph W. Thornton
University of Chicago, University of Chicago, University of Chicago

Why certain phenotypes evolve, and why they do so in some lineages but not others, has long fascinated biologists. Drift and selection can only act on variation produced by mutation, so evolutionary history may be shaped by the spectrum of phenotypic variation accessible to a biological system and by changes in that spectrum as lineages evolve. We addressed this question at the biochemical level by combining deep mutational scanning to experimentally measure all possible phenotypes of all possible genotypes within a region of sequence space, with ancestral protein reconstruction to characterize how this genotype-phenotype (GP) map changed during history. We characterized all possible variation in binding specificity that can be produced by the protein-DNA interface of steroid hormone receptors and their cognate response elements using two ancestral receptors that flank a historical change in specificity. In each map, the wild-type specificity is encoded more frequently than all others, and specificities are clustered in sequence space. This structure affects the propensity for particular phenotypes to evolve from others over short and long time scales. In the deepest ancestor, a sparsely connected GP map dominated by the ancestral specificity favored conservation of the ancestral phenotype. In its descendant, a change in connectivity and phenotypic bias made acquisition of the derived specificity the most likely outcome. Our findings show that the spectrum of accessible phenotypic variation is nonrandom and can vary over short and long scales of genetic divergence, contributing to the direction of historical evolution.

399 - Phenotype bias in ancestral sequence space shaped the diversification of a protein family * This is a joint submission for the Graduate Student Excellence Award. Co-leading authors: Santiago Herrera-Álvarez and Jaeda E. J. Patton

Santiago Herrera-Álvarez

Santiago Herrera-Ã•lvarez, Jaeda E.J. Patton, Joseph W. Thornton
University of Chicago, University of Chicago, University of Chicago

Why certain phenotypes evolve, and why they do so in some lineages but not others, has long fascinated biologists. Drift and selection

can only act on variation produced by mutation, so evolutionary history may be shaped by the spectrum of phenotypic variation accessible to a biological system and by changes in that spectrum as lineages evolve. We addressed this question at the biochemical level by combining deep mutational scanning to experimentally measure all possible phenotypes of all possible genotypes within a region of sequence space, with ancestral protein reconstruction to characterize how this genotype-phenotype (GP) map changed during history. We characterized all possible variation in binding specificity that can be produced by the protein-DNA interface of steroid hormone receptors and their cognate response elements using two ancestral receptors that flank a historical change in specificity. In each map, the wild-type specificity is encoded more frequently than all others, and specificities are clustered in sequence space. This structure affects the propensity for particular phenotypes to evolve from others over short and long time scales. In the deepest ancestor, a sparsely connected GP map dominated by the ancestral specificity favored conservation of the ancestral phenotype. In its descendant, a change in connectivity and phenotypic bias made acquisition of the derived specificity the most likely outcome. Our findings show that the spectrum of accessible phenotypic variation is nonrandom and can vary over short and long scales of genetic divergence, contributing to the direction of historical evolution.

405 - From diffusion to network: A Neurexin view of the Origin of Neural Synapse

Christine Guzman

Christine Guzman, Kurato Mohri, Hiroshi Watanabe

University of Fribourg, Okinawa Institute of Science and Technology, Okinawa Institute of Science and Technology, Okinawa Institute of Science and Technology

In the nervous system, direct information exchange occurs through specialized intercellular junctions called synapses, whose protein organization and functional properties have been studied in Bilateria. However, the knowledge of when or how the functional synapse first emerged is still lacking. In this study, we used the non-bilaterian model Nematostella vectensis to investigate the role of Neurexins (Nrxns), a family of core pre-Synaptic Cell Adhesion Molecules (SAMs), in an early nervous system. We observed two classes of Nrxns: a broadly expressed 'classical' Nrxn (cNrxn), and a neuronal-specific Nrxn (nNrxn). Functional analysis of the cNrxn revealed its major involvement in cell-cell adhesion between ectodermal and endodermal epithelia. Meanwhile, knockdown of nNrxns resulted in abnormal behaviors associated to polypsâ€™ muscle contractions. Further experiments suggested that nNrxns are involved in chemical transmission rather than peptidergic signaling. This study provides molecular, functional, and cellular insights into the ancestral non-neural function of Nrxns and may explain how and why this cell adhesion molecule family was employed in the synaptic machinery of the early nervous system.

415 - Disentangling the Ancestral History and Population Genetic Structure of Modern Iranian Populations

Nada Mohamed Ibrahim Elsharkawy

Nada M Elsharkawy, Mark Penjueli, Aashish Jha, Carlos D. Bustamante, Roxana Daneshjou, Alexander Ioannidis

The Genetic Heritage Group, New York University Abu Dhabi, The Genetic Heritage Group, New York University Abu Dhabi, The Genetic Heritage Group, New York University Abu Dhabi, Department of Biomedical Data Science, Stanford University, Department of Biomedical Data Science, Stanford University, Department of Biomolecular Engineering, University of California Santa Cruz

Iran is considered a crucial hub for human migration throughout history, yet the Iranian populations have not been adequately represented in genomic studies. This limitation has restricted our understanding of the contributions Iranian peoples have made to modern human genomic diversity and population histories. This study aims to bridge this gap by analyzing whole genome sequences (WGS) of 73

contemporary Iranians from 24 diverse ethnic populations, in the context of WGS data from 4,000 individuals from worldwide human populations. Our findings reveal distinct admixture patterns with populations from neighboring countries, showcasing unique genetic influences and admixture patterns reflective of the region's historical interconnections, notably with Europe, Western Asia, South Asia, and Central Asia. However, these admixture patterns appear to vary depending on ethnic affiliation. By integrating new genomic data with existing datasets, we examine three critical aspects: gene flow into Iran, gene flow out of Iran, and the temporal dynamics of these processes, and whether these dynamics differ across the different ethnic affiliations within Iran. The research contributes to a broader understanding of the pivotal role Iran has played in human prehistoric migrations and lays the groundwork for future analyses to further understand the evolutionary pathways that have shaped the population structures in this pivotal geographical area.

416 - Hi-reComb: constructing recombination maps from bulk sperm Hi-C sequencing

Milan Malinsky

Milan Malinsky
University of Bern

Recombination is central to genetics and to evolution of sexually reproducing organisms. However, obtaining accurate estimates of recombination rates, and of how they vary along chromosomes, is not straightforward. To advance our ability to estimate recombination rates, we present Hi-reComb, a new method and software for estimation of recombination maps from bulk sperm chromosome conformation capture (Hi-C). With a straightforward workflow, Hi-reComb produces robust results with resolution surpassing other gamete sequencing approaches and will be applicable to a broad range of studies and model organisms in genetics and evolutionary biology.

434 - Rapid evolution and functional diversification of the Yellow/MRJP gene family in Lepidoptera

Krushnamegh Kunte

Saurav Baral, Krushnamegh Kunte
National Centre for Biological Sciences - TIFR, National Centre for Biological Sciences - TIFR

Gene duplication is a crucial evolutionary mechanism that generates new genetic material for evolutionary innovation. While gene duplications often risk non-functionality, they can also lead to significant evolutionary advancements. The Yellow/MRJP gene family exemplifies such a case, where duplication and subsequent divergence have contributed to the evolution of novel phenotypes. Despite being a highly conserved gene family present in all insect species, the Yellow/MRJP gene family exhibits significant variation in copy number even among closely related species. This paradoxical nature makes it an ideal candidate for studying the evolutionary dynamics underlying gene duplication and copy number variation. We analyzed the patterns of molecular as well as protein structural evolution across twelve members of the Yellow/MRJP gene family in Lepidoptera. We found that each member of the gene family has distinct expression pattern â€“ indicative of functional divergence â€“ and molecular evolutionary dynamics. Some genes within the family have a developmental stage specific expression, whereas some are expressed in high quantity across all developmental stages, indicative of a housekeeping role. Such a dynamic expression profile was mirrored by dynamism in rates of molecular evolution. Some genes show an active evolution with recent and species-specific duplications, while others have remained as single-copy genes. These findings suggest that the Yellow/MRJP gene family has retained highly conserved ancient paralogs alongside rapidly evolving homologs indicating an ongoing process of active evolution and functional diversification, which plays a considerable role in insect evolution.

448 - Error correction of of viral genomic sequences using a phylogenetic prior

Monica Arniella

Monica Arniella, Lenore Pipes, Rasmus Nielsen
University of California Berkeley, University of California Berkeley, University of California Berkeley

Over the course of the SARS-CoV-2 pandemic, there has been a fast-paced and widespread effort to sequence clinical and wastewater samples around the world. Analysis of these sequences has quickly revealed potentially problematic sites with high error rates, necessitating careful filtering or masking of regions in the genome. Additionally, the error rate of genomic sequencing can be highly variable between experiments, and areas of low sequencing depth make base calling even more error-prone due to ambiguity in the identity of nucleotides. In the case of viral sequences, we can use additional information from phylogenies in order to infer the most probable nucleotide at a position. Here, we use the global SARS-CoV-2 phylogeny (involving millions of sequences from around the world) and the phylogenetic method tronko by Pipes et al. to calculate prior probabilities for each nucleotide along randomly chosen genomes, and use the observation of sequencing reads to choose the bases with the highest posterior probability. To test the performance of this method, we test varied sequencing conditions using real samples downloaded from the Sequencing Read Archive (SRA), simulated reads from genomes obtained from the GISAID repository, and compare the performance of our method with standard viral base calling methods. Using this approach, we evaluate the utility of using phylogenetic information for more confident base-calling at uncertain sites and increasing the accuracy of downstream analyses.

449 - Stickwitu: a macroevolutionary approach reveals adhesive organs associated traits in exclusive meiofauna phyla

Sónia CS Andrade

ThainÃ¡ Cortez, Maikon Di Domenico, SÃ³nia CS Andrade
Universidade de Sao Paulo, Universidade Federal do ParanÃ¡, Universidade de Sao Paulo

The meiofauna encompasses all animals ranging from 40 to 1000 Î¼m in size with representatives in 23 metazoan phyla. Some include clades with few representatives of meiofauna (as Annelida, Nemertea, Mollusca), while others are exclusively meiofaunal (Kinorhyncha, Micrognathozoa, Gnathostomulida, and Loricifera). Certain traits are well-established as potentially convergent among these groups, characterizing adaptations to the interstitial environment, such as the presence of adhesive glands for adhesion to grains. To understand the molecular mechanisms potentially associated with these adaptations, we used a comparative phylogenomic approach to identify 'novel' and 'ancestral' genomic regions among these species, which may be associated with the traits of interest. For this purpose, we analyzed genomic data from 162 marine invertebrates taxa, being 64 meiofaunal. After identifying 577,966 homologous groups (HG) with MCL, we employed PAPS pipeline (Phylogenetic_Aware_Parsing_Script) to infer ancestral HGs and novel-core HGs. A significant number of HGs were identified for the branches with exclusive meiofaunal species. Several gene ontologies with relevant metabolic roles were found, including some associated with body movement, such as muscle contraction, regulation of cilium movement, as well functions related to adhesion and secretion production, such as cell-cell adhesion via plasma membrane adhesion molecules, and proteoglycan biosynthetic process, pathways that might be linked to the functions of adhesive organs. Therefore, by comparing these genomic regions, we were able to assess molecular mechanisms at genomic scale, providing insights into the evolutionary processes underlying the key traits essential for the success of meiofauna in the interstitial environment.

457 - Elucidating the Evolution of the Recombinational Landscape of Placental Mammals Using Comparative Genomics

Isabella Rose Childers

Isabella R Childers, Nicole M Foley, Kevin R Bredemeyer, Terje Raudsepp, William J Murphy
Texas A&M University, Texas A&M University, Texas A&M University, Texas A&M University, Texas A&M University

Genomes are composed of varying evolutionary histories, influenced by processes like gene flow and ILS. Emerging studies have shown that the regions of the genome with lower meiotic recombination rates possess less phylogenetic conflict and have a tendency to preserve species histories. However, we have a poor understanding of how recombination rates have evolved over tens of millions of years and how we can use this information to better select informative genomic regions for phylogenomics. Here, we performed a comparative genomic analysis of recombination maps from divergent placental mammals to identify genomic regions that have retained historically low and high recombination rates. Low-rate regions should be useful for resolving species relationships, while high-rate regions may be useful for adaptive gene studies. To track the evolution of recombination rates, we reconstructed the ancestral placental mammal karyotype from mammalian genomes previously identified as having slowly evolving karyotypes and representing the four superordinal clades. We combined the ancestral chromosomes and recombination maps from extant species to identify syntenic regions where high and low recombination rates in the ancestor have persisted for 90-100 million years of independent evolution and characterized phylogenetic signal and functional attributes from these regions. These findings will enhance our understanding of the chromosomal architecture of speciation and natural selection within mammals.

459 - Hill-Robertson interference effects bias the inference of fitness effects of new mutations in partially-selfing populations

Parul Johri

Austin Daigle, Parul Johri
Department of Biology, Department of Genetics, University of North Carolina, Department of Biology, Department of Genetics, University of North Carolina

The accurate estimation of the distribution of fitness effects (DFE) of new mutations is critical for population genetic inference but remains a challenging task. While various methods have been developed for DFE inference using the site frequency spectrum of putatively neutral and selected sites, their applicability in species with diverse life history traits is not well understood. Selfing is common among eukaryotic species and can lead to decreased effective recombination rates, increasing the effects of selection at linked sites, including interference between selected alleles. We employ forward simulations to investigate the limitations of current DFE estimation approaches in the presence of selfing and linked effects of selection. We find that distortions of the site frequency spectrum due to Hill-Robertson interference in highly selfing populations lead to mis-inference of the deleterious DFE of new mutations. While accounting for the decrease in the effective population size due to linked effects of selection largely accounts for the observed bias in populations with moderate levels of selfing, this correction is unable to accurately estimate the DFE in highly selfing populations. In addition, the presence of cryptic population structure and uneven sampling across subpopulations leads to the false inference of a deleterious DFE skewed towards effectively neutral/mildly deleterious mutations. Finally, the proportion of adaptive substitutions estimated at high rates of selfing is substantially overestimated. Our observations apply broadly to species and genomic regions with little/no recombination.

460 - Nucleotide bias between leading and lagging strand in bacteria is caused by asymmetric mutagenesis

Eldar Badamshin

Eldar Badamshin, Alexey Skudnov, Stepan Denisov, Konstantin Gunbin , Konstantin Popadin
Higher School of Living Systems, Immanuel Kant Baltic Federal University , Higher School of Living Systems, Immanuel Kant Baltic Federal

University , The University of Manchester, Higher School of Living Systems, Immanuel Kant Baltic Federal University , Higher School of Living Systems, Immanuel Kant Baltic Federal University

The Mutation Accumulation Experiment (MAE) is a commonly employed technique to obtain the most pure mutational spectrum (Anjali Mahicar et al, 2022). These studies usually provide a 6-component mutational spectrum (2 base pair types*3 possible mutations) for different bacterial species.
According to the parity rule 1 (hypothesis) frequencies of complementary mutations have to be equal (Sueoka, 1995). For this reason the 6-component mutation spectrum is commonly used in research. The approach doesn't differentiate between complementary mutations (e.g. C>A and G>T), thus reducing 12 possible mutations to 6.
Using MAE data we created 12-component mutation spectrum of four species: Vibrio cholerae, Aliivibrio fischeri, Bacillus subtilis and Dienococcus radiodurans (Dillon et al., 2016, Sung et al., 2015, Long et al., 2015), taking into account replication asymmetry and that many mutations may arise due replication errors (Horton et al., 2023).
We showed that frequencies of complementary mutations are not equal on leading and lagging strands. We discovered that the GT-rich leading strand is characterized by C>T and A>G mutations that outnumber the corresponding complementary mutations. This observation is true for all chromosomes (main and secondary) for C>T and A>G mutations in MutS/L-deficient lines, and for C>T in wild-type of all studied species. We assume that C>T, A>G mutations occur more frequently in the template ssDNA for lagging strand (Frederico et al., 1990), which spends more time being single-stranded compared to the leading strand template.

467 - Localization of light-sensitive proteins, melanopsin/opn4, in neuroepithelial cells and radial glial cells in sablefish (Anoplopoma. fimbria) brain during development

Niloufar Mokariasl

Niloufar Mokariasl, Hayley V Barnes, Brent E Gowen, Gursimran Gill, John S Taylor
University of Victoria, University of Victoria, University of Victoria, University of Victoria, University of Victoria

Opsin proteins, categorized as G protein-coupled receptors, facilitate light sensitivity in vertebrates. They are typically divided into two groups, visual opsins, which contribute to vision or image formation, and non-visual opsins.
Non-visual opsins greatly outnumber visual opsins, yet our understanding of their functions remains limited. Most non-visual opsins are well conserved across vertebrates, especially fish. Consequently, animals that are very distantly related to one another, and living in completely different light environments, have almost identical non-visual opsins repertoires. Thus, their roles, whatever they may be, are also expected to be highly conserved. We are investigating melanopsin gene expression in sablefish. Sablefish is a deep-sea fish that lives in the Pacific Ocean and experiences varying light conditions throughout its lifecycle. We introduced sablefish (A. fimbria) as a model for studying opsins because of its unique life history and availability, from unfertilized egg to adulthood, from a local aquaculture facility. Using qPCR we quantified the expression of five melanopsin genes in 22 stages in sablefish. Furthermore, using immunohistochemistry, we have localized subset of these proteins in neuroepithelial cells and radial glial cells within the brain, as well as in all layers of the retina at four different larval stages and in juveniles, which indicates that these opsins might play a role in brain development. Moreover, we developed a brain atlas, employing hematoxylin and eosin (H&E) staining, and compared eye and brain development in sablefish, zebrafish, and medaka.

480 - A proposal of the Ur-RNAome

Miryam Palacios-Pérez

Miryam Palacios-PÃ©rez, Marco V. JosÃ©

Universidad Nacional AutÃ³noma de MÃ©xico, UNAM, Network of Researchers on the Chemical Emergence of Life, NoRCEL, Universidad Nacional AutÃ³noma de MÃ©xico

We portray here a plausible collection of ancient RNAs following an RNY pattern (where R indicates purine, N means any of the four bases, and Y indicates pyrimidine), since Eigen and Schuster suggested in the early 1970s, based on logical deduction of the thermodynamic properties of such polyribonucleotides in the RNA world, that this was the first genetic code.

We selected 26 species (13 Bacteria and 13 Archaea), and we began by extracting an RNY RNAome from a contemporary organism and the resulting smaller RNAome was used as a query in order to obtain a list of ancient RNAs encoded by RNY codons. We found that the free energy of RNY hairpins was consistently lower than their corresponding shuffled controls.

We found traces of the 3 ribosomal RNA (16S, 23S and 5S), tRNAs, 6S RNA, and the RNA moiety of RNAseP and of Signal Recognition Particle (SRP). Yet, at his stage of evolution there was not a genetic code as there was the complete absence of Peptidyl Transferase Centre, and not even vestiges of anti-Shine-Dalgarno 16S RNA. Interestingly, we detected the anticodons of both glycine (GCC) and threonine (GGU) in their hairpins proto-tRNA.

It is widely accepted that the earliest RNA molecules folded into hairpins or mini-helixes. Herein, we depict the 2D and 3D conformation of those earliest RNA molecules with only RNY triplets (Ur-RNAome).

499 - Revealing functional consequences of eukaryotic histone H2A repertoire innovations

Pravrutha Raman

Pravrutha Raman, Hana Khan, Kelsey Woodruff, Amy Hamada, Toshio Tsukiyama, Harmit Malik
Fred Hutchinson Cancer Center, Fred Hutchinson Cancer Center, University of Washington, Fred Hutchinson Cancer Center, University of Washington, Fred Hutchinson Cancer Center, University of Washington, Fred Hutchinson Cancer Center, Fred Hutchinson Cancer Center, Howard Hughes Medical Institute

Despite their essential roles in fundamental biological processes eukaryotic histones display repeated genetic innovations that could confer lineage-specific adaptive advantages. We investigate the fusion of two conserved histone H2A variants. Most eukaryotic H2A repertoires include canonical H2A and two distinct variants H2A.Z and H2A.X, important for gene regulation and DNA damage response (DDR), respectively. However, Drosophila encode a single H2Av gene, a functional chimera of H2A.X and H2A.Z. Our evolutionary analyses show that H2Av arose in a subset of Diptera by H2A.Z acquiring the H2A.X C-terminal tail containing â€œSQ motifâ€• required for DDR. This confines the DDR and gene regulation functions of H2A.X and H2A.Z to the same genomic location. We hypothesize that this fusion may be advantageous to limit DDR- dependent transposition to H2A.Z-enriched heterochromatin. To investigate, we are reversing D. melanogaster to its ancestral repertoire by engineering distinct H2A.X and H2A.Z, uncoupling their functions to reveal how genomic functions are altered by their fusion. We also recreated this evolutionary event in yeast by appending its H2A.X C-terminal tail onto its H2A.Z. These â€˜fly-likeâ€™ yeast have a dramatically reduced DDR efficiency that cannot be attributed entirely to a lack of SQ phosphorylation, since phosphomimic mutations fail to enhance DDR. These data insinuate that an H2A+H2Av repertoire might lead to a trade-off of compromised DDR in fly genomes and to the co-evolution of other fly chromatin factors to facilitate DDR. Our work reveals the functional consequences of histone innovation on genome instability.

511 - Pseudogenization-driven gene loss shapes genome evolution in Hanseniaspora

Luis Fernando García-Ortega

Luis F. GarcÃ-a-Ortega, Angela MarÃ-a GarcÃ-a-Acero, Alexander DeLuna, LucÃ-a Morales, Eugenio Mancera, Luis Delaye
Departamento de IngenierÃ-a GenÃ©tica, Unidad Irapuato, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Departamento de IngenierÃ-a QuÃ-mica y Ambiental, Universidad Nacional de Colombia, Departamento de IngenierÃ-a GenÃ©tica, Unidad Irapuato, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Unidad de GenÃ³mica Avanzada (Langebio), Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano (LIIGH), Universidad Nacional AutÃ³noma de MÃ©xico, Departamento de IngenierÃ-a GenÃ©tica, Unidad Irapuato, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Departamento de IngenierÃ-a GenÃ©tica, Unidad Irapuato, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional

Gene family contractions and gene loss have played significant roles in the evolution of fungal genomes. Among the mechanisms driving gene loss, pseudogenization stands out as a major contributor; high mutation rates during this process lead to rapid genome reduction. Fungi from the Hanseniaspora genus are a great model to understand gene loss since some spices are characterized by the smallest genomes among the budding yeasts. Furthermore, Hanseniaspora can be separated into two lineages, a slow evolution lineage and a fast evolution lineage with reduced genomes. In this study, we isolated and sequenced 16 Hanseniaspora isolates from different geographic regions across Mexico. Among them, ten were identified as H. lachancei, two as H. pseudoguilliermondii, two as H. guilliermondii, and one as H. opuntiae. Comparative analysis of these genomes revealed the dynamic nature of gene family sizes and compositions, with metabolic, DNA repair, and cell cycle control functions being the most affected. Interestingly, the distribution of pseudogenes does not correlate with genome size or gene count, but instead shows lineage specificity, indicating that distinct remodeling processes operate during the evolution of each organism's genome. Overall, we observed a wide range of pseudogene integrity, suggesting that each lineage may exhibit varying degrees of partial activity. This hypothesis will be further tested by phenotyping the isolates under a variety of growing conditions. Our findings highlight the importance of pseudogenization-driven gene loss as a mechanism that shapes genome evolution in autochthonous Hanseniaspora populations, providing insights into their adaptation strategies.

513 - A Structural Proteome Screen Identifies Protein Mimicry in Host-Microbe Systems

Gabriel Alejandro Penunuri

Gabriel A Penunuri, Shelbi Russell , Russ Corbett-Detig, Pingting Wang
UCSC, NA, UCSC, UCSC, UCSC

Host-microbe systems are unique evolutionary niches that produce complex biological interactions that neither participant could evolve on their own. Over the last two decades the influence of host-microbe interactions on global health has begun to emerge and with it the greater biological research community's interest in host-microbe interactions and the evolution of these systems. However, these systems have historically been a complex and difficult field of biological research. Impactful advances in global health will be obtained by progressing our understanding of these systems and laying the groundwork for future bioengineering projects. Protein mimicry, wherein a microbe encodes a protein that closely resembles and mimics the structure and function of a hostâ€™s, is an important aspect of host manipulation in these systems. We have utilized novel protein structure prediction and alignment tools in order to explore the structural proteome for examples of protein structure mimicry previously undetectable by sequence based approaches. By leveraging the Legionella pneumophila proteome and its many known structural mimics, we have developed and validated a framework that can be applied to virtually any host-microbe system to uncover signals of protein mimicry, identifying candidates of host control in microbial proteomes. We successfully identified undiscovered effector candidates as well as determined host targets of previously known effectors in three microbe proteomes important to human health and disease; Legionella pneumophila, Helicobacter pylori, and Wolbachia. We also further

supported the evidence of one of the Wolbachia mimic candidates through functional assays in the Drosophila-Wolbachia system.

515 - Characterising the detectable and invisible fractions of genomic loci under balancing selection

Max Reuter

DÃ©bora Brandt, Filip Ruzicka, Aida AndrÃ©s, Tim Connallon, Max Reuter
University College London, Institute of Science and Technology Austria, University College London, Monash University, University College London

Balancing selection refers to a suite of population genetic processes that push allele frequencies away from the boundaries and towards a polymorphic equilibrium state, thereby helping to maintain heritable variation in fitness and its components. Empirical studies have provided iconic examples of balanced polymorphisms, yet the general prevalence of balancing selection within genomes remains a topic of longstanding debate. Genome-wide scans for signals of balancing selection suggest that it is rare. However, current inference methods are notoriously conservative and are thought to identify only a subset of loci evolving under balancing selection. What is unclear is what proportion of loci under balancing selection is essentially invisible to detection and how our appreciation of the biological properties of balanced allelic variation (stability and equilibrium frequency, or selection and dominance coefficients) is biased by limitations in the power of inference. Here, we address these questions using a combination of analytical modelling and population genetic simulations. We begin by reviewing classic population genetics theory for the maintenance of polymorphism under three scenarios of balancing selection, overdominance, antagonistic pleiotropy and sexual antagonism. We then extend this theory to make predictions about the fraction and properties of loci under each balancing selection scenario that remain stably polymorphic for extended periods of timeâ€”and hence most visible to selection scans. Finally, we determine the portion of the parameter space in which current methods are able to detect balancing selection. We discuss the implications of our predictions for the interpretation of empirical data from genome scans.

521 - Unveiling reference genes for gene expression analyses in Heliconius erato lativitta

Paola Fernanda Calderón

Paola Fernanda CalderÃ³n, Angel Ojeda, Caroline Bacquet
Universidad Regional AmazÃ³nica Ikiam, Universidad Regional AmazÃ³nica Ikiam, Universidad Regional AmazÃ³nica Ikiam

The Neotropical butterfly Heliconius erato lativitta has been used as a model organism for the study of evolutionary adaptations, speciation, and genetic bases. This species is mainly recognized for its wing color pattern mimicry. However, the molecular mechanisms involved in adaptation to temperature changes remain unidentified. In this sense, quantitative real-time PCR (qPCR) has resulted one of the most useful tools to elucidate the molecular mechanism through the quantification of the mRNA of genes involved in the adaptation of this species. Nevertheless, the accuracy and reliability of qPCR data depends on the proper selection of reference genes, whose expression profiles must remain stable across different experimental conditions. Therefore, in this work six genes (ACT1, ACT2, ANX, AK, eEf1Î±, and eEf1Î±2) from H. e. lativitta were used as candidate reference genes and their expression level was assessed at different temperatures in three type of tissues both for males and females. The stability of reference genes was evaluated using three algorithms geNorm, NormFinder and RefFinder. The results obtained with these algorithms indicate that AK and ACT2 are stables under temperature stress. However, at the tissue level, only ACT1 and AK were the most stable genes, while at the sex level, only ACT1 and ACT2 exhibited a high stability. Hence, this study provides a standardized set of reference genes for future gene expression analysis using qPCR in H. e. lativitta

523 - Illuminating the Abyss: Cultivation Strategies for Uncharted Cave Dwarf Bacteria

Karen Sevilla Landaverde

Karen Sevilla Landaverde, Corina Diana Ceapa, Arturo JimÃ©nez Sanchez, Colin Bonduelle, Daniel RÃ-os Barrera
Institute of Chemistry, Institute of Chemistry, Institute of Chemistry, University of Bordeaux, Biomedical Research Institute

Microbial Dark Matter is a term used to describe the vast majority of microorganisms on Earth that are yet-to-be-cultured, representing over 99% of the estimated microbial diversity. While these microorganisms are abundant and diverse, particularly in soils, there is a significant proportion of them that remain unexplored, especially in unique environments such as caves. Among this unexplored microbial diversity, dwarf bacteria, which have cell sizes smaller than 0.6Âµm, are particularly challenging to grow due to their nutrient-hoarding ability and limited microenvironment access.
To address this challenge, cultivation methods have been proposed to simulate the conditions found in caves, which have yielded two populations of cultivated bacteria: larger and facultative small-cell size bacteria, and dwarf bacteria. While the former have been successfully isolated, the cultivation rate and morphotype diversity for the latter appear to be reduced.
To accurately estimate the cell size of the bacteria, four new fluorescent molecules have been tested and compared against DAPI in this study, resulting in improved size estimation. Using this methodology, several cell populations of dwarf bacteria that grow in caves were isolated and characterized.
This research highlights the significance of exploring unexplored environments to expand our understanding of microbial diversity, which could lead to the discovery of novel microbial species. The proposed cultivation method could also be useful for studying other environments, such as extreme habitats and even well-known ones. The findings of this study emphasize the importance of developing novel cultivation methods to study the vast majority of unexplored microbes.

534 - Genetic regulatory effects in response to a high-cholesterol, high-fat diet in baboons

Wenhe Lin

Wenhe Lin, Jeff Wall, John L VandeBerg5, Michael Olivier, Laura A Cox, Yoav Gilad
University of Chicago, UCSF, University of Texas Rio Grand Valley, Wake Forest University School of Medicine, Wake Forest University School of Medicine, University of Chicago

Steady-state eQTLs account for only a fraction of disease-associated loci identified through GWAS, while eQTLs involved in gene-by-environment (GxE) interactions have rarely been characterized due to experimental challenges. Here, we leverage baboonsâ€™ genetic and physiological resemblance to humans and explore gene-by-diet effects on gene regulation in a longitudinal study.

We collected liver, muscle, and adipose samples from 99 baboons before and after they were fed a high-cholesterol, high-fat (HCHF) diet for two years. We obtained transcriptomic profiles from each tissue at the two time points. We identified 2,714 diet-responsive eQTLs (DR eQTLs) with a variable effect after HCHF. We found that DR eQTLs exhibit a localization pattern that is distinct from non-DR eQTLs, featuring a more evenly distributed distance to the TSS. Similar to GWAS hits, DR eQTLs are enriched among genes with complex regulatory landscapes, enduring stronger selective constraints, and participating in diverse biological functions. Furthermore, the human orthologs associated with DR eQTLs are enriched for GWAS genes associated with human metabolic traits, suggesting that context-responsive eQTLs with more complex regulatory effects are likely to explain GWAS hits that do not seem to overlap with standard eQTLs.

Our results highlight the complexity of genetic regulatory effects and the distinct characteristics of dynamic eQTLs that are responsive to

environmental perturbations. We provide additional support for the notion that eQTLs with disease-relevant GxE interactions are crucial for a better understanding of non-coding GWAS signals for human complex disease.

544 - Chloroplast Phylogenomics and the Taxonomy of Mesoamerican Agaves

Elsa Peters Ruiz de Chávez

Elsa Peters Ruiz de ChÃ¡vez, AbisaÃ- GarcÃ-a-Mendoza, Alejandro Casas, Daniel PiÃ±ero, Logan Kistler, Marcela Sandoval-Velasco
Posgrado en Ciencias BiolÃ³gicas, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, JardÃ-n BotÃ¡nico, Instituto de BiologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de Investigaciones en Ecosistemas, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de EcologÃ-a, Universidad Nacional AutÃ³noma de MÃ©xico , Anthropology Department, National Museum of Natural History, Smithsonian Institution, Centro de Ciencias GenÃ³micas, Universidad Nacional AutÃ³noma de MÃ©xico, Anthropology Department, National Museum of Natural History, Smithsonian Institution

Agaves are one of the most emblematic plant groups in Mexico, where they thrive and dominate various ecosystems. Their importance resides not only in the great diversity and broad ecological interactions these plants have but also in the multiple uses humans have given to them for thousands of years.

Consequently, Agaves have been in the spotlight of intensive research for many years which has led to discussions about their taxonomy and attempts to reconstruct the molecular phylogeny of the genus. However, most studies have been based on few and small chloroplast and nuclear DNA sequences and important phylogenetic relationships have been left undetermined due to a lack of data resolution, which limits testing some relevant evolutionary hypotheses.

Since taxonomy and molecular systematics provide frameworks for understanding biological interactions within a genus, both should be considered when trying to understand evolutionary relationships. However, for Agave these approaches have been insufficiently integrated.

In this study, we generated chloroplast data for 160 agave samples from both herbarium and living specimens representing 65 species. We de novo-assembled and circularized chloroplast genomes which we use to reconstruct a robust phylogeny. We use this phylogeny as a biological framework for testing evolutionary hypotheses about niche conservatism and niche diversification for species historically managed by humans in Mesoamerica. Such integration looks to understanding how human management has impacted the geographic distribution and molecular phylogenetic relationships among Agave species, which in turn has modified the ecological niche and evolutionary history of the genus.

548 - The extent of incomplete lineage sorting in divergences of major groups of Neoaves

Naoko Takezaki

Naoko Takezaki
Kagawa University

Owing to large-scale molecular studies, it has become clear that extant birds are separated into three clades: Palaeognathes, Galloanserae and Neoaves. However, the higher-level phylogenetic relationships of Neoaves are difficult to resolve likely due to rapid diversification. Previous studies recognized ten major groups in Neoaves, â€œthe magnificent sevenâ€• and three orphan groups. Their relationships have been differently estimated by different studies and different types of sequence data such as coding and non-coding.

This study investigated the extent of incomplete lineage sorting using different types of data [coding sequences, introns, ultra-conserved elements (UCEs), and untranslated regions (UTRs)] from three previous studies using gene concordance factor (proportion that the relationship or the grouping appeared in gene trees) and site concordance factor (proportion of sites that supported them by parsimony-based criterion). . Most of the relationships of the major groups appeared in estimated phylogenies with concatenated sequences and even clustering of the major groups were supported by less than few percent of gene concordance factors and 30% - 40% of site concordance factors, many of which were close to 33.3%. The values of both concordance factors were similar for all the different types of sequence data. This result indicates that in the relationships of many of the major groups the extent of incomplete lineage sorting is extreme and that their divergences are close to multifurcation.

558 - Exploring independent evolution of light-dependent redox regulation in the plastids of Paulinella

Duckhyun Lhee

Duckhyun Lhee, Hwan Su Yoon
SungKyunKwan University, SungKyunKwan University

Redox regulation plays an important role in the control of metabolic pathways in response to environmental changes. Within chloroplasts, thiol oxidation is linked to the activation of photosynthetic electron transport, allowing the modulation of chloroplast functions in response to light stimuli. In addition, evolutionary trends suggest an expansion in the number of redox-sensitive cysteines within the proteome, particularly coinciding with the emergence of plastids. In this study, we investigate the redox regulatory system in the plastid of Paulinella micropora, an organism possessing photosynthetic organelles derived from an independent primary endosymbiosis of Archaeplastida plastids. Using a quantitative proteomic approach employing resin-assisted enrichment coupled with isobaric tandem mass tag labeling, we investigate the redox dynamics of cysteine sites under light and dark conditions. Our results suggest that in Paulinella, the light-dependent redox regulation of key metabolic pathways within plastids has evolved independently from that observed in plants. However, the mechanism underlying the transfer of light-induced electrons to target proteins remains unclear. Furthermore, our analysis reveals a high percentage of oxidation among cysteines located in the plastid transit peptide regions of Paulinella, suggesting a possible role for in these cysteines.

569 - Regulatory evolution and gene expression under hypoxia

Shagufta Qamar

Shagufta Qamar
OIST

Oxygen is hypothesized as one of the critically defining elements that led to the diversification and evolution of complex life forms. However, the exact mechanisms of how early animals evolved under low atmospheric oxygen levels remains poorly understood.
The starlet sea anemone, Nematostella vectensis, belongs to Cnidaria, an early-branching (basal) metazoan phyla that likely emerged around the same time as the rise of atmospheric oxygen levels, yet retains much of the genetic complexity of their common ancestor, as is used as the model in our studies.
Using a comparative analysis approach, the transcriptomes of the N. vectensis embryos cultured under hypoxia are analysed to i) identify key regulatory pathways involved in the hypoxia response in a basal metazoan, and ii) permit functional interrogation of the ancestral molecular toolkit employed in embryonic development under hypoxic conditions.
Our work on a basal metazoan group will provide a more complete understanding of the evolutionary processes that guided the emergence of early metazoans.

581 - Hsp90 rescues the function of acetyl-CoA synthetase carrying mutations in a highly conserved residue

Chun-Hsin Yang

Chun-Hsin Yang, Jun-Yi Leu

Institute of Molecular Biology, Academia Sinica, Institute of Molecular Biology, Academia Sinica

Maintaining protein homeostasis is vital for cellular survival. Heat shock protein 90 (Hsp90), a highly conserved molecular chaperone, facilitates the maturation of a wide variety of protein clients, and therefore engages in numerous cellular process. By maintaining protein stability and/or folding, Hsp90 can buffer the detrimental effects of genetic mutations. Cryptic genetic variations were released under stress conditions of Hsp90 activity insufficiency. Yet, our knowledge in the molecular mechanism underlying Hsp90 buffering is limited.
In this study, we aimed to seek mutations which can be buffered by Hsp90. Through EMS-mutagenesis, we collected clones with growth defects under Hsp90 inhibition condition. To further investigate the causal mutations, we conducted bulk-segregant analysis coupling with whole-genome sequencing. We identified a missense mutation on ACS2, which is an essential gene encoding acetyl-coA synthetase, as one of the Hsp90 buffering targets. Remarkably, this amino acid exhibits a high degree of conservation across taxa ranging from fungi to vertebrates, indicating its evolutionary significance. Via error-prone PCR library screening, we found another two putative single mutations on ACS2 buffered by Hsp90. We hypothesized that the Hsp90 buffering system allows cells to accumulate additional mutations to compensate the deleterious effect of the initial one during evolution. In the future, we would like to further dissect the molecular mechanism of Hsp90 buffering effect and its impact on evolution.

583 - Environmentally driven changes in hunter-gatherer population interconnectivity explain African genetic diversity

Cecilia Padilla-Iglesias

Cecilia Padilla-Iglesias, Michela Leonardi, Zhe Xue, Johanna Paijmans, Margherita Colucci, Anahit Hovhannisyan, Lucio Vinicius, Andrea B Migliano, Andrea Manica

Human Evolutionary Ecology Group, University of ZÃ¼rich, Evolutionary Ecology Group, Department of Zoology, University of Cambridge, Evolutionary Ecology Group, Department of Zoology, University of Cambridge, Evolutionary Ecology Group, Department of Zoology, University of Cambridge, Evolutionary Ecology Group, Department of Zoology, University of Cambridge, Human Paleosystems Group, Max Planck Institute for Geoanthropology, Evolutionary Ecology Group, Department of Zoology, University of Cambridge, Trinity College Dublin School of Genetics and Microbiology, Human Evolutionary Ecology Group, University of ZÃ¼rich, Human Evolutionary Ecology Group, University of ZÃ¼rich, Evolutionary Ecology Group, Department of Zoology, University of Cambridge

Emerging archaeological, anthropological and genetic evidence have challenged the traditional view that Homo sapiens originated from a single African region. Instead, it has been proposed that our species emerged and diversified within numerous, geographically distinct populations in Africa that became in contact at different points in time, exchanging genes and culture. Whilst analyses of contemporary and ancient genomes from African hunter-gatherer groups support deep divergence times between them, coupled with intermittent episodes of gene flow, these events are lacking a geographic and, perhaps more importantly, the climatic context in which they took place. Our study employs a Climate Informed Spatial Genetic Modelling (CISGeM) framework, integrating Species Distribution Models (SDMs) built from African hunter-gatherer archaeological assemblages and paleoclimatic reconstructions with data from all available contemporary and ancient African hunter-gatherer genomes. This allows formally assessing whether inferred geographic range changes from SDMs, inter-regional migrations and demographic fluctuations can be reconciled with genetics. We utilise Approximate Bayesian Computation to determine which demographic parameters (migration rates, admixture dates, growth ratesâ€¦) are most supported. Our

findings suggest that temporal patterns of African genetic diversity are attributable to climatic influences on population dynamics, and do not require the necessity for hypothesized introgression from â€œghostâ€• populations. We pinpoint key environmental factors promoting and hindering gene-flow between African regions and outline the historical geographical range of ancient hunter-gatherer populations. This research highlights the value of incorporating genomic data into spatially explicit models to elucidate the complex processes shaping human evolutionary trajectories.

584 - Population genomic analysis reveals cryptic population structure in the commercially important Lake Malawi cichlid Copadichromis mloto

Sawasawa

Wilson Sawasawa

University of Antwerp, University of Antwerp

Fish is an important source of animal protein for many people in Malawi and the regions of Mozambique and Tanzania bordering the lake. We investigated population genetic structure in the Malawi cichlid species C. mloto based on whole-genome data of a relatively large and geographically widespread sample set. We found that what was previously thought to be a single species of utaka â€“ in itself a complex of semi-open water, planktivorous cichlid species â€“ shows unexpected patterns of population structure, pointing to the presence of at least three widely distributed, genetically distinct C. mloto clades. Generalising from this, our results indicate that taxonomic diversity in the large and understudied group of benthic and pelagic Lake Malawi cichlid species might be higher and more complex than previously thought, despite generally wide distribution ranges and seemingly continuous habitats. Such diversity might often not be easily discernible by morphological examination and calls for broad genomic investigation and monitoring

601 - MYB transcription factors involved in Ophrys orchids labellum development

Sara Edith Garcia-Morales

Sara E Garcia-Morales, Mattia Alessandrini, Alessia Russo, Philipp SchlÃ¼ter

Department of Plant Evolutionary Biology, Institute of Biology, University of Hohenheim, Department of Plant Evolutionary Biology, Institute of Biology, University of Hohenheim, Department of Plant and Microbial Biology and ZÃ¼rich-Basel Plant Science Centre, University of Zurich, Department of Systematic and Evolutionary Botany and ZÃ¼rich-Basel Plant Science Centre, University of Zurich, Department of Plant Evolutionary Biology, Institute of Biology, University of Hohenheim

The sexually deceptive Ophrys orchids attract specific pollinator species, deceiving them via mimicry of floral traits such as odor, color, and shape, specifically at the labellum. Our study focuses on O. sphegodes, whose highly specific pollinator interaction has led to a mechanism that promotes reproductive isolation. Deciphering these traits' genetic regulation is necessary to better understand the molecular basis of pollinator attraction. MYB transcription factors (TFs) are known to play an essential role in molecular processes. In orchids, several MYB TFs have been studied in well-known biosynthetic pathways and show precedents that pigment accumulation displays spacio-temporal specificity. To study the molecular mechanisms behind labellum pigmentation, we first identified the O. sphegodes MYB TFs. Previously published genomes had come mainly from orchids distant from the Ophrys genus. Since the recent publication of the Plathantera and Ophrys genomes, where duplication events were described across their genomes, a characterization of the MYB TFs was needed to identify which clades underwent duplication events in Ophrys sphegodes. We then analyzed gene duplication, selection, and gene expression in clades associated with labellum development. This data will contribute to our understanding of the molecular processes that shape the complex Ophrys labellum.

603 - Rapid evolution of recombination landscapes during the divergence of cichlid ecomorphs in Lake Masoko

Marion Talbi

Marion Talbi, Milan Malinsky

Institute of Ecology and Evolution, University of Bern, Institute of Ecology and Evolution, University of Bern

Meiotic recombination is fundamental to evolution of sexually reproducing organisms and differences in recombination rates are important during rapid adaptation and organismal diversification. However, despite a substantial body of theory, many unknowns remain regarding how and why recombination rates evolve in nature. Here, we reconstruct recombination maps based on linkage disequilibrium and use subsampling and simulations to show that fine-scale recombination landscapes differ substantially between two cichlid fish ecotypes of Astatotilapia calliptera that diverged only ~2,500 generations ago. We show that these recombination differences are partially associated with genomic islands showing high allele frequency divergence. We also found 47 large haplotype blocks that are polymorphic in Lake Masoko, cover 21% of the genome, and contribute disproportionately to the evolution of recombination. The observed results are not driven by PDRM9 whose binding sites do not show any relationship to recombination rates in this species. Overall, our work provides a holistic view of changes in recombination landscapes during early stages of speciation with gene flow and advance our understanding of the combinatorial basis of evolution.

605 - The activity of transcription factor is easily evolvable

Richard Ron Resurreccion Litan

Richard Ron R Litan, Po Chen Hsu, Jun Yi Leu

Institute of Molecular Biology, Academia Sinica, Institute of Molecular Biology, Academia Sinica, Institute of Molecular Biology, Academia Sinica, nstitute of Molecular Biology, Academia Sinica, nstitute of Molecular Biology, Academia Sinica

The initiation of transcriptional rewiring is influenced by changes in both the cis- and trans-regulatory elements. However, the process on how perturbations in trans-acting elements instigate transcriptional rewiring remains elusive. Previous report showed that Sef1 in Lachancea kluyveri (LkSef1) facilitates non-functional binding to its major target genes. On the other hand, fusing LkSef1 with the transcriptional activator VP16 not only increased its basal activation but also conferred control to other genes originally unaffected by it. Thus, changes such as in the form of mutations that can improve the activity of a transcription factor may mediate new target genes and consequently elicit rewiring. Here, we combined random mutagenesis in LkSef1 and iterative fluorescence-based sorting to select for mutations that can readily enhance its transcriptional activity. Using an LkSef1-enhanced GFP reporter system, gain-of-function mutants were selected, and point mutations were found to increase its transcriptional activity. Also, all mutations are concentrated outside the conserved domains suggesting undefined regions responsible for its activity. Once reconstituted in the genome, these high activity LkSef1 can potentially create novel networks and may ultimately provide mechanisms for transcriptional rewiring.

613 - Effects of pesticides are driven by differences between species, not their chemical structure

Alicja Witwicka

Alicja Witwicka, Yannick Wurm, Federico Lopez-Osorio, Courtney May, Yeahji Jeong

Queen Mary University of London , UCL, Queen Mary University of London , Alan Turing Institute, Queen Mary University of London , Queen Mary University of London , Queen Mary University of London

The extensive use of pesticides unintentionally harms insect pollinators, but how their effects differ among species is not well

understood. Pesticide toxicity is typically assessed using model species. Results are then extrapolated to thousands of other insects, underscoring the need for immediate investigation into species-specific effects. Comparing behaviour across species is challenging due to significant morphological and life history variations. However, high-resolution genetic tools provide more reliable and consistent results. By examining brain gene expression data, we explored the impact of pesticides on four insect species from three different orders: a butterfly, a blowfly, a solitary bee, and a bumble bee. Contrary to expectations, sulfoxaflor, considered safe for bees, elicited stronger responses in non-bee species than clothianidin, a neonicotinoid known for its high toxicity. Intriguingly, while each species showed overlapping sets of genes and metabolic pathways affected by different pesticides, the overlap between species was minimal. We show that pesticide response is driven by evolutionary differences between species rather than the chemical makeup of pesticides. Our results reveal that it is incorrect to extrapolate the effects of pesticide exposure from one species to others and call into question the current pesticide classification system based on the chemical structure rather than biological targets or effects. Highlighting species-specific reactions to human-made stressors, our research underscores the importance of incorporating molecular data to enhance pesticide safety evaluations and inform broader conservation strategies.

617 - Experimental Evolution Reveals the Roles for Protein Translation and Gene Silencing in Regulating Protein Expression Noise

Wei-Han Lin

Wei-Han Lin, Florica J. G. Opoc, Chia-Wei Liao, Jun-Yi Leu

Institute of Molecular Biology, Academia Sinica, Doctoral Program in Microbial Genomics, National Chung Hsing University and Academia Sinica, Institute of Molecular Biology, Academia Sinica, Institute of Molecular Biology, Academia Sinica, Institute of Molecular Biology, Academia Sinica, Institute of Molecular Biology, Academia Sinica, Doctoral Program in Microbial Genomics, National Chung Hsing University and Academia Sinica

Non-genetic variations derived from expression noise at transcript or protein levels can result in cell-to-cell heterogeneity within an isogenic population. Although cells have developed strategies to reduce noise in some cellular functions, this heterogeneity can also facilitate varying levels of regulation and provide evolutionary benefits in specific environments. Despite several general characteristics of cellular noise having been revealed, the detailed molecular pathways underlying noise regulation remain elusive. Here, we established a dual-fluorescent reporter system in Saccharomyces cerevisiae and performed experimental evolution to search for mutations that increase expression noise. By analyzing evolved cells using bulk segregant analysis, whole-genome sequencing and mutation impact prediction, we identified two histone deacetylases, Hos2 and Sir2, as well as the catalytic subunit of the SWI/SNF chromatin remodeling complex, Snf2, as negative noise regulators. A hos2 mutant down-regulated multiple ribosomal protein genes and exhibited partially compromised protein translation, indicating that Hos2 may regulate protein expression noise by modulating the translation machinery. In addition, treating cells with translation inhibitors or introducing mutations into several Hos2-regulated ribosomal protein genes enhanced protein expression noise. Furthermore, deleting subunits of the Sir2 silencing complex led to noise increase, implying a close relationship between gene silencing and noise regulation. Taken together, our current study provides an effective strategy for identifying noise regulators and also sheds light on how cells regulate non-genetic variation.

630 - Establishment of newly-formed allopolyploids through transition to selfing

Uliana Kolesnikova

Uliana Kolesnikova, Alison Scott, Polina Novikova

MPI for Plant Breeding Research , MPI for Plant Breeding Research , MPI for Plant Breeding Research

Whole-genome duplication is a frequent event across the tree of life, and is especially common in plants. Despite this, newly formed polyploids may struggle to become established. The dramatic karyotypic change of a whole genome duplication comes with reproductive challenges, both due to improper segregation of chromosomes in meiosis and a lack of compatible mating partners. While autopolyploids are compatible with unreduced gametes produced in low amounts by diploid progenitors, allopolyploids (polyploid hybrids) are typically unable to cross with parental diploid species.

In Brassicaceae, allopolyploid establishment seems to be facilitated by a transition to self-compatibility. We explain this process using allotetraploid Arabidopsis kamchatica, which inherited the ability to self-pollinate from the diploid progenitor species A. lyrata. Specifically, the self-compatible parental A. lyrata lineage arose in Siberia, and lost self-incompatibility due to degradation in the male component of the self-recognition system. We have shown that this mutation is dominant. Having inherited the dominant non-functional self-incompatibility system variant from Siberian A. lyrata, A. kamchatica transitioned to self-compatibility immediately upon allotetraploid formation.

This mechanism of immediate transition to selfing in an allopolyploid due to the self-compatibility of a diploid ancestor is observed in other Brassicaceae (Arabidopsis suecica, Capsella bursa-pastoris, and Brassica napus). We propose that transition to selfing is a necessary step in the establishment of allopolyploids in Brassicaceae. In addition, our findings imply that loss of self-incompatibility in diploid species can impact the evolution of allopolyploids.

642 - Sexually antagonistic variation and balancing selection in the Drosophila melanogaster genome

Débora Brandt

DÃ©bora Brandt, Aida M. AndrÃ©s, Max Reuter

University College London, University College London, University College London

Selection for sexual dimorphism can give rise to an adaptive conflict between the sexes, when male and female traits are genetically correlated. This can lead to the accumulation of sexually antagonistic variants that are advantageous to one sex and deleterious to the other, and to their maintenance in the population through balancing selection. However, sexually antagonistic (SA) selection does not always result in long-term maintenance of polymorphism, and balancing selection can be caused by a variety of additional mechanisms (such as heterozygote advantage, or negative frequency dependent selection). In Drosophila melanogaster, evidence of sexually antagonistic (SA) polymorphism exists at genome-wide loci, posing the question of what is the overlap between SA loci and loci with signatures of balancing selection in nature. We are interested in the overlap of these processes from two perspectives: what proportion of loci under balancing selection can be attributed to SA selection? And what proportion of SA loci show signatures of long-term balancing selection? We identify the targets of balancing selection in three populations of D. melanogaster using, for the first time in this species, the most recent and powerful methods available. We assess the power of these methods in D. melanogaster using simulations, and identify loci showing signatures of both balancing and SA selection, assess their relative prevalence in the D. melanogaster genome, and discuss their biological function.

646 - Exploring genomic correlates of multi-trait phenotypic convergence in trap-jaw ants

Gaurav Agavekar

Gaurav Agavekar, Yoshihiko Suzuki, Michael Hiller, Gene Myers, Evan P Economo

Okinawa Institute of Science and Technology, Okinawa Institute of Science and Technology, LOEWE Centre for Translational Biodiversity

Genomics, Senckenberg Research Institute, Okinawa Institute of Science and Technology, Okinawa Institute of Science and Technology

Convergent evolution has produced remarkable adaptations throughout the tree of life. A central question is whether convergence at the genomic level underlies phenotypic convergence, and if so, what are the relative contributions of protein-coding vs. regulatory convergent changes? We ask this question in Strumigenys, a hyperdiverse ant genus in which trap-jaw mandibles (TJ), a phenotype characterized by morphological, behavioral, and neural adaptations, has evolved independently in different biogeographic regions. These ants employ TJ as a high-velocity weapon to hunt fast prey items such as springtails, and produce the fastest acceleration observed in resettable animal movements with their mandibles. We took a comparative genomic approach to analyze three independent instances of trap-jaw evolution in Strumigenys. To this end, we newly sequenced 18 chromosome-scale, highly accurate genomes (~99.9999% base-level accuracy) and generated comprehensive annotations. We then separately analyzed coding genes and conserved non-exonic elements (CNEs) to explore the genomic signatures of phenotypic convergence. In our analyses, we do not find strong evidence for convergence in protein-coding genes, but we do see multiple independent accelerations in CNEs in trap-jaw clades. These CNEs exhibit substantial overlap with ATAC-seq peaks, corroborating their functional significance in regulatory processes. Although our exploration of convergence within these CNEs is ongoing, preliminary findings suggest that regulatory regions, rather than protein-coding sequences, play an important role in multi-trait phenotypic convergence among Strumigenys ants.

656 - Primordial ferredoxin reduction by hydrogen with iron at the intersection between geochemistry and biochemistry

Natalia Mrnjavac

Natalia Mrnjavac, Max Brabender, Delfina P. Heinriques Pereira, Manon L. Schlikker, Martina Preiner, William F. Martin
Institute of Molecular Evolution, Heinrich Heine University, Institute of Molecular Evolution, Heinrich Heine University, Microcosm Earth Center, Max Planck Institute for Terrestrial Microbiology, Institute of Molecular Evolution, Heinrich Heine University, Microcosm Earth Center, Max Planck Institute for Terrestrial Microbiology, Institute of Molecular Evolution, Heinrich Heine University

The acetyl-CoA pathway of carbon fixation is considered to be the primordial metabolic route to carbon and energy in the first life forms. In modern cells, the direct electron donor for CO2 reduction via the acetyl-CoA pathway is ferredoxin, a small protein that is itself reduced by molecular hydrogen. This reaction is catalyzed by hydrogenases, which had to evolve a complex and fine-tuned process, flavin-based electron bifurcation, in order to overcome the thermodynamic barrier to ferredoxin reduction by hydrogen at physiological conditions. Ferredoxin is a small and simple electron carrier harboring one or two FeS clusters, and is likely to have evolved before flavin-based electron bifurcation and hydrogenases. We experimentally showed that hydrogen can reduce ferredoxin with Fe(0) non-enzymatically in a reactor, with conditions mimicking those in serpentinizing hydrothermal vents. This finding suggests that serpentinizing vent environments on the early Earth could replace the physiological function of flavin-based electron bifurcation, forging a link between geochemical processes in Earthâ€™s crust and nascent metabolism. Under the serpentinizing hydrothermal vents hypothesis for the origin of life, hydrogen-dependent ferredoxin reduction by iron would reflect a phase of biochemical evolution in which protometabolic reactions required direct physical contact of proteins with geochemical surfaces, before the origin of self-compartmentalized free-living cells.

657 - Information transfer from Earth’s surface to species’ evolution

Greer Dolby

Douglas G Moore
39 Alpha Research

Physiographic features on Earthâ€™s surface are long thought to limit and structure populations. That structuring occurs through the non-random reduction of gene flow associated with those features, resulting in partitioning of genomic variation across space. This can be contextualized as a form of information transfer between Earthâ€™s landscape and the lineages evolving with that landscape. As a proof of this concept, we simulated genomic data of diverging lineages, varying migration, mutation and recombination rates, and effective population size. We calculated standard pairwise genetic metrics genome-wide and then calculated the partial information decomposition of those genetic metrics to determine if any of the decomposed nodes held information that predicted the evolutionary history of those lineages, namely whether they diverged with or without gene flow. Considering the 7,648,800 nodes, the information content of nodes involving Tajimaâ€™s D and nucleotide diversity performed predicted gene flow history better than current SFS-based tools. This suggests that the context of lineage divergence (with or without gene flow) can be recovered with information theoretic measures calculated on the evolving sequences. If physiographic features do limit some inter-population gene flow in nature, then this suggests there is information transfer from Earthâ€™s landscape to populations of species which is encoded in the partitioning of genomic variation across space. Each species is a separate receiver of this communication, where species evolving with high fidelity to the landscape receive this information through a â€˜noiselessâ€™ channel, while others (e.g., panmictic species) receive this message through a noisy channel.

664 - Comparing common strategies for ortholog selection used in phylogenomics

Mingzhu Yang

Mingzhu Yang, Mattia Giacomelli, Maria Eleonora Rossi, Jesus Lozano Fernandez, Philip Donoghue, Davide Pisani
University of Bristol, University of Bristol, University of Barcelona, University of Bristol, University of Barcelona, University of Bristol, University of Bristol

Phylogenomics rely on superalignments including hundreds or thousands of orthologous gene families. Many pipelines exist to identify orthologs from sets of proteomes. A strategy that has grown in popularity is extracting single-copy genes pre-identified in the BUSCO database from the species of interest (Seppey et al., 2019). Another commonly used strategy is using OrthoFinder (Emms et al., 2019) to define de novo orthologous families. Other similar strategies can be designed using different approaches to define the orthogroups (e.g. using OrthoMCL). One frequently overlooked problem is the correspondence between single-gene families inferred using different methods. While the expectation is that if different pipelines perform comparably well, they should identify the similar sets of single-gene families, available evidence suggests the opposite (Li et al., 2023).
Here we identified 954 single-gene families using BUSCO for a set of nine high-quality metazoan proteomes and three outgroups. For the same genomes, OrthoFinder identified 2937 single-gene families, 365 of which were present in all 12 taxa. We functionally annotated these families and looked at the pattern of correspondence between families inferred using BUSCO and OrthoFinder. We found that the BUSCO metazoan set is depleted in translational genes (KOG functional category J). Furthermore, there is no clear correspondence between orthogroups inferred using BUSCO and OrthoFinder. We conclude that the use of alternative software for ortholog detection can potentially lead to the inference of different phylogenies, and it is thus crucial to better understand which of these approaches is more reliable.

666 - Population Size Shapes the Impact of Linked Selection in Primates

Bjarke Meyer Pedersen

Bjarke Meyer Pedersen, Juraj Bergman, Mikkel Heide Schierup
Bioinformatics Research Center, Aarhus University, Department of Biology - Ecoinformatics and Biodiversity, Bioinformatics Research

Center, Aarhus University

Selective sweeps and deleterious mutations affect nucleotide diversity at linked sites, termed linked selection. Linked selection shapes nucleotide diversity through background selection and sweeps; however, the quantitative effect is not well understood. Through genetic linkage, natural selection reduces surrounding nucleotide diversity by modulating nearby allele frequencies. The span of this reduction depends on recombination rate, creating an association between recombination rate and diversity. Using the resulting relationship, we investigate the quantitative effects of linked selection within genomes and across species and how the impact changes with effective population size.

We studied diversity patterns from recently published genome-wide diversity data from 233 primate species as a function of recombination rate. We lifted the data to hg38 coordinates and superimposed the human pedigree-based recombination map at a 100 kb scale where it is well conserved across primates. We then inferred the relationship between recombination rate and diversity in each species and modeled the effective population size for each species, to study how this relationship depends on an estimate of the overall Ne. We find a strong relationship between diversity and recombination rate in all species, highlighting the importance of linked selection shaping diversity in all primates. Additionally, the relative reduction in diversity in low/high recombination regions and the strength of the diversity recombination relationship increases with estimated population size, supporting the theory of stronger/more selection in larger populations.

669 - Human ZEB2 orchestrates a network of genes with stronger contribution to neuronal development than non-human primate ZEB2

Jeong-Eun Lee

Jeong-Eun Lee, Vladimir M Jovanovic, Amanda Jager Fonseca, Sebastian Streblow, Stefano Berto, Katja Nowick
Human Biology and Primate Evolution Group; Institute for Zoology; Department of Biology, Chemistry and Pharmacy; Freie UniversitÃ¤t Berlin, Human Biology and Primate Evolution Group; Institute for Zoology; Department of Biology, Chemistry and Pharmacy; Freie UniversitÃ¤t Berlin, Bioinformatics Solution Center; Freie UniversitÃ¤t Berlin, Human Biology and Primate Evolution Group; Institute for Zoology; Department of Biology, Chemistry and Pharmacy; Freie UniversitÃ¤t Berlin, Human Biology and Primate Evolution Group; Institute for Zoology; Department of Biology, Chemistry and Pharmacy; Freie UniversitÃ¤t Berlin, Medical University of South Carolina, Human Biology and Primate Evolution Group; Institute for Zoology; Department of Biology, Chemistry and Pharmacy; Freie UniversitÃ¤t Berlin

Humans differ from other primates in many phenotypes, despite their nearly identical protein sequences. Therefore, investigating human-specific changes at the level of transcriptional regulation becomes imperative for understanding species-specific traits and adaptations. In this study, we explore the functional evolution of ZEB2, a key transcription factor involved in neural development and other functions, across primates. Using ChIP-seq in lymphoblastoid cell lines (LCLs) from three great ape species, we demonstrate ZEB2â€™s conserved binding near transcriptional start sites but also species-specific variation in target site choice and binding affinity. Human-specifically altered genes followed by ZEB2 knockdown in LCL are involved in neural development, as well as synapse and neuron projection organization, pointing to a human-specific functional change related to brain developmental processes. Consequently, we analyzed single cell RNA-seq and bulk RNA-seq data of developing great ape brain organoids. We revealed species differences in ZEB2 co-expressed genes, which might underlie evolutionary alterations in ZEB2 coordinated developmental pathways related to brain size and cognitive abilities. Our study offers valuable candidates for future studies and furthers our understanding of the role of ZEB2 in the evolution of the human brain.

670 - Molecular responses of brain functions to chronic low dose ionizing radiation in wild mammals from the Chernobyl Exclusion Zone

Rebecca Saager

Rebecca S Saager, Vladimir M Jovanovic, Zbyszek BoratyÅ„ski , Phillip Lehmann, Katja Nowick, Anne Hartleib

Freie UniversitÃ¤t Berlin, Freie UniversitÃ¤t Berlin, CIBIO, UniversitÃ¤t Greifswald, Stockholm University, Freie UniversitÃ¤t Berlin, Freie UniversitÃ¤t Berlin

Ionizing radiation (IR) impacts cognition, brain size, and brain development in humans and other animals. Cognition is a crucial trait for animals adapting to stressors like IR contamination. Here, we analyzed RNA-sequencing data from the frontal cortex, visual cortex, hippocampus, and amygdala of bank voles (Myodes glareolus) from different areas in the Chernobyl Exclusion Zone with varying levels of IR exposure, to discover molecular determinants in the response to increased radioactive contamination in an ecological realistic context. We assessed gradual and categorical gene expression patterns in relation to internal dose rate, identifying numerous genes potentially affected by radiation. Of the 198 genes with categorical expression changes affected by IR, four were differentially expressed across all studied tissues. The hippocampus, the brain region involved in learning and memory, showed the highest number of differentially expressed genes. Gene set enrichment within the affected hippocampal genes points to a change in the astrocyte function(s) in this brain region. This further highlights the impact of IR on modulating spatial memory, reward and fear behavioral response. We also performed weighted gene co-expression network analysis to elucidate molecular pathways involved in bank vole adaptation to radiation. We further validated our results using matching tissues of the yellow-necked mice (Apodemus flavicollis) from the same study area. Overall, this research advances our understanding of the ecological and physiological consequences of nuclear accidents and provides valuable insights into the broader impacts of IR on wildlife and humans alike.

675 - Mammalian implantation and subsequent pregnancy – a story of miRNA evolution and co-option.

Nkem Nwaobi

Nkem Nwaobi, Jonathan Fenn, James McInerney, Niamh Forde, Mary O'Connell

University of Nottingham, University of Nottingham, University of Nottingham, University of Leeds, University of Nottingham

MiRNAs have long been associated with animal innovation. Specifically, it is known that bursts of miRNA diversity coincide with functional and morphological innovation. Indeed, placental mammals have 13 miRNA families identified thus far that are unique and conserved in this clade. However, we know from gene regulatory network biology that â€œnewerâ€• miRNAs have fewer targets and major transitions such as the shift in reproductive strategy in mammals from oviparity to vivipary are most likely supported by a combination of new miRNAs and older miRNAs. Here we consider the role of â€œolderâ€• miRNAs (i.e., miRNAs that predate the origin of mammals) in the evolution and diversification of mammal reproductive strategies. Using a set of >400 mammal genomes, we have applied machine learning approaches to identify miRNAs and their homologs across all species in our set. From our large presence absence matrix, we extracted all miRNAs that emerged prior to mammal divergence, i.e. are â€œolderâ€•, and we mapped their evolutionary trajectories (e.g., expansions and contractions of miRNA gene families). Then using independent methods have identified the targets, pre- and post-mammal origin. Here we explore shifts in the regulatory network through time and to identify novel patterns coincide with the origin of mammals. The likely impacts of these â€œolderâ€• miRNAs and their target profiles on mammal implantation and subsequent embryonic development are discussed.

680 - Production of lanthipeptides with potential antimicrobial activity in E. coli obtained through

genomic mining

Jonathan Eduardo Gómez González

Jonathan E. GÃ³mez GonzÃ¡lez, Marlene Guerrero Acevedo, MoisÃ©s A. Alejo HernÃ¡ndez, Katia P. Villavicencio SÃ¡nchez, Carol Siseth MartÃ-nez Caballero, Diana Corina CeapÄƒ
UNAM, UNAM, UNAM, UNAM, UNAM, UNAM

Antimicrobial resistance (AMR) stands as one of the most significant public health challenges of our time, and the World Health Organization (WHO) has made it a priority to invest in new medicine, diagnostic tools, vaccines, and interventions to address it through its Global Action Plan on Antimicrobial Resistance. In this study, we aimed to investigate biosynthetic gene clusters (BGCs) associated with lanthipeptides, which are natural products belonging to RiPPs (ribosomally synthesized and post-translationally modified peptides). Using genomic mining, we aimed to identify LanM amino acid sequences characterized experimentally, consulting the MIBiG database and conducting homology analysis on 232 genomes via BLAST. We then utilized the antiSMASH platform to identify complete BGCs related to these LanM sequences. Certain BGCs were selected for manual characterization based on the presence of essential genes. To avoid redundancy, we analyzed the precursor peptide sequences of competent BGCs using sequence similarity networks with already characterized lanthipeptides. This process enabled us to identify candidates that strike a balance between being a part of an existing lanthipeptide network and possessing sufficient structural variation. Sixteen promising candidates were chosen and grouped into sets of six, each contained in a pRSFDuet-1 vector for heterologous expression using E. coli NiCo. Our team predicted the chemical structures of the candidates and intends to evaluate their antibacterial activity against clinically relevant multidrug-resistant bacteria. This study highlights genome mining's potential in discovering new class II lanthipeptides, potentially leading to novel antimicrobial agents' development.

685 - Insights into Human Neurotransmission through Molecular Evolution Studies in Bacteria

Paul C Taylor

Jake A Fowler, Jennifer A Miles, Andrew D Millard, David J Scanlan, Michael E Webb, Ana M Rojas-Mendoza, Paul C Taylor
University of Leeds, University of Leeds, University of Leicester, University of Warwick, University of Leeds, Centro Andaluz de BiologÃ-a del Desarrollo, University of Leeds

Neurotransmission is fundamental to human health. Disorders cause e.g. Parkinsonâ€™s disease and mental health conditions (10.1001/archneur.1986.00520100062016). While major advances in medicine resulted from study of species like E coli and yeast (10.15698/mic2022.04.773), unicellular organisms offer limited insights into neurotransmission. We research homologs of proteins involved in neurotransmission that are found in multicellular cyanobacteria (10.1016/j.bbrc.2021.09.071). We focus on the archetypal machinery of neurotransmission (glutamate receptors, calmodulin) and enzymes key to biosynthesis of neurotransmitters including dopamine and serotonin.

The dominant mode of neurotransmission involves activation of ionotropic glutamate receptors (iGluR) that allow calcium ions to enter neurons. Cyanobacteria are known to have homologs (GluR0) of iGluR (10.1016/j.jmb.2007.10.081), but to date there is no comprehensive molecular evolution study of the links between iGluR and GluR0. Using a combination of phylogenetic analysis, statistical coupling analysis and modelling, we identified key features of extant iGluR and GluR0 that emerged during critical stages of their evolution and that may have as yet unremarked medical significance.

Once in the neuron, calcium binds to calmodulin. We showed that the closest bacterial homologs to animal calmodulins are

calmodulin-like domains in cyanobacterial proteins. We used NMR spectroscopy to show how calcium affects the conformation of one such domain from Okeania.

GTP cyclohydolase I (GCH1) is key to biosynthesis of tetrahydrobiopterin, necessary for neurotransmitter biosynthesis. Reminiscent of GluR, calmodulin and our older work (10.1016/j.bbrc.2018.12.171;10.1007/s00239-020-09965-x), phylogenetic analysis shows cyanobacterial GCH1 are the most closely-related prokaryotic enzymes to animal GCH1. We will use them to probe GCH1 kinetics.

693 - A complex landscape of introgressions in a diverse lineage of neotropical yeasts

J Abraham Avelar-Rivas

J Abraham Avelar-Rivas, Luis Fernando GarcÃ-a Ortega, Ivan SedeÃ±o, Antonio Urban-Aragon, Claudio Lopez, Eugenio Mancera, LucÃ-a Morales, Alexander DeLuna

Unidad de GenÃ³mica Avanzada (Langebio), Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Irapuato, Mexico., na, Departamento de IngenierÃ-a GenÃ©tica, Unidad Irapuato, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Irapuato, Mexico., Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano (LIIGH), Universidad Nacional AutÃ³noma de MÃ©xico, Juriquilla, Mexico., 3Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano (LIIGH), Universidad Nacional AutÃ³noma de MÃ©xico, Juriquilla, Mexico., Unidad de GenÃ³mica Avanzada (Langebio), Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Irapuato, Mexico., Departamento de IngenierÃ-a GenÃ©tica, Unidad Irapuato, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Irapuato, Mexico., Departamento de IngenierÃ-a GenÃ©tica, Unidad Irapuato, Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Irapuato, Mexico., Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano (LIIGH), Universidad Nacional AutÃ³noma de MÃ©xico, Juriquilla, Mexico., Unidad de GenÃ³mica Avanzada (Langebio), Centro de InvestigaciÃ³n y de Estudios Avanzados del Instituto PolitÃ©cnico Nacional, Irapuato, Mexico.

Saccharomyces cerevisiae from the tropical Americas has scarce representation in genome-wide population studies of the species, which includes wild, domesticated, and feral yeasts from diverse habitats. To gain insights into the evolution of these yeasts living at the intersection of fermentative and wild contexts, we sequenced the genomes of over 200 strains isolated from open traditional agave fermentation across Mexico. Analysis of genetic variation in the context of isolates from all around the globe revealed a specific neotropical clade with strains isolated from Brazil to Northern Mexico. These genomic sequences formed a highly divergent branch with widespread introgressions from the sister species S. paradoxus. Phylogenomic analyses identified structured subpopulations within the clade, which correlated with geography and population parameters. Intriguingly, each subpopulation showed a specific genomic distribution of introgressions and the origins of such components were from different lineages of S. paradoxus. This suggests a still changing introgression landscape within the lineage, arguably reflecting ongoing hybridization, loss of heterozygosity, and backcrossing. Our findings underscore the contribution of population structure and introgressions to yeast diversification in a region characterized by its high biodiversity.

710 - Serpentinization as the source of energy, electrons, organics, catalysts, nutrients and pH gradients for the origin of LUCA and life

Loraine Schwander

Loraine Schwander, Max Brabender, Natalia Mrnjavac, Jessica L.E. Wimmer, Martina Preiner, William F. Martin

Institute of Molecular Evolution, Biology Department, Math. -Nat. Faculty, Heinrich-Heine-UniversitÃ¤t, Institute of Molecular Evolution,

Biology Department, Math. -Nat. Faculty, Heinrich-Heine-UniversitÃ¤t, Institute of Molecular Evolution, Biology Department, Math. -Nat. Faculty, Heinrich-Heine-UniversitÃ¤t, Institute of Molecular Evolution, Biology Department, Math. -Nat. Faculty, Heinrich-Heine-UniversitÃ¤t, Microcosm Earth Center, Max Planck Institute for Terrestrial Microbiology and Philipps-UniversitÃ¤t, Institute of Molecular Evolution, Biology Department, Math. -Nat. Faculty, Heinrich-Heine-UniversitÃ¤t

Serpentinization in hydrothermal vents is central to some autotrophic theories for the origin of life because it generates compartments, reductants, catalysts and gradients. During the process of serpentinization, water circulates through hydrothermal systems in the crust where it oxidizes Fe (II) in ultramafic minerals to generate Fe (III) minerals and H2. Molecular hydrogen can serve as a source of electrons for the reduction of CO2 to organic compounds, with suitable catalysts present. Using catalysts during serpentinization H2 reduces CO2 to formate, acetate, pyruvate, and methane. These compounds represent the backbone of microbial carbon and energy metabolism in acetogens and methanogens, anaerobic chemolithoautotrophs that use the acetyl-CoA pathway of CO2 fixation and that inhabit serpentinizing environments today. Serpentinization generates reduced carbon, nitrogen and probably reduced phosphorous compounds that were likely conducive to the origins process. It gives rise to microcompartments and proton gradients to support chemiosmotic ATP synthesis by the rotor-stator ATP synthase. This would help to explain why the principle of chemiosmotic energy harnessing is more conserved than the machinery to generate ion gradients via pumping coupled to exergonic chemical reactions, which in the case of acetogens and methanogens involve H2-dependent CO2 reduction. Serpentinizing systems exist in terrestrial and deep ocean environments and were probably even more abundant on the early Earth. Serpentinization once occurred on Mars and is likely still occurring on Saturnâ€™s icy moon Enceladus, providing a perspective on serpentinization as a potential source for life on other worlds.

717 - Insights into the genetic bases of differential organ size in cacti flowers

Isaura Rosas Reinhold

Isaura Rosas-Reinhold, Gustavo RodrÃ-guez-Alonso, Cristian Cervantes, Genaro RamÃ-rez-Castro, Ulises Rosas, Alma PiÃ±eyro-Nelson, Salvador Arias

Center for Genomics and Systems Biology, new york university, Centro de InvestigaciÃ³n en DinÃ¡mica Celular. Instituto de InvestigaciÃ³n en Ciencias BÃ¡sicas y Aplicadas. Universidad AutÃ³noma del Estado de Morelos, Centro de InvestigaciÃ³n en DinÃ¡mica Celular. Instituto de InvestigaciÃ³n en Ciencias BÃ¡sicas y Aplicadas. Universidad AutÃ³noma del Estado de Morelos., Instituto de BiologÃ-a, Universidad Nacional Autonoma de Mexico, Jardin botÃ¡nico, Instituto de BiologÃ-a, Universidad Nacional Autonoma de Mexico, Jardin botÃ¡nico, Instituto de BiologÃ-a, Universidad Nacional Autonoma de Mexico, Departamento de ProducciÃ³n AgrÃ-cola y Animal, Universidad AutÃ³noma Metropolitana-Xochimilco, Jardin botÃ¡nico, Instituto de BiologÃ-a, Universidad Nacional Autonoma de Mexico

Floral morphological variation is a main driver of angiosperm evolution, changes in perianth pigmentation, floral organ number and size have been related to diversification and speciation. Developmental genetic studies in model species unveiled genes that directly control organ size. The roles of homologs of these genes in angiosperms with complex ontogenies such as the flower-shoot of Cactaceae have not been analyzed. Here we present a comparative transcriptomic analysis across flower buds of different growth stages from two cacti species with contrasting flower size: Disocactus speciosus and D. eichlamii. Differential gene expression analyses show that genes associated with flower size have different expression patterns in these species. Furthermore, a Weighted Gene Co-expression Network Analysis showed differential enrichment in processes related to organ growth between the initial size and the final size. Cell measurements and scanning electron microscope photographs of tepals of growing flower buds suggest that different cellular dynamics underlie final perianth size and support the transcriptome results. Our data suggest that differential regulation of specific genes could play a major role in defining flower size. In particular, BIG BROTHER and BIG PETAL, which are negative regulators of cell proliferation and cell expansion, respectively, were found to be differentially expressed in D. eichlamii and D. speciosus. These results suggest that these

two genes could play an important role in flower size diversification in these species.

736 - Expression of Specialized Metabolites by an OSMAC Approach Applied to Streptomyces

Uriel Colin Camacho

Uriel Colin Camacho, Corina-Diana CeapÄƒ
Institute of Chemistry, UNAM, Institute of Chemistry, UNAM

The Streptomyces genus, found within the Actinobacteria phylum, possesses an average of 39 biosynthetic gene clusters (BGCs) responsible for producing diverse microbial specialized metabolites (MEMs). However, many of these BGCs remain inactive or 'silent' under laboratory conditions. The 'one strain, many compounds' (OSMAC) approach has been demonstrated to activate these BGCs, making it a promising strategy for developing new bioactive compounds. Therefore, we aimed to extract specialized metabolites from Streptomyces using an OSMAC approach. Streptomyces colonies obtained from environmental soil samples were cultured, isolated, and purified. They were then grown in liquid SCA medium along with 1% EtOH or 3% DMSO. Their specialized metabolites were extracted using ethyl acetate, and their antimicrobial activity against pathogenic bacteria was evaluated. A total of 9 isolates were obtained, of which 3 showed a higher potential to generate pigments by Petri dish assay when 1% EtOH or 3% DMSO was added. These 3 isolates will be used as the first batch to extract their metabolites and evaluate their antimicrobial activity.

738 - An interdisciplinary approach to ascertaining the utility of various skeletal elements in non-human aDNA studies

Sophy Charlton

Sophy Charlton, Kristina Tabbada, Evangelos Dimopoulos, Lachlan Scarsbrook, GiedrÄ— PiliÄ•iauskienÄ—, Povilas BlaÅ¾eviÄ•ius, Laurent Frantz, Greger Larson
University of York, University of Oxford, University of Cambridge, University of Oxford, Vilnius University, Vilnius University, Queen Mary University of London, Ludwig Maximilian University, University of Oxford

All ancient DNA (aDNA) studies are limited by the degraded and fragmented nature of aDNA, as well as the low endogenous DNA content of the majority of archaeological samples. In recent years, the targeted sampling of specific skeletal elements in humans - particularly the petrous portion of the temporal bone (pars petrosa) â€“ has been utilised to mitigate some of these issues. In contrast, the potential of similar approaches in the sampling of faunal populations is yet to be fully explored. Here, in order to empirically characterise the utility of petrous for aDNA analysis relative to other skeletal elements in non-human species, we present the first comparative aDNA study utilising paired dog (Canis lupus familiaris) skeletal elements (petrous vs. bone/tooth) (n=56). We informed our targeted sampling strategy through the use of micro-CT scanning, indicating denser areas of cortical bone. This allowed for the exploration of aDNA preservation across skeletal elements, considering both endogenous and exogenous fractions. Our results suggest that endogenous DNA yields from petrous may be up to 200x higher than other skeletal elements from the same individual. However, petrous appear to have lower microbial species richness when compared to other skeletal elements, including pathogens. We leverage the findings of our research to identify the most suitable substrate(s) for the genetic analysis of archaeological dog specimens and other mammalian fauna.

787 - Physiological and Transcriptomic Mechanisms of Hypoxia Acclimation in High-Altitude Deer Mice

Naim M Bautista

Naim Bautista, Nathanael Herrera, Ellen Shadowitz, Oliver Wearing, Zachary Cheviron, Graham Scott, Jason Storz
University of Nebraska Lincoln, University of Montana, McMaster University, McMaster University, University of Montana, McMaster University, University of Nebraska Lincoln

A fundamental question in evolutionary biology concerns the relative contributions of phenotypic plasticity vs. local adaptation (genotypic specialization) in enabling wide-ranging species to inhabit diverse environmental conditions. In species that are distributed across elevational gradients, plasticity in environmental tolerance can be expected to influence the ability of species to track upward shifts in temperature isoclines as a result of climate change. Here we conduct a long-term hypoxia-acclimation experiment to assess the relative roles of local adaptation and plasticity in enabling highland and lowland deer mice (Peromyscus maniculatus) to sustain aerobic thermogenesis at simulated elevations that span the speciesâ€™ current range and exceed the upper limits. We measured whole animal physiological performance (thermogenic VO2max) in conjunction with physiological traits and tissue-specific transcriptomes to gain insight into underlying mechanisms of plasticity. Our results demonstrate that highland natives exhibit a superior thermogenic capacity at the most severe levels of hypoxia, suggesting that the realized niche of the species and its ability to inhabit such a broad range of elevational zones is attributable to a combination of genetically based local adaptation and evolved changes in the plastic response to hypoxia.

804 - Consequences of indirect exposure to BPA during embryonic development in live-bearing fish

Yolitzi Saldivar Lemus

Yolitzi Saldivar Lemus, Saeid Panahi Hassan Barough, Caitlin Gabor
Texas State University, Texas State University, Texas State University

Bisphenol A (BPA) is a chemical primarily used in the production of polycarbonate plastics and can cause endocrine disruption, changes in DNA methylation patterns, gene expression, behaviour, and affect axonal growth and muscle development. Many studies on BPA, however, are often conducted with very high BPA concentrations that are not environmentally relevant, and, in the case of fish, are restricted to only oviparous species. We used the viviparous and highly matrotrophic fish Heterandria formosa to look at the effects of an environmentally relevant concentration of BPA (2 ug/L) on behavior, cognition, physiology and gene expression in newborn offspring and their mothers that were exposed to BPA during gestation. We found no significant difference in offspringâ€™s and motherâ€™s cortisol release, and in offspringâ€™s cognitive performance across treatments, however, BPA exposed offspring swam slower and spent more time in proximity to other fish. We also detected changes in expression of genes related to the immune system and cell proliferation and negative regulation of cell death. Offspring in the BPA treatment had genes related to the immune system downregulated, and genes related cell proliferation and negative regulation of cell death upregulated, which has been associated to cancer. Additionally, we conducted WGCNA analysis, and we obtained two modules that highly correlated with the BPA treatment (green and yellow). Interestingly, the yellow module that negatively correlated with BPA is mainly related to embryonic development and organ morphogenesis. Our results indicate that even low concentrations of BPA can have detrimental effects on developing organisms.

811 - A reformulation of measures of linkage disequilibrium and population structure under drift, inbreeding, and tight linkage: An approach through probabilities of identity by state

Marcy Uyenoyama

Marcy Uyenoyama
Duke University

Observations of genome- and planet-wide patterns in linkage disequilibrium (LD) and measures of population structure (especially FST) have served as the basis for inferring demographic history. Most definitions of FST that involve probabilities of identity-by-descent (IBD) or variance components take an approach through pedigrees, with reference to an ancestral base population. The dependence of classical measures of FST on allele frequencies can cloud their interpretation as indicators of population structure. A reformulation of Nei's GST in terms of identity-by-state (IBS) probabilities clarifies the meaning of Wright's F-statistics. In place of references to allele frequencies in the ancestral base population, this approach entails a locus-specific accounting of the mutation process itself.

Similarly, accounting for both the existence of variation as well as its pattern can simplify the inference framework for the analysis of associations among distinct loci. Under this approach, a compact method for deriving the probabilities of the joint observation IBS from a pair of loci provides diversity-based measures of LD from unphased samples. Of particular interest are associations that arise among sites separated by just tens of kilobases, implying rates of crossing-over on the order of 10^{-5}. Interactions among genetic drift, mutation, crossing-over, and regular inbreeding determine levels of within-locus genetic diversity and between-locus identity-by-state.

828 - Context-specific mutation rate variation across the tree of life

Fabian Ramos-Almodovar

Fabian Ramos-Almodovar, Benjamin F. Voight, Iain Mathieson
University of Pennsylvania, University of Pennsylvania, University of Pennsylvania

Germline mutations arise from a combination of DNA damage and replication errors. One hypothesis posits that variation in the efficiency of DNA repair in the germline explains the observed variation in mutation rates across species, although this has mainly been explored in mammals. Another possibility is that exogenous mutagens also contribute to mutation rate variation. Different mutational processes cause different types of mutation â€œsignaturesâ€•, and we can exploit this to measure their relative importance. To explore the distribution of mutational processes across a broad phylogenetic context, we applied Baymer, a Bayesian hierarchical tree model, to infer windows of nucleotide sequence contexts (up to 4 flanking nucleotides, i.e., â€˜9-mersâ€™) that capture variation in mutation rates across 70 eukaryote species using polymorphism datasets. We applied phylogenetic linear mixed models to explore the role of biological and environmental modifiers of C > T mutations at CpG sites. We confirmed that variation in CpG > T mutation rates is highly correlated with germline methylation levels, with patterns consistent with the phylogeny. Additionally, we analyzed mutational signatures from the Catalogue Of Somatic Mutations In Cancer (COSMIC), which are believed to represent different mutational mechanisms that could contribute to germline mutagenesis. We explore the distribution of the 86 COSMIC signatures across species, reflecting both genetic mechanisms and potential environmental mutagens. These results allow us to explore the factors affecting mutation rates with data that encompasses a wide range of biological and environmental diversity.

835 - Investigating the "neighbor repulsion" pattern in Patterson’s f-statistics

Mehmet Somel

GÃ¶zde AtaÄŸ, Mehmet Somel
METU, MPI-EVA, METU

Here we study a surprising pattern arising in population genetics analyses involving Patterson's outgroup-f3 and f4 statistics, where groups sharing a high fraction of common ancestry -often geographically neighbors- show higher apparent affinity to a more "distant" group. Such a pattern has been described for North African modern-day groups showing higher affinity to Sardinians than to their neighbors. We now present a new case involving Bronze Age East Mediterranean populations, where Anatolian groups choose

populations from Greece rather than their neighbors, in seeming contradiction with patterns observed in PCA and MDS analyses. We next perform coalescent simulations involving low-level gene flow independently into "sister" groups from an external population with high within-population diversity. This recapitulates the observed f-statistic patterns, where the sister groups show higher affinity to a more distant but unadmixed population. We also show that, in the same simulations, summaries of outgroup-f3 statistics using MDS can capture overall shared ancestry patterns, similar to our empirical observations. We note that this behavior of both outgroup-f3- and f4-statistics is theoretically expected as these are single estimates representing genetic correlations arising from mixed ancestry. Our results illuminate how low-level and heterogeneous admixture can produce unexpected population genetic statistics.

842 - Intrinsic expression and subcellular localization specificity in de novo emerged proteins.

Carly Houghton

Carly Houghton, Nelson Castilho Coelho, Annette Chiang, Aaron Wacholder, Allyson F. O'Donnell, Anne-Ruxandra Carvunis
University of Pittsburgh, School of Medicine, Department of Computational and Systems Biology, Center for Evolutionary Biology and Medicine, University of Pittsburgh, School of Medicine, Department of Computational and Systems Biology, Center for Evolutionary Biology and Medicine, University of Pittsburgh, Department of Biological Sciences, University of Pittsburgh, School of Medicine, Department of Computational and Systems Biology, Center for Evolutionary Biology and Medicine, University of Pittsburgh, Department of Biological Sciences, University of Pittsburgh, School of Medicine, Department of Computational and Systems Biology, Center for Evolutionary Biology and Medicine

De novo gene birth is the process by which new protein-coding genes arise from previously non-coding DNA. Recent findings have unveiled of thousands of de novo open-reading frames (ORFs) that are translated in cells. Upon emergence of a de novo ORF, the new protein encoded has had no prior natural selection on its protein sequence. Although certain de novo proteins adopt secondary structures and impact cellular fitness, little is known about their fate of following translation. Are they rapidly degraded or sequestered in specific subcellular niches? Do these unevolved proteins display a similar localization diversity to that observed for the conserved proteome? To address these questions, we used an integrated approach in which we genetically engineered yeast to express approximately 250 fluorescently-tagged de novo proteins under an inducible expression system. We then systematically determined their in vivo protein abundance and subcellular localization. Our findings reveal that, despite being expressed from a common promoter, de novo proteins exhibit sweeping differences in protein abundance, suggesting that they vary greatly in either translation efficiency or protein stability. Moreover, de novo proteins display distinct localization patterns compared to the conserved proteome, including being over and underrepresented in the mitochondria and the nucleus, respectively. Finally, we found that particular localizations are associated with unique sequence features within de novo proteins. This study provides the first systematic examination of de novo proteins in vivo and paves the way to understanding the potential phenotypic consequences and the evolutionary implications of their existence.

843 - Serological analysis of the immune response to pathogens in the Mexican Biobank

Mayte Celeste Cervantes-Hernandez

MaytÃ© C. Cervantes-HernÃ¡ndez, Luis Pablo Cruz-Hervert, Amanda Y. Chong, SofÃ-a Vieyra-SÃ¡nchez, Alexander J. Mentzer, Tim Waterboer, Lourdes GarcÃ-a-GarcÃ-a, AndrÃ©s Moreno-Estrada
National Laboratory of Genomics for Biodiversity (UGA-LANGEBIO), CINVESTAV, Universidad Nacional AutÃ³noma de MÃ©xico , Wellcome Trust Centre for Human Genetics. University of Oxford, National Laboratory of Genomics for Biodiversity (UGA-LANGEBIO), CINVESTAV, Wellcome Trust Centre for Human Genetics. University of Oxford, German Cancer Research Center, Instituto Nacional de Salud PÃºblica , National Laboratory of Genomics for Biodiversity (UGA-LANGEBIO), CINVESTAV

Population studies of pathogens are vital for understanding the distribution of infectious diseases. When coupled with DNA biobanks, antibody response can be correlated with genetic and clinical factors, leading to potential public health interventions. Here, we analyze serology data for 41 antigens specific to 20 pathogens in 5,798 individuals from the Mexican Biobank. Samples were distributed nationwide across Mexico and were collected as part of the 2000 National Health Survey including socio-demographic data. Serology measurements were obtained using a multiplex panel based on Luminex beads technology previously utilized to measure a number of pathogens similar to that of the UK Biobank. We used antibody titers cutoffs for each antigen to calculate the prevalence of medically relevant pathogens, including HPV, EBV, HIV, Hepatitis B and C, Helicobacter pylori, and Toxoplasma gondii. We compared our results with the UK Biobank as well other sources. We found that the pathogens with the highest prevalence in the MX Biobank also have a high prevalence globally, except for H. pylori which exhibits a higher prevalence in Mexico. In a logistic regression model, no significant differences were found in seropositivity of pathogens between sexes. However, significant differences were found when comparing speakers vs. non-speakers of indigenous languages and comparing rural vs urban areas. This work represents a valuable resource for epidemiological studies in Mexico and future studies of genetic association with immune response to certain pathogens.

846 - The first reference level genome of the Caribbean long-spined black sea urchin (Diadema antillarum)

Audrey Majeske

Audrey Majeske, Carlos Farkas Pool, Juliet Wong, Alejandro Mercado Capote, Alondra Diaz-Lameiro, Jose Eirin-Lopez, Nikolaos Schizas, TarÃ¡s Oleksyk

Oakland University, University of Puerto Rico MayagÃ¼ez, Universidad CatÃ³lica de la SantÃ-sima ConcepciÃ³n, Duke University, University of Puerto Rico MayagÃ¼ez, University of Puerto Rico MayagÃ¼ez, Florida International University, University of Puerto Rico MayagÃ¼ez, Oakland University

We generated the first reference level nuclear genome assembly of the keystone Caribbean black long-spined sea urchin species (Diadema antillarum). We employed different assembly strategies using whole genome sequences acquired from multiple sequencing platforms (Illumina, PacBio and Oxford Nanopore) to generate a high-quality near-complete assembled genome. We compared the output metrics of several different assembly programs (hifiasm, flye and SPAdes), and then improved these through scaffolding using the RagTag tool. The efficacy of our assembly strategy was underscored by a BUSCO completeness score of 98.4% (based on metazoa_odb10), with the genome comprising 2789 contigs.Â The length of this resultingÂ genome assembly isÂ 1.75Â Gbps with an N50 of 1.57 Mbps.

848 - Modeling the emergence of gene clusters under interlocus gene conversion to elucidate the evolutionary history of a heavy metal adaptation in Arabidopsis halleri

Yannick Schäfer

Yannick SchÃ¤fer, Yichen Zheng, Akash Chandra Parida, BjÃ¶rn Pietzenuk, Piyal Karunarathne, Laura Rose, Ute KrÃ¤mer, Thomas Wiehe

University of Cologne, University of Cologne, Ruhr University Bochum, Ruhr University Bochum, Heinrich Heine University DÃ¼sseldorf, Heinrich Heine University DÃ¼sseldorf, Ruhr University Bochum, University of Cologne

Plants have evolved a plethora of adaptations to extreme environmental conditions. One such extremophile organism, which colonizes heavy metal-rich soils, is Arabidopsis halleri. A. halleri possesses a cluster of three copies of the Heavy Metal ATPase4 (HMA4) gene encoding a Zinc and Cadmium pump which facilitates root-to-shoot transport of these metals.

Analysis of genomic fragments from the HMA4 cluster suggested that the gene copies maintain highly similar sequences due to interlocus gene conversion (IGC).

PacBio sequencing was used to generate long read sequence data for multiple accessions of A. halleri, uncovering further evidence for IGC among the three HMA4 gene copies. Moreover, the data provides a basis for studying IGC in additional gene families that underwent recent expansion.

We conducted extensive simulations with a model of sequence divergence between newly duplicated gene copies, informed from the experimental data. In this model, we tested the effects of recombination, IGC and selection. In particular, an interaction between purifying selection and IGC rate decreasing with increasing sequence divergence leads to a mosaic pattern of inter-copy diversity.
Given specific parameters, exons and short introns can enter an equilibrium between IGC and mutations, while long introns escape the equilibrium and diverge continuously. This effect was also observed in the HMA4 genes. Estimating duplication time from divergence level is a complex task due to overestimation caused by reciprocal recombination and underestimation caused by IGC. Additionally, we explored population genetic estimators within and between simulated copies to detect and quantify genetic signals of IGC and selection.

856 - Comparative genomics and the salivary transcriptome of the redbanded stink bug shed light on its high damage potential to soybean

Federico G. Hoffmann

Hunter K. Walt, Federico G. Hoffmann
Mississippi State University, Mississippi State University

The redbanded stink bug, Piezodorus guildinii (Westwood) (Hemiptera: Pentatomidae), is a significant soybean pest in the Americas, inflicting more physical damage on soybean than other native stink bugs. Studies suggest that its heightened impact is attributed to the aggressive digestive properties of its saliva. Despite its agricultural importance, the factors driving its greater ability to degrade plant tissues have remained unexplored in a genomic evolutionary context. In this study, we hypothesized that lineage-specific gene family expansions have increased the dosage of digestive genes expressed in the salivary glands. To investigate this, we annotated a previously published genome assembly of the redbanded stink bug and performed a comparative genomic analysis on 11 hemipteran species and reconstructed patterns of gene duplication, gain, and loss in the redbanded stink bug. We also performed RNA-seq on the redbanded stink bugâ€™s salivary tissues, along with the rest of the body without salivary glands. We identified hundreds of differentially expressed salivary genes, including a subset lost in other stink bug lineages but retained and expressed in the redbanded stink bugâ€™s salivary glands. These genes were significantly enriched with protein families involved in proteolysis, potentially explaining the redbanded stink bugâ€™s heightened damage to soybeans. Contrary to our hypothesis, we found no support for an increased dosage of digestive genes in the salivary glands of the redbanded stink bug. Nonetheless, these results provide insight into the evolution of this important crop pest, establishing a link between its genomic history and its agriculturally important physiology.

866 - Genome evolution in parasitic plants; Insights from the first chromosome assemblies in Cuscuta

Juan Daniel Cerda

Juan Daniel Cerda, Claude dePamphilis, Eric Wafula, Michael Axtell, Jim Westwood, Paula Ralph
The Pennsylvania State University, The Pennsylvania State University, The Pennsylvania State University, The Children's Hospital of

Philadelphia, The Pennsylvania State University, Virginia Tech, The Pennsylvania State University

The parasitic plant Cuscuta play both positive (diversity drivers) and negative (pests) roles in the ecosystem and in agriculture and is a model for mobile RNA (miRNA & microRNA) between host and parasite. Current studies on Cuscuta genome discovered large scale gene loss as well as numerous genes associated to the haustoria (feeding structure), but no study has done an extensive exploration of gene evolution in Cuscuta, much less genome evolution. In this study, we will attempt to answer the questions: How did parasitism evolved in Cuscuta? And how does a parasitic lifestyle affect the genome evolution of Cuscuta gronovii? To answer these questions, we created a chromosome level genome assembly for Cuscuta. gronovii and Cuscuta campestris. We incorporated the genomes of Cuscuta australis, five Ipomoea species and Solanum lycopersicum with a comparative genomics approach to discern gene and genome evolution. Chromosome level assembly showed chromosome number close to expected for Cuscuta campestris (n = 30 vs n = 31), but half the previously established numbers for Cuscuta gronovii (n = 30 vs n = 15). We found that 46% and 81% of the genomes of Cuscuta campestris and Cuscuta gronovii respectively are comprised of repetitive elements. We deduced that the extreme genome size in Cuscuta gronovii is caused by a recent expansion of LTR repeats. This study has shown that Cuscuta genome plasticity is a complex issue worth revisiting in depth.

872 - A vast evolutionarily novel translatome participates in conserved cellular processes

Aaron Wacholder

Aaron Wacholder, Saurin Parikh, Nelson Castilho Coelho, Omer Acar, Carly Houghton, Lin Chou, Anne-Ruxandra Carvunis
University of Pittsburgh School of Medicine, University of Pittsburgh School of Medicine, University of Pittsburgh School of Medicine, University of Pittsburgh School of Medicine, University of Pittsburgh School of Medicine, University of Pittsburgh School of Medicine, University of Pittsburgh School of Medicine

Ribosome profiling experiments demonstrate pervasive translation of eukaryotic genomes outside of annotated coding sequences. This â€œnoncanonicalâ€• translatome encodes thousands of proteins that potentially participate in biological processes but have never been studied in depth. We developed a novel approach to identify the noncanonical translatome at high sensitivity by integrating ribosome profiling reads across hundreds of published experiments. Applying this approach to S. cerevisiae, we identified a yeast noncanonical translatome consisting of over 18,000 unannotated coding sequences. We then characterized the yeast translatome by evolutionary properties and phenotype.

Using comparative genomics, we find that the yeast translatome can be divided into two evolutionary classes. The smaller class consists of coding sequences that are conserved over long evolutionary time; these are mostly annotated and tend to be long and highly expressed. The larger class consists of evolutionarily young coding sequences that emerged recently from noncoding sequences; most are short, lowly expressed and unannotated. Employing a genetic screen in which translation is experimentally disabled, we find that a large fraction of both canonical and noncanonical recently-emerged coding sequences show substantial reductions of fitness despite lacking evolutionary conservation. Moreover, the potential roles of recently-emerged ORFs can be inferred from their interactors in the genetic interaction network. Novel ORFs are rapidly integrated into cellular networks, and involved in processes including RNA processing, DNA repair and cell growth.

Overall, our results demonstrate the existence of thousands of evolutionarily unconserved coding sequences, most of which are unannotated, that appear to play key roles in major biological processes.

894 - Evolutionary systems biology highlights diterpenoids as anti-herbivore defenses in above- and belowground rice tissues

Taryn Sena Dunivant

Taryn S Dunivant, Simon C Groen

University of California Riverside, University of California Riverside

Rice is a critical staple crop for global food security. However, rice is highly vulnerable to pests and pathogens, resulting in major yield reductions. Pest and pathogen pressure is predicted to increase with the heightened frequency of weather extremes and anticipated decreases in water availability associated with climate change. This concern becomes even more acute with the global push for direct-seeded rice cultivation because this is not possible without resistance to pests and pathogens. Knowledge of molecular resistance mechanisms is essential for developing novel approaches for pest and pathogen management in rice production. Studies that leveraged an evolutionary systems biology approach demonstrated its ability to predict mechanisms that elucidate genotype-phenotype relationships. Our reanalysis of transcriptomic data from a field-based systems biology study revealed pathways that show evidence of selection in rice sub-populations that experience relatively high pest and pathogen pressure and that are potentially involved in rice defense responses to both root-knot nematodes (RKNs, Meloidogyne) and Fall Armyworms (FAWs, Spodoptera). We identified that differentially expressed genes in rice roots and shoots responding to RKNs and FAWs, respectively, were significantly enriched within fitness-linked modules of a gene co-expression network constructed from traditional rice varieties with genetically and geographically diverse backgrounds. Key plant defense-related pathways enriched in these modules included the diterpenoid phytoalexins. To determine the role of various classes of diterpenoids in rice responses to above- and belowground herbivory we performed functional genetic tests on rice under attack by RKNs and FAWs.

904 - Recombination rate variation and the evolution of genomic landscapes in nematodes

Gavin Woodruff

Gavin C. Woodruff, Kimberly A. Moser

University of Oklahoma, University of Oklahoma

Genomes often exhibit heterogeneity across chromosomes, and distributions of protein-coding genes, repetitive elements, and polymorphisms are not uniform along chromosomes in multiple species. One explanation for these patterns is recombination rate variation. As recombination interacts with selection to shape the evolutionary fates of alleles, variation in recombination rate within and between chromosomes could promote differences in the distribution of genomic features across chromosomes. In the nematode Caenorhabditis elegans, recombination rate correlates with multiple genomic features that are non-uniformly distributed along chromosomes. Specifically, in C. elegans, recombination rates are higher on chromosome ends compared to chromosome centers. Its closest known relative, C. inopinata, harbors a radically altered genome with nearly uniform chromosomal distributions of repetitive elements and protein-coding genes; it likewise has a less heterogeneous chromosomal distribution of polymorphisms. Is this dramatic change in genomic organization in any way connected to the evolution of recombination rates? Here, we describe ongoing efforts to infer genetic maps in C. inopinata. We performed whole-genome sequencing on 180 individual F2 recombinants. Tentative, preliminary results reveal some chromosomes may have conserved recombination rate domain structure whereas other chromosomes may harbor divergent, more uniform recombination rate distributions. We intend to compare these chromosome-level recombination rates with genomic landscapes of genes, repeats, and polymorphisms across multiple species to understand the extent to which recombination rates predict genomic organization.

908 - On the evolution of gene product diversity

Daohan Jiang

Daohan Jiang, Nevraj Kejiou, Tina Qiu, Alexander Palazzo, Matthew Pennell
University of Southern California, University of Toronto, University of Toronto, University of Toronto, University of Southern California

RNA and protein expressed from the same gene can have diverse isoforms due to various post-transcriptional and post-translational modifications, yet functional significance of the substantial gene product diversity generated by these processes remains largely unknown, raising the question whether it is generally adaptive or errors in biochemical processes. Additionally, testing these hypotheses has been challenging because it is often unclear how variation isoform abundance relates to the underlying evolutionary processes. Here, we introduce a generalized theoretical model where cis- loci (sequence motifs recognized by a modification enzyme), trans- loci (loci underlying the modification enzymeâ€™s expression and/or activity), gene expression level, and isoform decay rates collectively shape isoform abundances to predict evolutionary dynamics under different regimes of selection. We derive distributions of isoform abundances across species and genes at mutation-drift-selection balance, focusing on two types of gene product diversity: the first, exemplified by RNA editing, generates modified isoforms that may or may not be functional; the second, exemplified by RNA splicing, produces functional isoforms alongside potentially non-functional ones. We demonstrate that factors beyond the selection regime significantly influence evolutionary dynamics; importantly, a population is more likely have a sub-optimal mean phenotype when the effective population size is small, the number of cis-acting loci is large, and/or different genes are subject to different regimes of selection. Our modeling framework provides a quantitative approach for investigating gene product diversity in future empirical studies to advance our understanding of the evolutionary forces shaping molecular diversity.

909 - Epistasis from conformational ensembles shapes evolution

Michael J Harms

Michael J Harms, Brennan A Fitzgerald, Natalie M Jaeger
Institute for Molecular Biology, University of Oregon, Department of Chemistry & Biochemistry, University of Oregon, University of Oregon, University of Oregon

Macromolecules exist as ensembles of interchanging conformations. Such dynamics are important for molecular function and regulation. We hope to convince you these ensembles are also profoundly important for protein evolution. Using theoretical and computational approaches, we found that subtle changes in ensemble composition can profoundly alter the effects of future mutations. Using experiments on the lac repressor and an RNA riboswitch, we revealed that these predicted effects are detectable both in vitro and in vivo. Using phylogenetic reconstructions and experimental studies of the innate immune protein S100A9, we revealed that molecular ensembles have indeed shaped the historical evolution of natural proteins. We are currently working on computational models to predict these effects for specific proteins, simulate protein evolution when taking conformational ensembles into account, and to define bioinformatic signatures of this phenomenon in multiple sequence alignments. Together, this work is revealing an intimate connection between molecular ensemblesâ€”an inescapable biophysical feature of macromoleculesâ€”and how macromolecules evolve. Specifically: subtle, functionally invisible, changes to the energies of low-population conformations can open and close future mutational trajectories, thus dramatically altering evolutionary outcomes.

914 - Sex-Specific Host Response to Staphylococcus aureus Skin and Soft Tissue Infection

Marijan Posavi

Marijan Posavi, Seth M Daly, Pamela R Hall
University of New Mexico CETI, University of New Mexico COP, University of New Mexico COP

Host sex can affect pathogen-induced immune responses. Evidence suggests that men have a higher risk of both colonization and infection with Staphylococcus aureus (MRSA), with an odds ratio of ~2.4 compared to women. The mechanisms underlying increased male predisposition to S. aureus infection remain unclear. To examine the genetic basis of this bias, we conducted a 2x2 factorial experiment and analyzed RNA-seq data in S. aureus-induced dermo-necrosis (SSTI) in mice. We examined differentially expressed (DE) genes between infected and control male and female mice. The RNA-seq read counts of each gene were modeled as a function of an effect of Sex (males vs. females), Condition (infected vs. control) and Sex-by-Condition interaction, while accounting for overdispersion using the negative binomial distribution employing DESeq2 R package. The Sex-by-Condition interaction unveiled 4034 DE genes indicating that the effect of infection on gene expression differs between males and females. Among these, 2486 were upregulated and 1548 were downregulated in infected males compared to infected females. These genes underwent enrichment analysis of GO and KEGG. Over-representation and GSEA analyses revealed 26 significantly enriched immune response GO biological processes. While the mechanisms behind these sex-specific DE mouse genes remain elusive, our analysis suggests that this bias stems from sex-specific genetic architecture, characterized by gene-sex interactions. Tailored approaches that consider sexual dimorphism - such as differences in susceptibility to infectious diseases, vaccination efficacy, and treatment responses - could significantly improve patient outcomes.

922 - UV light exposure in amphibians leads to CC:GG>TT:AA substitutions and non neutral amino acid shift

Valeria Vadimovna Lobanova

Valeria V. Lobanova, Alina G. Mikhailova, Konstantin Popadin
Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University, Immanuel Kant Baltic Federal University

External exposure to mutagens is one of the factors driving the molecular evolution of species and classes. For example, one of the strong mutagens influencing the occurrence of DNA damage and, as a consequence, substitutions and rearrangements, is ultraviolet light.

It has been shown that the signature of UV light has a known pattern, doublet base substitution CC>TT (DBS1 COSMIC sign, https://doi.org/10.1002/humu.22341). We hypothesize that since amphibians may inhabit solar environments and have transparent eggs (https://doi.org/10.1016/j.mrgentox.2019.03.003) they may have increased UV light explosion and the following increase in rates of CC:GG>TT:AA, observed in neutral sites (4-fold positions). Analysing whole nuclear genomes of dozens of amphibian species we have shown that indeed genomes accumulate TT- and AA-rich codons following the increased frequencies of CC:GG>TT:AA. It leads to an amino acid shift toward an increase in Asn, Lys and Leu, Phe, while Pro and Gly become losers.

943 - Coevolution between mRNA and protein abundances: resolving the roles of mutation and natural selection in shaping gene expression evolution using phylogenetic comparative methods

Alexander L Cope

Alexander L Cope, Joshua G Schraiber, Matthew W Pennell
Rutgers University, University of Southern California, University of Southern California

Gene expression variation is a major contributor to phenotypic variation across species. Numerous studies have sought to understand both the evolutionary processes and molecular mechanisms shaping across-species variation in gene expression. At the molecular level,

mRNA abundances are a key regulator of protein abundances, which are the â€œfunctional unitsâ€• of gene expression. Although these two levels of gene expression are undoubtedly correlated, comparative analyses revealed distinct evolutionary patterns of mRNA and protein abundances. Of particular note is the greater divergence across macroevolutionary timescales of mRNA abundances compared to protein abundances, leading to speculation of compensatory evolution across regulatory levels to maintain stable protein abundances. We developed a novel, mechanistic phylogenetic model of trait evolution to analyze the coevolutionary dynamics between mRNA and protein abundances across 11 mammalian species. Our model was fit to the data in a hierarchical manner, allowing us to share information across genes and compensating for the relatively small number of species analyzed. By doing so, we tease apart the relative contributions of mutation and natural selection in shaping mRNA-protein correlations across genes and species.

945 - Mean-field interacting multi-type birth-death processes with a view to applications in phylodynamics

William S. DeWitt

William S. DeWitt, Steven N. Evans, Ella Hiesmayr, Sebastian Hummel
University of California, Berkeley, University of California, Berkeley, University of California, Berkeley, University of California, Berkeley

Multi-type birth-death processes underlie approaches for inferring evolutionary dynamics from phylogenetic trees across biological scales, ranging from deep-time species macroevolution to rapid viral evolution and somatic cellular proliferation. A limitation of current phylogenetic birth-death models is that they require restrictive linearity assumptions that yield tractable likelihoods, but that also preclude interactions between individuals. Many fundamental evolutionary processes -- such as environmental carrying capacity or frequency-dependent selection -- entail interactions, and may strongly influence the dynamics in some systems. Here, we introduce a multi-type birth-death process in mean-field interaction with an ensemble of replicas of the focal process. We prove that, under quite general conditions, the ensemble's stochastically evolving interaction field converges to a deterministic trajectory in the limit of an infinite ensemble. In this limit, the replicas effectively decouple, and self-consistent interactions appear as nonlinearities in the infinitesimal generator of the focal process. We investigate a special case that is amenable to calculations in the context of a phylogenetic birth-death model, and is rich enough to model both carrying capacity and frequency-dependent selection.

947 - Molecular Footprints on Osmoregulation?Related Genes Associated with Freshwater Colonization by Cetaceans and Sirenians

Mariana F. Nery

Mariana Nery, Elisa Ramos
University of Campinas, University of Campinas

This research investigates the genetic foundations of adaptive physiological mechanisms in mammals that transitioned to aquatic lifestyles, with a particular emphasis on species moving from marine to freshwater habitats. Such transitions pose osmotic challenges, necessitating adaptations in the osmoregulatory system. We focused on the selective pressures acting on coding and regulatory regions of 20 genes related to osmoregulation in cetaceans and sirenians, which include species in both marine and freshwater environments. Our findings highlight positive selection in the vasopressin (AVP) gene in lineages that have adapted to freshwater habitats and in aquaporins across both types of environments. Dolphins displayed a higher number of positively selected sites compared to the Amazonian manatee, with AQP5 and AVP genes showing selection signals across several independent lineages. Vasopressin gene sequences in river dolphins were surprisingly similar across their independent evolutionary paths. The study also examined Transcription Factors in the promoter regions of these genes, revealing phylogenetic conservation among sister species. Genes such as ACE, AQP1,

AQP5, AQP7, AVP, NPP4, and NPR1, essential for osmotic regulation, showed signs of accelerated evolution in freshwater mammals. This research enriches our understanding of the genetic evolution underpinning osmoregulation in aquatic mammals, shedding light on the complex genetic adaptations to environmental changes.

969 - Inferring distributions of fitness effects of wild house mice from allele frequency spectra

Olivia Fernflores

Olivia Fernflores, David Castellano, Emanuel Masiero da Fonseca, Travis J Struck, Ryan N Gutenkunst
University of Arizona, University of Arizona, University of Arizona, University of Arizona, University of Arizona

The distribution of fitness effects (DFE) of new mutations is a key input into the evolutionary process. We aim to infer the DFE among multiple populations of wild house mice, so that the extensive knowledge of mouse molecular biology can be leveraged to understand the biological basis of the DFE. Here we present distributions of fitness effects for pairs of populations from Iran and France, France and Germany, and Germany and Heligoland. We find that for all population pairs, the best DFEs are those that infer a high correlation between mutational fitness effects in the two populations.

974 - Role of linc01089 and linc01140 in Transcriptional Stress Response in Myelogenous Leukemia Cells

Victor Ulises Plascencia Perez

Victor U Plascencia Perez, Mario E Zurita Ortega
Instituto de BiotecnologÃ-a - UNAM, Instituto de BiotecnologÃ-a - UNAM

The inhibition of transcriptional machinery by various small molecules has emerged as a promising strategy for targeting cancer cells, which exhibit heightened transcriptional activity. Triptolide, among these small molecules, is known to inhibit the pre-initiation complex component TFIIH, particularly its XPB subunit. Treatment of cancer cells with triptolide induces a state of stress, resulting in alterations in global gene expression. While many transcripts are down-regulated, others are upregulated, suggesting their potential involvement in stress response, and possibly contributing to the development of resistance in cancer cells. Notably, among the upregulated transcripts are several long non-coding RNAs (lncRNAs). It has been shown that lncRNAs can participate in key cellular processes by promoting changes in chromatin architecture, acting as enhancers, forming biomolecular condensates, etc.
In this study, we focus on the upregulation of the long intergenic non-coding RNAs (lincRNAs) linc01089 and linc01140 following triptolide treatment. Through a strategy involving heterochromatinization of the loci of these lincRNAs, we aim to elucidate their potential roles in the transcriptional stress response. Additionally, we are conducting experiments to investigate their cellular localization and their impact on cell cycle progression, apoptosis, and proliferation.
Furthermore, we plan to perform bioinformatic analyses to explore the expression patterns of linc01089 and linc01140 across various cancer types. Understanding the involvement of these lincRNAs in transcriptional stress response mechanisms may offer insights into novel therapeutic targets for myelogenous leukemia and other cancers.

975 - Uncovering the molecular mechanism of rapidly evolving species-specific fertilization in abalone

Stella Myra Lai

Stella M Lai, Damien B Wilburn

The Ohio State University, The Ohio State University

Fertilization, or the fusion of sperm and egg, is essential for most sexually reproducing organisms. The genes associated with reproduction undergo rapid evolution yielding tremendous diversity in reproductive proteins between taxa and closely related species. Despite more than a century of research, the molecular mechanisms by which sperm bind and dissolve the extracellular matrices surrounding the ovum remain unclear. Conserved across all animals, the egg coat is a species-specific barrier that restricts access to the ovum plasma membrane and limits polyspermy, which is lethal in most sexually reproducing species. As broadcast spawners that undergo external fertilization, abalone are a classic model for studying the molecular and evolutionary basis of fertilization where sperm lysin non-enzymatically dissolves the egg coat by binding its egg protein partner VERL. Here, we investigate lysin and VERL interactions in two abalone species (Haliotis rufescens and H. fulgens) to gain a mechanistic understanding of species-specific egg coat dissolution. VERL is a highly repetitive molecule with multiple tandem domains: the two most N-terminal domains (VN1, VN2) are distinct and rapidly evolving, while the remaining domains are homogenized and conserved. We found that species specificity of egg coat dissolution is mediated by weak but specific lysinâ€“VN1 interactions. However, mechanical dissolution of the egg coat requires tight but non-specific lysinâ€“VN3 binding, which induces a conformational change that destabilizes and restructures the egg coat. Structural modularity between species specificity and mechanistic dissolution within VERL may facilitate speciation of abalone by rapid sexual coevolution of gametic interactions.

976 - Evolutionary diversification of dimeric transcription factors by evolving DNA binding specificity

José Fabricio López Hernández

Jose Fabricio LÃ³pez HernÃ¡ndez, Brian P Metzger, Joseph W Thornton
University of Chicago, Purdue University, Deparment of Biological Sciences, University of Chicago. Deparment of Ecology and Evolution, University of Chicago. Deparment of Human Genetics

Variation in gene regulation contributes to the phenotypic diversity present in extant species. To regulate distinct gene targets, transcription factors (TFs) bind to specific DNA sequences. Estrogen receptors (ERs) and ketosteroid receptors (kSRs), critical for sexual development and physiology, regulate transcription by forming homodimers. Each subunit binds to one of two response elements (RE) using their DNA binding domain (DBDs). These receptors shared a common ancestor (AncSR1) before a gene duplication event. While ERs retained the ancestral profile binding to DNA, the ancestor of kSRs (AncSR2) recognized a different RE by acquiring substitution in the DNA recognition helix, the homodimer interface, and the rest of the DBD. Using an evolutionary approach that measures the effect of each substitution and their interactions will help us identify what changes were important during their functional transition.

To answer this question, we are conducting a comprehensive study to quantify the effect of each possible substitution and their combinations, between AncSR1 and AncSR2. To understand the genetic architecture of the functional transition, we are using deep mutational scanning (DMS), to infer the effect of all combinations of substitutions between AncSR1 and AncSR2. We assay the function of mutants using cell sorting, by GFP expression, coupled with high-throughput sequencing of each sorted bin, to measure the ability of DBD mutants to induce transcription GFP activation using distinct REs. We will infer the effect of each substitution and their interactions to model the possible evolutionary paths during the functional transition of AncSR1 to AncSR2.

1023 - The fitness landscape of an ancient signaling pathway

José Aguilar-Rodríguez

JosÃ© Aguilar-RodrÃ-guez, Daniel F. Jarosz, Dmitri A. Petrov

Stanford University, Stanford University School of Medicine, Stanford University

Recent advances in high-throughput technologies are bringing to life the metaphor of the adaptive landscape at unprecedented resolution. However, a limitation of most studies of fitness landscapes is their focus on the coding region of single biomolecules. Here we move beyond single-gene approaches by probing the landscapes of complex biological pathways at large scale. We use CRISPEY-BAR, which permits both high-throughput precision genome editing and high-resolution barcode-tracking of the edits in cell populations, to map the fitness of tens of thousands of single-nucleotide variants in the Ras/PKA and TOR/Sch9 pathways which control cell growth in all eukaryotes. Across four environments, we measured fitness of both coding and non-coding variants segregating in more than a thousand wild and domesticated yeast strains to understand the properties of standing genetic variation in these pathways. While we find that most variants behave neutrally, hundreds show significant positive and negative fitness effects. These variants tend to be missense and affect protein stability or conserved regions. We also find that rare variants and variants at low frequency tend to have a higher fitness impact than more common variants. Finally, we uncover numerous variants that change fitness across environments, which suggests that local adaptation may be common since they tend to be associated to phylogenetically related strains occupying particular ecological niches. These findings illustrate how massively parallel genome editing can be used effectively to interrogate the functional properties of natural variation in key biological pathways.

1032 - The global modifier Hsp90 shapes adaptation by controlling the fitness consequences of regulatory variation

José Aguilar-Rodríguez

Christopher M. Jakobson, JosÃ© Aguilar-RodrÃ-guez, Daniel F. Jarosz
Stanford University School of Medicine, Stanford University, Stanford University School of Medicine

Global modifiers influence the mapping of genotypes onto phenotypes and fitness through their epistatic interactions with genetic variants on a massive scale. The canonical example of a global modifier is the essential stress-responsive chaperone Hsp90 that impacts development and adaptation from microbes to humans, yet the molecular mechanisms by which Hsp90 alters the consequences of mutations remain unclear. We exploited the power of nucleotide-resolution genetic mapping in Saccharomyces cerevisiae to uncover more than a thousand natural variant-to-phenotype associations governed by this molecular chaperone. Strikingly, Hsp90 more frequently modified the phenotypic effects of cis-regulatory variation than variants that altered protein sequence. Nearly all interacting variants fell within clients of Hsp90 or targets of its direct binding partners showing how the effect of a global modifier can percolate throughout cellular networks. Hsp90â€™s influence was especially potent on evolutionarily young genes, highlighting its influence on variation central to evolutionary novelty. Reconciling the diverse epistatic effects of this chaperone, synthetic transcriptional regulation, and reconstructions of natural alleles by genome editing revealed a central role for Hsp90 in regulating the fundamental relationship between molecular activity and fitness for many dosage-sensitive genes. Our findings establish that non-coding variation is a core driver of Hsp90's influence on heredity, offering a mechanistic explanation for the chaperone's strong effects on evolution and development across species. They illustrate how global modifiers can pervasively change the adaptive value of thousands of variants segregating in natural populations.

1034 - The Genetic History of Himalaya

Mark A Penjueli

Mark A Penjueli, Rita Fahmy, Eaaswarkhanth Muthukrishnan, Yoshina Gautam, Dinesh Bhandari, Sarmila Tandukar, Shirley Sutton, Kaitlyn Spees, Genevieve L Wojcik, Jeevan B Sherchand, Carlos D Bustamante, Alexander G Ioannidis, Aashish R Jha

New York University Abu Dhabi, New York University Abu Dhabi, New York University Abu Dhabi, New York University Abu Dhabi, Public Health Research Laboratory, Institute of Medicine, Public Health Research Laboratory, Institute of Medicine, Department of Biomedical Data Science, Stanford University, Department of Biomedical Data Science, Stanford University, Department of Biomedical Data Science, Stanford University, Public Health Research Laboratory, Institute of Medicine, Department of Biomedical Data Science, Stanford University, Department of Biomedical Data Science, Stanford University, University of California Santa Cruz, New York University Abu Dhabi, Center for Genomics and Systems Biology, New York University Abu Dhabi, Public Health Research Center, New York University Abu Dhabi

Sandwiched between Eurasia and the Indian sub-continent, the Himalayan region harbors a rich and complex demographic history. In this study, we investigate patterns of relatedness and population substructure using a comprehensive dataset of 1,200 individuals from 66 Himalayan populations, global references, and more than 400,000 genome-wide SNP markers. Results from global ancestry methods reveal that indigenous Himalayan populations share East Asian (EAS) and South Asian (SAS) ancestries. Local ancestry-based analyses reveal that the EAS component is shared between Trans-Himalayan language speakers, including Tibetans, and differs from those in mainland East and Southeast Asia. Interestingly, the SAS component in Himalaya is closest to the Tribal Indian populations of Dravidian-like ancestry, indicating that the ancestral Dravidian homeland extended to the foothills of the Himalaya prior to the Indo-European expansion. Admixture dating reveals that the EAS and SAS components admixed within the last 2,000 years, which is consistent with anthropological records. Genome-wide scans for positive selection revealed strong signals in immunity-related genes that replicated in multiple Himalayan populations indigenous to the malaria endemic Terai plains (altitude <400m). We also identified signals of positive selection on height-related genes in two indigenous Tibeto-Burman populations of the Middle Hills (1000-2500m). SNPs with strongest signals of selection are nearly fixed in both Himalayan populations for the allele associated with shortness in previous GWAS. However, very few of these candidate SNPs or genes overlap between the two populations, suggesting convergent evolution of shorter stature in Himalaya.

1040 - Genomic Insights into Diving Adaptations of Extreme Pinniped Divers: A Comparative Study of Elephant and Weddell Seals

Diana D Moreno-Santillan

Diana D Moreno-Santillan
University of California Berkeley

Extreme pinniped divers are extraordinary models to study adaptations to hypoxia/reoxygenation events derived from extreme fluctuations in oxygen levels during deep diving. We analyzed the evolutionary adaptations of the Weddell seal (Leptonychotes weddellii) and the Northern and Southern elephant seals (Mirounga angustirostris and M. leonina), which have exceptional diving abilities, reaching depths of over 1000 meters for extended periods up to 90 minutes. To understand the molecular mechanisms underpinning these remarkable adaptations, we conducted a comprehensive comparative genomic analysis of thirteen pinniped species. Our gene family evolution analyses revealed a significant loss of function in genes involved in the enzymatic degradation of glutathione, an endogenous antioxidant that our functional analysis have previously shown to increase rapidly in response to hypoxia exposure in primary cells derived from extreme divers, likely to protect cells against oxidative damage.We also detected adaptive signatures in pathways involved in the optimization of mitochondrial respiration, energy metabolism, DNA repair and cardiovascular regulation in these extreme divers. Notably the Weddell seal exhibited more genes under positive selection compared to the Elephant seals, possibly reflecting adaptations to extreme environmental pressures. Pathways enriched in Weddell seals include those related to fat metabolism, oxidative damage against hydrogen peroxide and membrane lipid remodeling.
Our analysis sheds light on the remarkable genomic adaptations that enable extreme pinniped divers to tolerate oxygen fluctuations

associated with deep diving while offering valuable insights into their unique physiological and biochemical strategies for overcoming the challenges posed by hypoxia and reperfusion damage.

1068 - Doing Less with More: Immune Modulation in Native and Non Native House Sparrows

Kailey Marie McCain

Kailey M McCain, Lynn B Martin
University of South Florida , , University of South Florida

Immune modulation is crucial for invasive species' success and house sparrows (Passer domesticus) exemplify this. Colonization likely relies on inflammation control, with Toll-like receptors (TLRs) playing a key role in pathogen recognition. We found elevated TLR expression (TLR2 and TLR4) in invading sparrows; however, excessive expression of TLRs may lead to an overly robust inflammatory response.

We hypothesize that non-native house sparrows will exhibit an elevated baseline TLR expression, while concurrently relying on the co-expression of anti-inflammatory cytokines (IL-10) to attenuate downstream inflammatory pathways and non-native house sparrows will hasten the resolution phase of their inflammatory response.

To examine the differences in immune modulation, we captured house sparrows and measured TLR2, TLR4, IFN

1094 - Fitness Consequences in Artificial Selection for Yeast Cell Sizes

Wei-Chin Ho

Wei-Chin Ho, Kara Schmidlin, Zerrin Uzum, Adam Kindelin, Alexander Sastokas, Michael Lynch, Kerry Geiler-Samerotte
University of Texas at Tyler, Arizona State University, Arizona State University, Arizona State University, Arizona State University, Arizona State University, Arizona State University, Arizona State University

Traits are not independently evolved and are often constrained by each other. For example, during the artificial selection for desired traits in agriculture, the evolution of reduced survivorship or reproduction is frequently observed. While cases of constrained evolution have been observed, the underlying mechanisms are usually unclear. While the anticorrelation of many traits can be explained by first principles regarding physics or chemistry, the underlying genetic architecture may also play an unequivocal role, if the mutations contributing to the artificial selection lead to the fitness reduction through another trait. To study these questions, the artificial selection for larger cell sizes in budding yeast, Saccharomyces cerevisiae, was used as an efficient way of forward genetic screening. In particular, we used flow cytometry to sort cells with top 10% sizes as the inoculum for the next growth cycle every two days. Throughout the artificial selection for about 20 cycles, we tracked the evolutionary dynamics of cell sizes. The growth curves of evolved populations were also recorded to study their correlation with cell size evolution. Using the lineage-tracking technique with DNA barcodes, we will also dissect the genetic composition of the evolved population to identify the genetic basis of cell size variations and test whether or not the mutations affecting cell sizes simultaneously affect the growth and survival.

1102 - Parallel evolution in diverse strains of Pseudomonas aeruginosa

Keaton Sinclair

Keaton Sinclair, Alexandra Hicks, Rees Kassen

University of Ottawa, University of Ottawa, McGill University, University of Ottawa

There are many paths an organism can take in the process of adaptation. But as in mountain climbing, as they approach a peak, fewer directions will take them upwards. Knowing the likelihood that organisms repeat the same adaptive changes informs our expectations of how they may change in the future, which can be important in tracking the mutations of an infectious pathogen.
The gram-negative bacteria Pseudomonas aeruginosa are opportunistic pathogens which can infect and cause mortality in humans with compromised immune systems. The species is extremely widespread and thrives in moist environments from soil to living tissue, leading to a large diversity of locally adapted strains.
This study sought to determine the extent of parallel evolution by analyzing the adaptation of eight strains of Pseudomonas aeruginosa, taken from three broad sources: lab reference (PA14 & PAO1), environmental (Pae111, Jp54, & Jp1140), and clinical (JD313, JD322, JD328). Each strain was propagated for 750 generations in minimal growth media to simulate scarcity-based selection pressure; the fitness and genome of the 750th generation isolates were then compared to those of the starting generation. We hypothesized that parallel evolution would be inversely correlated with both the change in fitness and the genetic distance between the strains.
We found that parallel evolution was weakly inversely correlated with fitness change and genetic distance, varying in intensity between different strains. Shared genes were mutated in similar ways across the independent isolates of different strains, suggesting that they may be part of a common adaptive pathway.

1104 - Genetic Ancestry and the Urinary Metabolome of Extremely Premature Infants in the Neonatal Intensive Care Unit.

Thaybeth I Malavé-Méndez

Thaybeth I MalavÃ©-MÃ©ndez, Miguel Guardado, Martina Steurer, Cheryl Chapin, Ryan Hernandez, Phil Ballard, Dara Torgerson
University of California San Francisco, University of California San Francisco, University of California San Francisco, University of California San Francisco, University of California San Francisco, University of California San Francisco, University of California San Francisco

The human biofluid metabolome is influenced by environmental and genetic factors, including genetic ancestry-linked variation shaped by varying continental demographic histories. How these factors contribute to metabolic processes under extreme environmental exposures of prematurity remains unexplored. We aimed to identify urinary metabolites and pathways that vary with continental genetic ancestry in genetically admixed, extremely premature infants in the neonatal intensive care unit. Our study included 171 infants <29 weeks gestational age including individuals with maternal self-reported Hispanic White, non-Hispanic White, and Black/African American race/ethnicity. Untargeted metabolomics was performed using UHPLC-MS/MS on urines collected at two timepoints between days 6-14 and 23-30 postnatal age. Three-way continental genetic ancestry proportions were inferred using a reference panel including African, European, and Amerindigenous ancestry individuals. We tested for associations between proportions of genetic ancestry and quantitative measures of ~1,000 individual metabolites among all infants combined, and infants stratified by social categories of maternal race/ethnicity. At timepoint 1, we identified 119, 72, and 58 metabolites associated with African, European, and Amerindigenous genetic ancestry proportions, respectively, and 64, 33, and 39 metabolites, respectively, at time 2. Ten metabolites were associated with the same ancestry in the same direction at both timepoints, and several metabolites showed consistent effects across racial/ethnic social identities. Our study suggests that genetic variation shaped by varying continental demographic histories has an impact on metabolite variation of extremely premature infants, and highlights the importance of considering genetic ancestry in clinical studies of diverse populations.

1108 - Genetic Variation and Evolution of Gonad-Restricted Chromosomes in Sea Lampreys

Ilya Kisselev

Ilya Kisselev, Sara V. Good

Unversity of Winnipeg, Univeristy of Winnipeg, Univeristy of Manitoba

Sex determination (SD) and differentiation processes in sea lampreys present a complex interplay of genetic and environmental factors. Recent studies have shown that genes on highly repetitive small gonad-restricted chromosomes (GRCs) exhibit sex-biased expression in animals undergoing sexual differentiation, suggesting specialized role in SD similar to sex chromosomes in other species. The objective of this research project is to analyse structural and genetic variation of GRCs and their evolutionary patterns in sea lampreys.

We analyzed population-level data to compare genetic differentiation between anadromous and land-locked lamprey populations, focusing on both somatic and GRC genome parts. We investigated the mode of inheritance by examining sex-specific heterozygosity in somatic chromosomes versus GRCs and used long-read data to compare structural variations in the GRC between male and female genomes. Additionally, we compared gene evolution rates on GRCs and somatic chromosomes by calculating dn/ds ratios.

Our findings showed minimal differentiation between anadromous and land-locked populations in the GRC region and ten-fold higher Fst in the somatic region across sexes. Interestingly, at the macroevolutionary level, GRC genes evolve faster than those on somatic chromosomes, as shown by higher dn/ds values. Males exhibited two-fold higher GRC heterozygosity compared to females, and we identified 2 Mb of sequences exclusive to one sex, suggesting significant sex-specific structural variation.

Overall, our findings indicate distinct molecular and evolutionary properties of GRCs in sea lamprey, and imply the role of this genomic region in the male-specific sex determination and differentiation pathways.

1114 - The manifestation of hybrid sterility requires a greater number of sterility genes in heterozygous introgressing segments than in homozygous introgressing segments.

Ezel Jacome Galindo-Pérez

Ezel J. Galindo-PÃ©rez, Blanca E. ChÃ¡vez-Sandoval, Lilian M. Barandica-CaÃ±on, Francisco GarcÃ-a-Franco

Departamento El Hombre y su Ambiente. Universidad AutÃ³noma Metropolitana. Unidad Xochimilco., Departamento de Zoologia, de la Escuela Nacional de Ciencias BiolÃ³gicas. Instituto PolitÃ©cnico Nacional., Departamento de Ciencias BÃ¡sicas, Universidad de Cartagena., Departamento El Hombre y su Ambiente. Universidad AutÃ³noma Metropolitana. Unidad Xochimilco.

In this bioinformatic study we identified all the orthologous genes of D. buzzatti, D. mojavensis and D. melanogaster present in six homozygous sterility-producing segments, of different sizes and distributed along chromosome 4 of D. buzzatii.

By their association with the respective mutant sterility alleles, each of these genes was classified into one of three categories: female sterility genes, male sterility genes, and both sex sterility genes.

Surprisingly, we identified a greater number of female sterility genes and a smaller number of male sterility genes and both sexes sterility genes. All of them distributed heterogeneously along chromosome 4.

We conclude that: there are twice as many major oogenesis genes as there are major spermatogenesis genes, heterogeneously distributed across all autosomes; In each introgressive segment, between 3 to 11 spermatogenesis genes and at least 20 oogenesis genes are required to be placed in a homozygous recessive state to produce the respective sterility of each sex; each female sterility segment always includes the minimum number of male sterility genes sufficient to produce sterility in both sexes; To produce male sterility, at least 45 male sterility genes must be put into recessive heterozygosity (30% heterozygous segments, 15 times larger than the 2.21% homozygous segments); The large number of major oogenesis genes allows for a very robust oogenesis process that practically

prevents the manifestation of female sterility and/or lethality by introgresant heterozygous segments.

1119 - Obtaining chromosomal introgressions in a homozygous state reveals the genetic architecture of reproductive isolation barriers

Blanca Estela Chavez-Sandoval

Blanca E Chavez-SAndoval, Ezel J. Galindo-pÃ©rez, Lilian M Barandica, Francisco GarcÃ-a-Franco

Instituto PolitÃ©cnico Nacional, Universidad AutÃ³noma Metropolitana, Universidad de Cartagena, Universidad AutÃ³noma Metropoloitana

To contribute to the understanding of the mystery of mysteries, how species arise, for more than 100 years hybrid crossings between individuals of different sister species in various taxa have been the fundamental strategy of various evolutionary studies. Always with the objective of identifying speciation genes and their processes in the genetic architecture of reproductive isolation.

In this work, through a new strategy of interspecific crossings between individuals of two sister species of Drosophila, D. buzzatii and D. koepferae, for the first time, it was possible to obtain introgressive segments in a homozygous state.

Heterozygous introgressive segments only produce male sterility, while homozygous segments produce male sterility and female sterility.

Homozygous male sterility segments are smaller, homozygous female sterility segments are larger and always include at least a small segment of male sterility, producing sterility of both sexes.

We conclude that the architecture of sterility segments: masks and underestimates the number of female sterility segments and genes; both the male sterility segments and the female sterility segments require a particular minimum size to produce sterility, consequently with a particular minimum number of sterility genes to produce the respective sterility of each sex; The hybrid sterility architecture model corresponds to heterozygous and homozygous additive recessive mutations for the heterozygous and homozygous segments respectively, which produce partial reproductive isolation of the introgressed hybrid offspring.

1123 - Capacity of the skin microbiota of two critically endangered species of Ambystoma to inhibit the growth of the fungus Batrachochytrium dendrobatidis

Fabiola Itzel Loeza-Torrero

Fabiola Itzel Loeza-Torrero, Eria Rebollar-Caudillo, Yunuen Torres-Tapia, Ireri Suazo-OrtuÃ±o, Yurixho Maldonado-Lopez

Instituto de Investigaciones sobre los Recursos Naturales, Universidad Michoacana de San Nicolas de Hidalgo, Centro de Ciencias GenÃ³micas Universidad Nacional AutÃ³noma de MÃ©xico, Campus Morelos, Escuela Nacional de Estudios Superiores Unidad Morelia, UNAM, Instituto de Investigaciones sobre los Recursos Naturales, Universidad Michoacana de San Nicolas de Hidalgo, Instituto de Investigaciones sobre los Recursos Naturales, Universidad Michoacana de San Nicolas de Hidalgo

Ambystoma dumerilii and Ambystoma andersoni are microendemic salamanders from MichoacÃ¡n, classified as critically endangered species due to threats as emerging diseases such as Chytridiomycosis, caused by the fungus Batrachochytrium dendrobatidis, which has affected 501 amphibian species worldwide. However, amphibians are associated to skin bacteria with the potential to inhibit pathogens, through the capacity of biofilm formation and swarming mobility, that allows them to colonize and protect individuals. In our study, we analyzed the skin bacteria of Ambystoma dumerilii and Ambystoma andersoni and the capacity to inhibit Batrachochytrium dendrobatidis, biofilm formation and swarming movility. Sixteen samples of A. dumerilii and 30 of A. andersoni were cultured and we obtained 101 bacterial strains from A. andersoni and 47 from A. dumerilii that were identified molecularly and morphologically. We found bacteria species such as Micrococcus luteus, Bacillus velezensis, Streptomyces sp. that have shown previously, the capacity to inhibit the growth of Batrachochytrium dendrobatidis. Thirty strains of both Ambystoma species were able to form biofilm and 20 strains of A. dumerilli and

15 of A. andersoni presented swarm mobility. The inhibition tests show that some strains present anti-Batrachochytrium dendrobatidis inhibition, and in some cases, bacterial strains favor the growth of this pathogen. Our results suggest that the culturable bacteria of Ambystoma species exhibit important capacities that allows bacteria to colonize new spaces and provide protection against pathogens.

1132 - Comparative population genomics reveals convergent signatures of adaptation in avian brood parasites.

Ekaterina Osipova

Ekaterina Osipova, Christopher Balakrishnan, Jeffrey DaCosta, Wesley Warren, Michael Sorenson, Timothy Sackton
Harvard University, East Carolina University, Boston College, University of Missouri, Boston Uuniversity, Harvard University

Parental care is a widespread behavioral strategy in the animal kingdom that plays a crucial role in the offspring survival. Nevertheless, seven bird groups â€“ brood parasites â€“ have independently evolved to exploit other species to provide parental care to raise their young. The evolution of obligate brood parasitism is a radical life history transition, which is currently poorly understood at the genetic level. To study genetic patterns of recent adaptation associated with avian brood parasitism, we have generated population-scale resequencing data (n = 11-31 individuals per species) from five species of obligate brood parasites along with their nesting outgroups. These species include the cuckoo finches and wydahs, the greater honeyguide, and the brown-headed cowbird, which together represent three independent origins of brood parasitism. Using McDonaldâ€“Kreitman framework, we find evidence for repeated adaptation in genes involved in spermatogenesis and sperm function in multiple parasitic clades, but not in their sister, nesting clades. This is consistent with evidence for increased male-male competition in parasitic lineages as a result of allocating resources from parental investment to reproduction. In addition, we find evidence of recent selective sweeps near genes associated with nervous system development in parasitic lineages, potentially contributing to the increased spatial navigation skills in brood parasites necessary to locate and surveil hostsâ€™ nests.

1138 - A highly contiguous reference genome for the endangered Texas ocelot (Leopardus pardalis) for facilitating genomically-informed population management and recovery.

James Nicholas Neugebauer

James N Neugebauer, William J Murphy, Michael E Tewes
Caesar Kleberg Wildlife Research Institute, Texas A&M University â€“ Kingsville, Veterinary Integrative Biosciences, Texas A&M University, Interdisciplinary Program in Genetics & Genomics, Texas A&M University, Caesar Kleberg Wildlife Research Institute, Texas A&M University â€“ Kingsville

The ocelot (Leopardus pardalis) is a neotropical felid that ranges from Argentina to the United States. The species is listed as Endangered at the federal level in the United States, the northernmost region of its current range. In Texas, USA, populations are composed of fewer than 100 individuals that are split across two distinct sites in the southernmost region of the state. These populations are at risk due to high amounts of habitat fragmentation and urbanization in the surrounding landscape, resulting in vehicle mortalities and limited opportunity for dispersal thus decreasing gene flow. Forthcoming recovery efforts necessitate the exploration of the current genomic health of the populations and monitoring through time. To make inferences on population-level differences in genomic structure and identify regional adaptations, a high-quality reference genome is required. Additionally, only one other assembled genome has been published from the Leopardus genus, which includes most wild felid species in the Americas. To meet these needs, we present the preliminary results from a long-read PacBio Hi-Fi assembly and a strategy to assess genetic variation across the Texas population. This work will be beneficial in future decision-making regarding ocelot breeding and recovery as well as contribute to the broader scientific

knowledge of genomics in small felid populations and the Leopardus genus.

1139 - Genomic Offset and Climate Adaptations in Grouse: Relationships with Demographic History

Theodore E Squires

Theodore E Squires, Patrik RÃ¶din-MÃ¶rch, Jacob HÃ¶glund, Kristinn P MagnÃºsson
University of Akureyri, Uppsala University, Uppsala University, Uppsala University, University of Akureyri, Icelandic Institute of Natural History

Using whole genome re-sequencing data we study the effects of climate influenced declines in effective population size on the accumulation of deleterious mutations and the response to future climate change in populations of cold-adapted avian sister species from the Holarctic: rock ptarmigan (Lagopus muta) and willow ptarmigan (Lagopus lagopus). We reconstruct the demographic histories of the populations and determine their nucleotide diversity, past and present inbreeding, and mutation load. Genomic vulnerability to future climate change scenarios (also known as offset) is predicted for the populations. We show that relatively small and isolated populations have reduced nucleotide diversity, higher signatures of past and present inbreeding, and higher estimates of mutation load. Among the studied populations, the most vulnerable to a mismatch between current and predicted future environments are rock ptarmigan populations in East Greenland, Iceland, and Svalbard, while among willow ptarmigan, subspecies residing on the British Isles are the most vulnerable.

1140 - Comparative annotation of Drosophila genomes

Pankaj Dhakad

Pankaj, Bernard, Darren Dhakad, Kim, Obbard
University of Edinburgh, Stanford University, University of Edinburgh

Accurate prediction of genomic features lays the foundation for evolutionary analyses. Annotation methods typically fall into two categories: ab-initio approaches, which predict gene structures using statistical approaches such as Hidden Markov Models; and sequence alignment-based methods, which employ sequence alignment of known RNAs, cDNA, and proteins to discover transcripts.
In this study, we employed the Comparative Annotation Toolkit (CAT) pipeline, supplemented with Braker3, to annotate 304 Drosophila genomes. CAT offers a comprehensive approach, integrating transcript projection, transcriptome, and proteome alignments, along with simultaneous gene-finding, CAT provides a flexible way to simultaneously annotate entire clades and identify orthology relationships. Nevertheless, CAT's efficacy in identifying species-specific genes appears limited in the absence of RNAseq data or high-quality reference annotations. To address this limitation, we complemented CAT annotations with those generated by braker3, leading to substantial improvements in annotations.
To assess annotation quality across these Drosophila species, we employed phylogenetic mixed models to infer phylogenetic effects on gene number/length, and effect of availability of RNAseq and distance from a reference species on gene number/length. Finally, leveraging the comprehensive annotation data, we assessed the codon usage bias patterns in the genus Drosophila.
This study demonstrates the importance of advanced annotation pipelines for accurate genome annotation. The annotation dataset aids in understanding gene structures and species-specific features, while also facilitating research into evolutionary phenomena like codon usage bias (CUB) in Drosophila.

1141 - Uncovering novel Bacterial and Archaeal diversity: genomic insights from metagenome-assembled genomes in Cuatro Cienegas, Coahuila

Ulises Erick Rodriguez-Cruz

Luis E Eguiarte, Luis E Eguiarte, Valeria Souza
Instituto de Ecologia, Instituto de Ecologia, Instituto de Ecologia

A comprehensive study was conducted in the Cuatro CiÃ©negas Basin (CCB) in Coahuila, Mexico, known for its remarkable microbiological diversity and unique physicochemical properties. Within the CCB, the "Archaean Domes" (AD) site is remarkable due to the abundance of hypersaline, non-lithifying microbial mats. This study focused on analyzing the small domes and circular structures formed in AD by metagenome assembly genomes (MAGs) with the aim of expanding our understanding of the prokaryotic tree of life by uncovering previously unreported lineages, as well as analyzing the diversity of bacteria and archaea in the CCB. A total of 325 MAGs were identified, including 48 Archaea and 277 Bacteria. Remarkably, 22 archaea and 104 bacteria could not be classified at the genus level, highlighting the remarkable novel diversity of the CCB. Besides, AD site showed considerable diversity at the phylum level, with Proteobacteria being the most abundant, followed by Desulfobacteria, Spirochaetes, Bacteroidetes, Nanoarchaeota, Halobacteriota, Cyanobacteria, Planctomycetota, Verrucomicrobiota, Actinomycetes and Chloroflexi. In Archaea, the monophyletic groups of MAGs belonged to the Archaeoglobi, Aenigmarchaeota, Candidate Nanoarchaeota, and Halobacteriota. Among Bacteria, monophyletic groups were also identified, including Spirochaetes, Proteobacteria, Planctomycetes, Actinobacteria, Verrucomicrobia, Bacteroidetes, Candidate Bipolaricaulota, Desulfobacteria, and Cyanobacteria. These monophyletic groups were possibly influenced by geographic isolation, as well as extreme environmental conditions in the pond DA, such as stoichiometric imbalance of C:N:P of 122:42:1, fluctuating pH (5-9.8) and high salinity (5.28%).

1142 - A simulation study to examine the impact of recombination onnphylogenomic inferences under the multispecies coalescentnmodel

Tianqi Zhu

Tianqi Zhu, Ziheng Yang
Academy of Mathematics and Systems Science, CAS, University College London

recombination within locus and free recombination among loci. Yet, in real data sets intralocus recombination causes different sites of the same locus to have different genealogical histories so that the model is misspecified. The impact of recombination on various coalescent-based phylogenomic analyses has not been systematically examined. Here, we conduct a computer simulation to examine the impact of recombination on several Bayesian analyses of multilocus sequence data, including species tree estimation, species delimitation (by Bayesian selection of delimitation models) and estimation of evolutionary parameters such as species divergence and introgression times, population sizes for modern and extinct species, and cross-species introgression probabilities. We found that recombination, at rates comparable to estimates from the human being, has little impact on coalescent-based species tree estimation, species delimitation and estimation of population parameters. At rates 10 times higher than the human rate, recombination may affect parameter estimation, causing positive biases in introgression times and ancestral population sizes, although species divergence times and cross-species introgression probabilities are estimated with little bias. Overall, the simulation suggests that phylogenomic inferences under the multispecies coalescent model are robust to realistic amounts of intralocus recombination.

1143 - Population genomics reveals complex genetic history of North Asian human populations

Vadim Stepanov

Vadim Stepanov, Vladimir Kharkov, Nikita Kolesnikov, Aleksey Zarubin

Institute for Medical Genetics, Tomsk National Research Medical Center, Institute for Medical Genetics, Tomsk National Research Medical Center, Institute for Medical Genetics, Tomsk National Research Medical Center, Institute for Medical Genetics, Tomsk National Research Medical Center, Institute for Medical Genetics, Tomsk National Research Medical Center

The whole genome research on genetic diversity in human populations is accumulating significant impact for medical genetics and personalized medicine. Objective of the study is to identify population structure and reveal signals of adaptive evolution in native populations of North Asia. We also aim to demonstrate the importance of incorporating population genetic diversity and ancestry information into present biomedical research.
Genome-wide analysis of genetic diversity using own WGS and SNPs array data was performed on a total sample over 2000 individuals representing 42 populations. PCA demonstrates general pattern of genetic diversity distributed according geographic locations of the populations. Major ancestral components can be distinguished as general Western- or Eastern Eurasian, and specific such as proto-Uralic, paleo-Asian, and Beringian. Uneven distribution of ROH among populations demonstrated founder effect in most North Asian populations. Analysis of genetic ancestry with ancient populations reveals significant contribution of Upper Paleolithic North Asian cultures to South and North Siberian people. Paleolithic European hunters-gatherers and Neolithic farmers contributed to modern native South Siberian and North European populations. We demonstrate the population-specific distributions of medically relevant variants in diverse human populations.
In conclusion, complex genetic history shaped by early migrations, isolation by distance, founder effects and natural selection has led to current genomic diversity patterns in human populations of North Asia. Future genetic studies and medical programs should take into account the genetic architecture of North Asian peoples.
The study was supported by the Russian Science Foundation grant No.22-64-00060.

1144 - Idiosyncratic evolutionary processes explain the distribution of genomic variation among co-distributed avian taxa in the Mexican tropical forests

Alexander Llanes-Quevedo

Alexander Llanes-Quevedo, Alicia Mastretta-Yanes, Luis Antonio SÃ¡nchez-GonzÃ¡lez, Adolfo Gerardo Navarro-Siguenza
UNAM, CONABIO, UNAM, UNAM

Avian diversity of Mexican tropical forests (MTF) has been strongly influenced by the presence of physical barriers to dispersal and the climatic cycles of the Pleistocene. While some studies suggest that differentiation patterns are associated to common geological and climate shifts, others propose more complex and species-specific responses to these changes; nevertheless, the understanding of processes generating these patterns is limited. Here, we analyzed genomic and ecological data for approaching spatial and temporal divergence patterns of three co-distributed bird species in the MTF: Saltator atriceps (Passeriformes: Thraupidae), Xiphorhynchus flavigaster (Passeriformes: Furnariidae) and Melanerpes santacruzi (Piciformes: Picidae). We performed population genomic analyses based on NextRAD SNP data to test if demographic dynamics and genetic differentiation are consequence of synchronic historical events or common physical barriers, as well as the influence of climatic variables in such differentiation. We found some spatial congruence in population structure of the species in the Tehuantepec Isthmus and wetlands of Tabasco. However, divergence times and populationsâ€™ dynamics are distinctive for each taxa. For M. santacruzi and S. atriceps, we obtained a significant correlation between genomic and environmental variation, while for X. flavigaster we inferred that genetic differentiation must be a consequence of the geographic distance and limited gene flow among populations. Our work provides evidence on the historical processes that have driven the evolution of MTF birds, and of the influence of the processes of local adaptation of the populations to different environmental conditions, which together have promoted the outstanding diversification of the Neotropical avifauna.

1146 - Environmental monitoring of zoonoses: its contribution to One-Health research.

Hugo Mendoza

Hugo Mendoza, Ana E Escalante

Instituto de EcologÃ-a (UNAM), Instituto de EcologÃ-a (UNAM)

Zoonoses are complex phenomena driven by the interaction of social-ecological processes and can be addressed through the inter/transdisciplinary One-Health approach, which considers that human and animal health are interdependent with environmental health. Although multiple zoonotic pathogens are environmentally transmitted and their identification has improved with the development of molecular analysis, monitoring has focused on humans and animals, leaving environmental monitoring lagging behind. Some authors have identified the lack of application of One-Health concept since environmental health has not been evaluated and inter/transdisciplinarity has not been conducted in research concerning One-Health. In this study, we analyzed the scientific literature specialized in molecular monitoring of zoonoses under the One-Health perspective to know the status of environmental monitoring, the degree of environmental health assessment, and the disciplinary character of the studies. We conducted a systematic search in Web of Science for scientific articles related to molecular monitoring of zoonoses that included One-Health in their title or abstract. We extracted data regarding the objects of monitoring (environment, humans, or animals), disciplines and sectors involved, and the type of health addressed (environmental, human, or animal), and analyzed them using Principal Component Analysis. We found that environmental monitoring of zoonoses is underrepresented and there is no suitable use of One-Health approach since there are no studies that evaluated environmental health. Likewise, the disciplines that dominated the studies were related to the biological sciences, and transdisciplinarity was addressed in only a few publications. The implications of our results are discussed.

1147 - Spatial population genetic models predict high genetic diversity loss in the Anthropocene even after habitat destruction ends

Moi Exposito-Alonso

Moi Exposito-Alonso, Kristy Mualim, Jeffrey Spence, Clemens Weiss, Oliver Selmoni, Meixi Lin

University of California Berkeley, Howard Hughes Medical Institute, Stanford University, Carnegie Institution for Science, University of California Berkeley, Stanford University, Carnegie Institution for Science, Stanford University, Stanford University, University of California Berkeley, Carnegie Institution for Science, University of California Berkeley, Carnegie Institution for Science

Anthropogenic habitat loss and climate change are reducing speciesâ€™ geographic ranges, increasing extinction risk and losses of speciesâ€™ genetic diversity. Although preserving genetic diversity is key to maintaining speciesâ€™ adaptability and is part of the post-2020 United Nationsâ€™ goals, we lack predictive tools and global estimates of genetic diversity loss across ecosystems. Bridging biodiversity theory and population genetics we built area-based predictive models with decreasing habitats. By using theoretical approaches and empirical data analyses of 10,095 georeferenced individuals from 20 plant and animal species, we estimate that more than 10% of genetic diversity may be lost in many threatened and non-threatened species, and an even larger fraction will be lost due to genetic drift even after habitat destruction stops. These results highlight the fundamental role of spatial population genetic models to meet global conservation targets at the genetic level.

1148 - Microbial Diversity Conservation and Characterization in the Yucatan Peninsula: Insights from Native Bee’s Honey and Water Sources in Calakmul Reserve.

Nelly . Sélem-Mojica

Nelly SÃ©lem-Mojica, Aurora Xolalpa-Aroche, David Alberto GarcÃ-a-Estrada, Haydee Contreras-Peruyero, Karina Verdel-Aranda, Enrique JesÃºs Delgado SuÃ¡rez
UNAM CCM, Universidad Intercultural Maya de QuintanaÂ Roo, „ CINVESTAV, UNAM CCM, Colegio de Posgraduados de Campeche, UNAM

The Yucatan Peninsula is part of the Mesoamerican Biological Corridor, where various conservation activities take place. Understanding our microbial diversity is crucial for its protection. This study examined microbial diversity in i) honey samples from native bee species Melipona beecheii and Scaptotrigona mexicana and ii) water sources within the Calakmul Biosphere Reserve.

Meliponiculture has deep roots in Yucatan, serving as an economic activity and cultural component. Notable is the honey of two native bee species: Melipona beecheii and Scaptotrigona mexicana. Honey possesses antibiotic properties, contrasting industrial sugar syrup. This project aims to valorize traditional Yucatan meliponiculture via microbial characterization of its honey. Metagenomes show Lactobacillus presence; one is significantly more abundant in M. beecheii honey and previously isolated from vinegar. No biosynthetic gene clusters were previously known. However, the honey-derived genome shows such antibiotic-producing gene clusters warranting further investigation.

Calakmul, a UNESCO World Heritage Biosphere Reserve, lacks permanent water bodies. Instead, seasonal rains fill "aguadas," temporary water formations vital for macrofauna, including endangered species. Samples were collected from three sites with varying protection levels during 2017-2019, establishing a baseline microbial diversity pre-Maya Train megaproject. Surprisingly, semi-urban samples showed higher alpha diversity than protected area samples.

1149 - Using archaeological samples to track the evolution of viral virulence

Steven R Fiddaman

Steven R Fiddaman, Laurent AF Frantz, Adrian L Smith, Evangelos Dimopoulos
University of Oxford, LMU, University of Oxford, Big Data Institute, Department of Veterinary Medicine, University of Cambridge

Pathogen genomes from archaeological animal remains help us to understand how pathogens have evolved over large spans of time. In this poster, I will demonstrate the power of ancient DNA analyses to understand the evolution of Mareká€™s Disease Virus (MDV), a highly contagious herpesvirus of chickens. Modern MDV causes widespread visceral tumours, immunosuppression and paralysis, and the chicken industry spends billions of dollars annually to control its spread. But it wasnâ€™t always this way. When Mareká€™s disease was first described in the 1900s, it was a comparatively mild disease, causing some limb paralysis with low mortality. Before this study, we lacked any MDV strains from prior to the 1960s, by which time MDV was much more virulent than at its first description. To investigate the genetic changes that underlie the virulence increase, we screened ~1000 archaeological chicken bones for traces of MDV. We recovered MDV genomes from 15 Eurasian samples from the past millennium. Using a genome-wide scan for positive selection, we identified a number of genes that were selected in the branch leading to modern MDV strains. Significantly, we found that the key driver of viral oncogenesis â€“ Meq â€“ had undergone adaptation between ancient and modern strains. Functional comparison between reconstructed ancient Meq and modern (virulent) Meq revealed that ancient Meq is a much weaker transcriptional regulator. These data provide strong evidence for the molecular changes that underpinned the evolution of modern hypervirulent strains of MDV.

1150 - IMPACT OF MECHANISM-VARYING ANTIBIOTICS ON HOW SPONTANEOUS MUTATIONS AFFECT E. COLI GROWTH

Ameya Kulkarni

Ameya Kulkarni, Michael Lynch, Wei-Chin Ho
Arizona State University, Arizona State University, University of Texas at Tyler

Antibiotics have, for a long period of time, been widely used to treat microbial infections. Therefore, antibiotics highly dictate microbial evolution. However, the interactions between antibiotics influence and the features of de novo mutations in microbes, which are the fundamental sources of evolution, remains largely misunderstood. The field of evolutionary microbiology, genetic and environmental interactions1 on mutational effects is an especially under-researched area. Specifically, interactions may exist between antibiotic exposure and the mutations, which may combinationally lead to non-additive changes in microbe growth rate. Due to this interaction, we hypothesized that antibiotic exposure will increase the detrimental effects of deleterious de novo mutations2. To test this hypothesis, the growth curves of 47 MMR-deficient Escherichia coli strains with 1800 days worth of de novo mutation accumulation3 are exposed to and measured with or without a sub-lethal dose of antibiotic of the ÃŸ-lactam, ampicillin and the aminoglycoside, streptomycin. Through the analysis of growth curves, the relative growth between wild-type and the antibiotic and between antibiotics is determined. This is then compared with the analogous data in the benign environment without antibiotics. These results can be used to link how mechanisms of genetics and environment interaction may operate, especially in comparing between similar yet mechanistically different environments. Additionally, this research helps in understanding the trends of microbial evolution under certain antibiotic treatments, and how they associate with corresponding mutations in the bacterial strains.

1151 - The Evolution Trends of Structural Variants at A Large Population Scale

Jen-Yu Wang

Jen-Yu Wang, J. J. Emerson
University of California-Irvine, , University of California-Irvine

Long-read sequencing simplifies genome assembly and enhances the discovery of structural variants (SVs). Employing Drosophila melanogaster as our model, we investigated SV evolution at a population level using over sixty whole-genome long-read sequencing datasets. Comparative analysis of contemporary SV-calling methods yielded population statistics, unfolded site frequency spectra and selection signals.

Traditionally, single nucleotide polymorphisms (SNPs) have been pivotal in inferring phylogeny, population evolution, and disease causation, owing to their detectability with cost-effective methods. Unlike SNPs, SVs, which often span and impact multiple genes, are more likely to influence an organism's phenotype. Despite their significance, SVs have eluded scrutiny due to their complex nature and the limitations of short reads from next-generation sequencing (NGS).

This study bridges the gap by leveraging state-of-the-art long-read sequencing to unveil SVs in Drosophila melanogaster populations worldwide. Utilizing Pacific Biosciences and Oxford Nanopore Technology sequences from NCBI SRA, we assessed two major SV detection approaches. Mapping-based programs, though computationally efficient, exhibited lower sensitivity, while assembly-based methods, dependent on well-constructed genomes, captured larger-scale changes. Our validated pipelines are open source on GitHub. Diversity and selection signals varied among SV types, and we sought SVs contributing to population differentiation and chromosomal evolution. SVs demonstrated signals of negative selection while a small fraction was probably positively selected. Additionally, we constructed the largest Drosophila pangenome for comprehensive representation of polymorphisms.

1152 - A geological timescale for bacterial evolution and oxygen adaptation

Gergely J Szöll?si

Gergely J SzÃ¶llÅ‘si, Adrian A Davin, Philip Hugenholtz, Tom A Williams
Okinawa Institute of Science and Technology Graduate University, Institute of Evolution, Centre for Ecological Research,, University of Tokyo, The University of Queensland, University of Bristol

Microbial life has dominated the biosphere throughout Earthâ€™s history but has left a meagre fossil record, greatly hindering our understanding of evolution in deep time. However, the co-evolution of life and the Earth system has left signatures of microbial metabolism in the geochemical record, most conspicuously the Great Oxidation Event (GOE) ~2.33 billion years ago (Ga, (Poulton et al. 2021)). The GOE transformed Earthâ€™s biosphere from dominantly anoxic to oxic via oxygenic photosynthesis and carbon burial via tectonism (Eguchi, Seales, and Dasgupta 2019). Here, we combine machine learning and phylogenetic reconciliation to infer ancestral transitions to aerobic lifestyles during bacterial evolution. Linking these transitions to the GOE provides new constraints to infer the timetree of Bacteria. We find that extant bacterial phyla are truly ancient, tracing their diversity back to the Archaean and the Proterozoic: the oldest include Bacillota (Firmicutes) that radiated 3.1-3.7 Ga, Cyanobacteriota (3.3-3.5 Ga) and Patescibacteria (3-3.5 Ga). We show that most bacterial phyla were ancestrally anaerobic and transitioned to aerobic lifestyles after the GOE. However, we infer that the earliest oxygen-adapted Bacteria pre-dated the GOE and facilitated the evolution of oxygenic photosynthesis in the cyanobacterial ancestor.

1154 - Conservation Genomics of North American Jaguars

Eldridge Wisely

Eldridge Wisely, Natalie Payne, Ivonne Cassaigne, Karla Vargas, Victor H Luja, Melanie Culver
University of Arizona, University of Arizona, Primero Conservation, Arizona State University, University of Arizona, Universidad AutÃ³noma de Nayarit, U.S. Geological Survey, University of Arizona

The jaguar (Panthera onca) was once found throughout the Americas. Now, jaguars are found in only a fraction of their former range. They suffer from habitat fragmentation, human-wildlife conflict, and illegal wildlife trafficking driving them toward extinction. Along with habitat fragmentation comes the potential for increased inbreeding and deterioration of genetic diversity - necessary for continued adaptation to a changing environment and climate. Jaguars living at the edge of their range are even more at risk because of reduced migration opportunities, but may also harbor unique adaptive and neutral genetic variation due to selection or drift.
Whole genome sequencing can aid management decisions for endangered species through increased resolution compared to traditional population genetics, which improves the ability to scan for local adaptations or alleles involved in reduced fitness. We used an individual-based sampling scheme to investigate the genomic characteristics of 14 free-ranging North American jaguars representing 7 Mexican states and Arizona over the last 30 years. We sequenced whole genomes of these jaguars from blood or tissue samples and mapped reads to the annotated tiger (Panthera tigris) genome. We inferred levels of connectivity and gene flow between the North American sub-populations represented by these jaguars, which reflects the level of landscape connectivity in North America, with respect to jaguar dispersal and migration. Additionally, we leveraged 30 years of temporal sampling to trace the levels of heterozygosity and inbreeding of these North American jaguars over time. Here we present our goals, methods, and preliminary findings.

1157 - The Codon Statistics Database: A Database of Codon Usage Bias

David Alvarez-Ponce

Krishnamurthy Subramanian, Bryan Payne, Felix Feyertag, David Alvarez-Ponce
University of Nevada, Reno, University of Nevada, Reno, University of Nevada, Reno, University of Nevada, Reno

We present the Codon Statistics Database, an online database that contains codon usage statistics for all the species with reference or representative genomes in RefSeq (over 15,000). The user can search for any species and access two sets of tables. One set lists, for each codon, the frequency, the Relative Synonymous Codon Usage, and whether the codon is preferred. Another set of tables lists, for each gene, its GC content, Effective Number of Codons, Codon Adaptation Index, and frequency of optimal codons. Equivalent tables can be accessed for (1) all nuclear genes, (2) nuclear genes encoding ribosomal proteins, (3) mitochondrial genes, and (4) chloroplast genes (if available in the relevant assembly). The user can also search for any taxonomic group (e.g., â€œprimatesâ€•) and obtain a table comparing all the species in the group. The database is free to access without registration at http://codonstatsdb.unr.edu.

1158 - Using saturation mutagenesis to map the evolutionary forces that shaped human regulatory elements

Omer Ronen

Omer Ronen, Simon Fishilevich, Yoav Navon, Georgii Bazykin, Tomas Marques-Bonet, Fumitaka Inoue, David Gokhman
Department of Molecular Genetics, Weizmann Institute of Science, Department of Molecular Genetics, Weizmann Institute of Science, Department of Molecular Genetics, Weizmann Institute of Science, Department of Biomedical Informatics, Harvard Medical School, Institute of Evolutionary Biology (UPF-CSIC), Institute for the Advanced Study of Human Biology (WPI-ASHBi), Kyoto University, Department of Molecular Genetics, Weizmann Institute of Science

Adaptation via natural selection is the process by which the incredible fit between species and their environment has evolved. However, despite the good understanding of evolutionary adaptation at the phenotypic level, we have a limited understanding of the genetic changes that made us human. Particularly, we are still far from understanding how adaptation presents itself at the molecular level. This task is even more difficult in the noncoding parts of the genome, where most adaptations occur.
Saturation mutagenesis is an approach that allows researchers to compare true evolutionary changes to the full spectrum of possible mutations. We apply it using massively parallel reporter assays (MPRAs), allowing us to generate the largest saturation mutagenesis library of regulatory regions in the human genome, consisting of 500 diverse regulatory elements. To identify signatures of selection, we leveraged and expanded classical evolutionary tests, such as the McDonald-Kreitman test, to apply them to noncoding regulatory elements. We also developed meta-analysis approaches to gain insight not only into specific regions, but also into classes of regulatory elements.
We identified signals of positive and negative selection in various elements, and discovered specific variants that have likely driven human-specific phenotypes, e,g., in IRF4, a key transcription factor that affects pigmentation and skin diseases.
This work allowed us to study how selection has shaped human noncoding regulatory elements, and to shed light on the role of key substitutions in human adaptations.

1159 - Molecular Characteristics of NTRK Fusion Genes

Andrea Zambrano Zambrano

Andrea Zambrano, Karen Granizo
Pontificia Universidad CatÃ³lica del Ecuador, CISeAL, -, Pontificia Universidad Catolica del Ecuador

NTRK fusions have aroused interest among oncologists since they are present in different cancer tumors. However, the complexity of

NTRK fusions due to their rarity and diverse fusion partners makes the diagnosis in medical settings challenging. Understanding the molecular characteristics, such as the common breakpoints, expected fusion partners, sites of fusions, and exon numbers, could simplify the detection, reducing costs and resources. The challenges persist in low- and middle-income countries where the lack of information, validation process, and costs impede the widespread implementation of molecular technologies. We obtained data from genome GRCh38.p5 and we acquired 20 features covering a range of structural, functional, regulatory, expression, and interaction from COSMIC, cBIO Portal, and FusionGBD. Databases were cleaned and processed using R script. A list of 118 fusion events detected in different types of cancer with different ages and sex. Fusion events were classified as recurrent and non-recurrent. Human genes were in GTF format and protein FASTA sequences.Gene fusions were prevalent in 43% of males and 57% of females. The types of cancer with more fusion events with NTRK are (Breast Invasive Carcinoma (n=41); Lung Adenocarcinoma (n=11), Salivary carcinoma (n= 26), and Papillary Thyroid Cancer (n= 67).NTRK fusions involve NTRK1, NTRK2, NTRK3 genes and a fusion partner. We identify the most prevalent fusion NTRK3::ETV6 and are in 3â€™ and the partner in 5â€™. NTRK3::ETV6 was the most prevalent in the population followed by TPM3::NTRK1

1160 - Pilot investigation of historical yak “crusties” using paleogenomics

Meghana Natesh

Natividad Lupianez Corpas, Meghana Natesh, Maanasa Raghavan
University of Chicago, University of Chicago, University of Chicago

Museum specimens offer an invaluable resource for studying species' biogeography and evolutionary histories. Coupling this resource with recent improvements in paleogenomic analyses has opened the doors to inferences on the past that were previously inaccessible, though biomolecular preservation and contextual gaps in museum records may limit the scope of such studies. Here, we present results from a pilot on â€œcrustiesâ€•, or dried bits of tissue, scraped from historical bones curated at the Field Museum. While crusties offer a minimally-destructive source of DNA from museum samples, they may not be as rich in endogenous DNA as bone or more intact tissues. Our pilot focuses on yak (Bos grunniens), a high-altitude adapted bovid species restricted today to the Tibetan plateau and surrounding regions. Wild yak (Bos mutus) populations persist on the Tibetan Plateau, though in dwindling numbers. Museum records for these yak samples are incomplete, noting the 1960s-1970s as the time of acquisition from the Chicago Zoological Society with no reference to the original sampling locations. Moreover, it is unclear whether the sampled yak are of a wild or domestic origin. We generate shotgun sequencing data from four individuals to evaluate DNA preservation and the feasibility of the generated data for establishing their genetic affinity with published wild and domestic yak across their range. Our pilot showcases the utility of paleogenomics on museum collections to supplement historical records in addition to offering a snapshot of genetic diversity in the recent past.

1161 - Towards Simulation Optimization: An Examination of the Impact of Scaling in Coalescent and Forward Simulations

Tessa Ferrari

Tessa Ferrari, Xinjun Zhang, Jazlyn Mooney
University of Southern California, University of Michigan, University of Southern California

Scaling is a common practice to increase computational efficiency. However, there exists a dearth of standardized guidelines for best practices. Few studies have examined the effects of scaling on diversity and whether the results are directly comparable to unscaled and empirical data. We examine the effects of scaling in two model populations, modern humans and Drosophila melanogaster. The reason is twofold: 1) human populations require moderate-to-no scaling for simulations while more dramatic scaling is required for Drosophila; and 2) model populations have empirical data for comparison. We determine whether estimates of diversity, the site frequency spectra,

deleterious variation, coalescence, run time, and memory are affected by scaling, the length of the simulated segment, and burn-in time. For humans we test both additive and recessive models of dominance and in Drosophila we test the additive model only. We find that the typical 10N generation burn-in is often not sufficient for full coalescence to occur in human or Drosophila simulations. As expected, memory and run-time increase as the scaling coefficient decreases and the length of the simulated segment increases. Simulating larger segments in humans is preferable. Diversity estimates have less noise, and concatenating these segments together to generate a genome is effective. Conversely, simulating smaller segments in Drosophila is preferable for achieving levels of diversity similar to empirical data; due to larger selective forces being employed from more aggressive scaling and causing background selection to occur at an increased rate.

1162 - Toward Understanding Salt Stress Response: Bioinformatic Analysis of Pseudocrossidium replicatum Transcriptome Reveals Divergence in Transcription Factors Across Species

Teresa Romero Luna

Teresa Romero Luna, Selma RÃ-os MelÃ©ndez, Mishael SÃ¡nchez PÃ©rez, Santiago ValentÃ-n GalvÃ¡n Gordillo, Analilia Arroyo Becerra, Miguel Angel Villalobos LÃ³pez

Centro de InvestigaciÃ³n en BiotecnologÃ-a Aplicada- Instituto PolitÃ©cnico Nacional, Centro de InvestigaciÃ³n en BiotecnologÃ-a Aplicada- Instituto PolitÃ©cnico Nacional, Centro de Ciencias GenÃ³micas, Universidad Nacional AutÃ³noma de MÃ©xico, Centro de InvestigaciÃ³n en BiotecnologÃ-a Aplicada- Instituto PolitÃ©cnico Nacional, Centro de InvestigaciÃ³n en BiotecnologÃ-a Aplicada- Instituto PolitÃ©cnico Nacional, Centro de InvestigaciÃ³n en BiotecnologÃ-a Aplicada- Instituto PolitÃ©cnico Nacional

Plants, being sessile organisms, are exposed to abiotic stress (cold, drought, flooding, salinity). However, they have evolved over time, acquiring adaptive strategies to tolerate stress. Mosses have been the first plants to adapt and populate the land. Pseudocrossidium replicatum is a moss classified by our group as totally tolerant to desiccation. This moss can also recover from 10 days of exposure to high salinity (700 mM NaCl) or ultra-freezing conditions (-80oC). Our group is interested in identifying the key genes responsible for abiotic stress tolerance in this moss. In the present work, we performed a transcriptomic and bioinformatic analysis of the protonemal tissue exposed to 200 mM NaCl for 3 hours. Nine NaCl-responsive transcription factors (TFs) were found. Three TFs belong to the family APETALA/ETHYLENE RESPONSIVE FACTOR (AP2/ERF). To make function comparisons between P. replicatum, other non-vascular vascular species and vascular model plants and algae, a phylogenetic analysis of these AP2 TFs (ERF3, DREB3, and DREB2A) was carried out. The three P. replicatum AP2 genes analysed conserved more similarity to the corresponding homologous from Physcomitrella patens, followed by genes from vascular plants and algae. The P. replicatum interactome revealed that it is also more related to the predicted interactome for P. patens; it diverges with respect to the functions in vascular plants, where these genes are mainly involved in regulating seed development. This analysis could infer information about the conservation of structures and functions throughout evolution and identify possible specific evolutionary novelties of P. replicatum.

1164 - Peruvian Population Genomics: Diversity, Admixture Dynamics, and the Impact of Indigenous American Ancestries on Complex Phenotypes

Timothy D OConnor

Victor Borda, Omar Caceres, Cesar Sanchez, Carlos Padilla, Diego Veliz-Otani, Marla Mendes Aquino, Bradley Maron, Eduardo Tarazona-Santos, Heinner Guio, Timothy D OConnor

Institute for Genome Science, University of Maryland School of Medicine, Institute for Health Computing, University of Maryland School of Medicine, Laboratorio de InnovaciÃ³n y Desarrollo, Instituto Nacional de Salud, Laboratorio de InnovaciÃ³n y Desarrollo, Instituto Nacional

de Salud, Laboratorio de InnovaciÃ³n y Desarrollo, Instituto Nacional de Salud, Institute for Genome Science, University of Maryland School of Medicine, SickKids Institute, Institute for Health Computing, University of Maryland School of Medicine, Pulmonary Hypertension Center University of Maryland Medical Center, Instituto de Ciencias Biologicas, Universidade Federal de Minas Gerais, Universidad de Hunuco, INBIOMEDIC, Institute for Genome Science, University of Maryland School of Medicine, Program in Health Equity and Population Health

Latin American populations have a recent history of extensive migration and admixture that shaped their genetic diversity. Particularly, admixed Peruvians are recognized as having predominantly Indigenous American ancestry. Phenotypically, Peruvians tend to be shorter and, as in many low and middle-income countries, Peru has shown an increasing rate of obesity. We generated genome-wide data for 336 admixed individuals combined with 515 previously described samples representing 13 regions to explore population structure, demographic history, and the relationship between Indigenous American ancestries and variation with complex phenotypes (i.e., height and waist-to-hip ratio). We identified that Indigenous American ancestries in admixed Peruvians mirror the north-south interaction known from archeology and linguistics previously described for Peruvian Natives. We detected the most recent admixture events between European and Native Peruvians dating to 8-10 generations ago which occurred with populations already admixed between distinct original populations. We observed differences between continental ancestry proportions between autosomes and chromosome X suggesting sex bias admixture with a predominant contribution of Indigenous American females. We found a correlation between small stature and a higher proportion of Indigenous American ancestry but no specific locus reached GWAS significance. We did find a significant association between locus 7p21.1 and the waist-to-hip ratio. Interestingly, the affected allele is common in African populations but rare in other 1000 Genomes populations. Our fine-scale population study provides subtle details of the genetic history and diversity, serving as one more step to understanding the broader human genetic variation landscape.

1165 - Resolving the regulatory circuitry driving gall development in response to root-knot nematode herbivory in tomato

Tyler R Inskeep

Tyler R Inskeep, Damaris Godinez-Vidal, Simon C Groen
Department of Botany and Plant Sciences, University of California, Riverside, Center for Plant Cell Biology, Institute for Integrative Genome Biology, University of California, Riverside, Department of Nematology, University of California, Riverside, Center for Plant Cell Biology, Institute for Integrative Genome Biology, University of California, Riverside, Department of Nematology, University of California, Riverside, Department of Botany and Plant Sciences, University of California, Riverside, Center for Plant Cell Biology, Institute for Integrative Genome Biology, University of California, Riverside

Plants are parasitized by many species, one of the most ubiquitous being the root-knot nematode (RKN; Meloidogyne spp.). We believe that RKNs exert spatiotemporally variable selective pressure on their host populations because their cold-blooded lifestyles establish latitudinal clines in RKN activity. Despite the presence of RKN-responsive resistance genes in populations of tropical crops such as tomato (Solanum lycopersicum) and rice (Oryza sativa), substantial variation in resistance is still observed across populations lacking resistance genes. Genome-wide analyses across these populations reveal QTLs associated with resistance in regulatory regions or within linkage disequilibrium of known transcription factors, suggesting that rewiring of gene regulatory networks (GRNs) might facilitate host adaptation to RKN parasitism. To evaluate this hypothesis, we focus on the interaction between tomato and the RKN Meloidogyne incognita. We are specifically interested in how RKNs manipulate host development to establish permanent feeding structures, galls, in their hostsâ€™ roots due to our detailed understanding of root developmental GRNs and the hypothesis that root developmental variation might facilitate adaptation to novel environments, across which RKN abundance varies. By profiling tissue-specific cell division patterns in

galls, weâ€™re gaining an understanding of the cell types contributing to this interaction. To resolve the GRNs mediating these processes, weâ€™re performing single cell RNA-sequencing of developing tomato galls. To understand how evolution has shaped these processes, weâ€™re comparing gall size across tomato landraces and using genome-wide mapping approaches to understand the network components required for a susceptible interaction.

1166 - Population genomics and spatial modelling provide insights into African human history

Cesar A Fortes-Lima

Cesar A Fortes-Lima, Carina M Schlebusch

Johns Hopkins University, Uppsala University, Uppsala University, University of Johannesburg, SciLifeLab

For the last decade, computational methods have been developed to investigate mass migrations in human history using genetic data. Here, we applied allele-frequency, haplotype-based, and spatial modelling approaches to reconstruct and visualise putative migration routes throughout the African continent. We first collected genetic and geographical data for 5.600 individuals from representative African populations from the whole continent as well as comparative Eurasian populations. Genetic data was generated in previous studies using Illumina arrays of over two million variants and whole-genome sequencing methods. Our spatial visualisation of ancestral components estimated using ADMIXTURE and effective migration rates estimated using FEEMS show i) putative migration routes of Bantu-speaking populations from west-central to east and south sub-equatorial Africa, ii) complex migration patterns of populations with different linguistic affinities across the Sahel belt, and iii) low migration patterns in North African populations due to low-migration rates across the Sahara Desert. We investigated haplotype data using MAPS, to infer both migration rates and population sizes of the studied populations and to disentangle their migration patterns over time. To better understand large expansions of African populations, we tested events of spread-over-spread, genetic continuity and genetic replacement. Our spatial modelling apporach refines our understanding of corridors of higher and lower human migration or expansion across the African landscape that were confirmed with other sources. Therefore, this research highlights how spatially explicit analyses in population genomics can complement and investigate hypotheses from archaeology, biological anthropology, and historical linguistics.

1168 - Unveiling the Functional Dynamics of Noncoding DNA: A Deep Learning Approach with Arabidopsis thaliana

Kristy Surja Mualim

Kristy S Mualim, Moises Exposito-Alonso

Stanford University, Carnegie Institute of Science, University of California, Berkeley, University of California, Berkeley, Howard Hughes Medical Institute

Despite most Genome-Wide Association Studies (GWAS) variants falling within the noncoding regions of the genome, the functional implications of these variants and their contribution to fitness effect within accessible chromatin regions is still poorly understood. These challenges are compounded by difficulties in experimental perturbation - most cis-regulatory elements have small effect sizes that are hard to detect and most studies lack power, sample size or scale. By leveraging deep learning techniques, we describe a deep-learning framework to decode the intricate molecular relationships of how natural sequence variation within Arabidopsis thaliana accessions impacts regulatory function. This framework aims to open new avenues for understanding evolutionary dynamics of noncoding DNA.

1169 - Exploring progenitors of protochordate ocellus as an evolutionary origin of the neural crest cells through single-cell transcriptomics.

Yuliia Fatieieva

Yuliia Fatieieva, Rozalina Galimullina, Igor Adameyko
Medical University of Vienna, Medical University of Vienna, Medical University of Vienna/ Karolinska Institutet

The neural crest, a highly versatile group of embryonic stem cells exclusive to vertebrate embryos, played a crucial role in the evolutionary advancement of vertebrates by providing them with foundational tissues like cartilage, bone, dentition, and a sophisticated nervous system. Despite its significance, the evolutionary origins of neural crest cells remain largely unknown. In this study, we revisit our hypothesis suggesting that neural crest cells may have originated from neuroectodermal precursors associated with the pigmented ocelli in chordates resembling Amphioxus (Ivashkin and Adameyko, EvoDevo 2013). We argue that changes in the requirements of pigment-shaded photoreception led the multipotent precursors of photoreceptors to acquire the ability to migrate beyond the central nervous system, ultimately giving rise to a diverse range of neural, glial, and pigmented photosensory cell types at the body periphery.
To explore this idea, we analyzed single-cell RNA sequencing datasets of pigmented cells originating from the neural crestâ€”melanocytes, and retinal pigment epithelial cells (RPEs) derived from the brain neuroectodermâ€”from both mouse and humans. Our analysis revealed a shared expression profile of multiple transcriptional controllers, beyond those associated with pigmentation, across these cell types. Constructing a cladistics tree based on their transcriptomes grouped these two cell types together, suggesting a common ancestry of their global gene expression programs. To investigate what kind of cell type could be ancestral for vertebrate melanocytes and RPEs we researched scRNAseq data of other chordates and found a supporting association with photosensory pigment-producing cells.

1170 - Enhanced DNA Damage Response in a Clade of Long-Lived Bats Resolved Using Chromosome-Length Genome Assemblies

Juan Manuel Vazquez

Juan M Vazquez, Mary E Lauterbur, Peter J Sudmant, David Enard
University of California, Berkeley, University of Arizona, University of California, Berkeley, University of Arizona

Lifespan is one of the most variable traits across the entire tree of life, and especially in mammals. Differences in lifespans between closely-related species provides a promising avenue for discovering novel pro-longevity pathways using evolutionary techniques. Previous studies focused on the evolution of longevity-associated traits, such as DNA damage response, have been hampered by a combination of low-quality genomes, low-phylogenetic coverage, or long evolutionary times, all of which can negatively affect their power to detect genes associated with longevity. In order to comprehensively study the evolution of aging and aging-associated traits in bats, we generated chromosome-level reference genomes and primary cell line libraries from a 10-million-year-old clade of 9 California Myotis species spanning a 3-fold range of lifespans. Increases and decreases in longevity independent of body size have evolved multiple times in this clade, providing a dynamic range which can be studied through functional genomics. Leveraging both genomes and cell lines, we identify genetic variants in several pathways specifically associated with longevity - including a polymorphic TP53 whole-locus duplication. We additionally find adaptations in other longevity-associated traits such as DNA repair and immunity; and show that these changes are associated with cellular resistance to various forms of chemically-induced DNA damage. These pathways represent new targets for exploration using primary cell cultures, and contribute to our understanding of how both agonistic and antagonistic pleiotropy play a role in the evolution of longevity.

1171 - The genomic and metabolomic architecture of extremophilic archaeon Halobacterium salinarum

Steph Varghese

Steph Varghese, Nahui Medina-ChÃ¡vez, Michael Travisano
University of Minnesota, University of Minnesota, University of Minnesota

Archaea constitute the most understudied domain of life, with different cellular structures, genetic processes, and metabolic pathways than Eubacteria or Eukaryota. This has led to unusual adaptive characteristics in archaeal taxa, including their ability to thrive in extreme environments. Archaea also represent the most metabolically diverse branch of life, having great potential as bioconverters and bioremediators. Recently, Medina-ChÃ¡vez & Kalambokidis et al. (in preparation for Evolution) found that Halobacterium salinarum, an extremophilic archaeon, can adapt quickly to multiple co-occurring environmental stressors by expanding its niche. This pivotal result challenges the expectation that adaptation requires specialization and resultant tradeoffs. A critical follow-up question is: What are the genomic and metabolic bases for H. salinarumâ€™s rapid adaptation that enables this archaeon to thrive in extreme environments? To explore this, we integrate whole genome sequencing and metabolomic analyses to compare biomolecular architectural changes between ancestral and derived populations. This has enabled us to (1) identify localized changes in H. salinarum DNA at the resolution of single-nucleotide polymorphisms; (2) characterize shifts in gene expression related to phenotypic adaptation; and (3) delineate the adaptive potential of this lineage for generating biotechnologically-relevant metabolic products. This work provides the genomic and metabolomic bases for the first-ever long-term archaeal evolution experiment. This foundational insight is necessary for utilizing the most understudied domain of life as a platform for biotechnological innovation.

1172 - Contrasting environments and introgression shape the evolutionary history of three South American fox species

Eduardo J Pizarro

Eduardo J Pizarro, Fabiola Leon, Daly Noll, Bernardita Julio-Kalajzic, Constanza Napolitano, Gerardo Acosta-Jamett, Daniel Gonzalez-AcuÃ±a, Juan Carlos Marin, Javier Cabello, Nicole Sallaberry-Pincheira, Javier MillÃ¡n, Cristian Bonacic, Patricio Pliscoff, Juliana A Vianna
Millennium Institute Center for Genome Regulation (CGR), Millennium Nucleus of Patagonian Limit of Life (LiLi), Departamento de EcologÃ-a, Facultad de Ciencias BiolÃ³gicas, Pontificia Universidad CatÃ³lica de Chile, Millennium Institute Biodiversity of Antarctic and Subantarctic Ecosystems (BASE), Millennium Institute Center for Genome Regulation (CGR), Millennium Nucleus of Patagonian Limit of Life (LiLi), Departamento de EcologÃ-a, Facultad de Ciencias BiolÃ³gicas, Pontificia Universidad CatÃ³lica de Chile, Millennium Institute Biodiversity of Antarctic and Subantarctic Ecosystems (BASE), Millennium Institute Center for Genome Regulation (CGR), Millennium Nucleus of Patagonian Limit of Life (LiLi), Departamento de EcologÃ-a, Facultad de Ciencias BiolÃ³gicas, Pontificia Universidad CatÃ³lica de Chile, Millennium Institute Biodiversity of Antarctic and Subantarctic Ecosystems (BASE), Departamento de EcologÃ-a, Facultad de Ciencias BiolÃ³gicas, Pontificia Universidad CatÃ³lica de Chile, Departamento de Ciencias BiolÃ³gicas y Biodiversidad, Universidad de los Lagos, Institute of Ecology and Biodiversity (IEB), Cape Horn International Center (CHIC), Programa de InvestigaciÃ³n Aplicada en Fauna Silvestre, Facultad de Ciencias Veterinarias, Universidad Austral de Chile, Facultad de Ciencias Veterinarias, Universidad de ConcepciÃ³n, Laboratorio de GenÃ³mica y Biodiversidad, Departamento de Ciencias BÃ¡sicas, Universidad del Bio-BÃ-o, Facultad de Medicina Veterinaria, Campus Pichi Pelluco, Universidad San SebastiÃ¡n, Unidad de RehabilitaciÃ³n de Fauna Silvestre, Escuela de Medicina Veterinaria, Facultad Ciencias de la Vida, Universidad AndrÃ©s Bello, Facultad de Ciencias de la Vida, Universidad AndrÃ©s Bello, Instituto Agroalimentario de AragÃ³n-IA2 (Universidad de Zaragoza-CITA), FundaciÃ³n ARAID, Departamento de Ecosistemas y Medio Ambiente, Facultad de AgronomÃ-a e IngenierÃ-a Forestal, Pontificia Universidad CatÃ³lica de Chile, Departamento de EcologÃ-a, Facultad de Ciencias BiolÃ³gicas, Pontificia Universidad CatÃ³lica de Chile, Millennium Institute Biodiversity of Antarctic and Subantarctic

Ecosystems (BASE), Millennium Institute Center for Genome Regulation (CGR), Millennium Nucleus of Patagonian Limit of Life (LiLi), Departamento de EcologÃ-a, Facultad de Ciencias BiolÃ³gicas, Pontificia Universidad CatÃ³lica de Chile, Millennium Institute Biodiversity of Antarctic and Subantarctic Ecosystems (BASE)

Evolutionary history of recently diversified species challenges genomic studies with processes, such as introgression and incomplete lineage sorting (ILS), that hinders the inference of the true evolutionary relationship between populations and species. Lycalopex is a genus of recently diversified South American foxes with inconsistent evolutionary relationships that are potentially related to introgression and ILS, especially the highly sympatric species, the Andean fox, the South American gray fox (herein SA gray fox), and the Darwinâ€™s fox. Here, we sequenced 35 whole-genomes and performed past and present species distribution models of the three species to decipher the evolutionary processes that shaped their divergence and to better understand the biogeography of the region. Also, 12 published genomes of South American canids were added to the analyses to better understand the evolutionary history of the group. Phylogenetic reconstruction and population analysis support introgression rather than ILS as the main source of phylogenetic inconsistencies. The Andes Mountains and the Pleistocene glacial cycles generated contrasting environments that promoted the diversification of the group, with a parapatric speciation of the Andean fox into desert and high-altitude environments, and a peripatric speciation of the Darwinâ€™s fox from island colonization. Low population structure was detected in the Andean and SA gray fox, with two evolutionary significant units found in the former, a pattern of bioclimatic differentiation in the latter, and introgression between both species in two contact zones. Together with exploratory selection analyses, the results support a divergence favored by geographic isolation and adaptation to different environments.

1173 - Population-level degeneration of young Y chromosomes in the mountain pine beetle.

Ryan R Bracewell

Ryan R Bracewell, Medhavi Verma, Barbara Bentz, Scott Geib
Indiana University, Indiana University, USDA Forest Service, USDA ARS

Neo-sex chromosomes provide a unique opportunity to study the cascade of genomic changes that occur when autosomes transform into heteromorphic sex chromosomes. Much of the work to date has focused on longer-term changes in gene content and gene regulation. Much less is known about natural variation within a species and how neo-sex chromosome variation might segregate in natural populations. The mountain pine beetle is a unique system to explore intraspecific variation because there appears to be three highly differentiated neo-Y chromosomes that have been associated with patterns of reproductive incompatibility between populations. To fully understand the extent of neo-XY differentiation and degeneration, we used a combination of PacBio HiFi, Hi-C, and RNA-seq data to assemble and annotate haplotype resolved male and female assemblies from three distinct neo-X/Y types of mountain pine beetle. We find that the fusion that led to the neo-XY likely involved a fusion between an â€˜oldâ€™ X chromosome and three autosomes. We find an astonishing level of neo-Y chromosome degeneration and structural changes in the form of both large- and small-scale inversions. Transposable elements are enriched on the neo-Y and we find population specific patterns of degeneration and gene loss suggesting that certain neo-X/neo-Y combinations could have gene dosage issues, thereby providing a potential mechanism explaining reproductive incompatibilities.

1174 - Satellite-based phenotyping and spatial population genomics of trees

John T .

John Burley
Australian National University

Trees are ideal empirical systems for many questions in spatial population genetics for the simple reason that individuals are fixed in space. Less widely appreciated is that rapid advances in satellite remote sensing and computer vision can enable individual-level observations of ecologically and evolutionarily important parameters (e.g. flowering time, population density) at the scale of species ranges, at least for certain taxa. Combining targeted spatial population genomics approaches guided by satellite-based detection and phenotyping of trees presents unprecedented opportunities to understand the genetic and environmental basis of trait variation; relationships between demographic and microevolutionary processes; and how flowering time variation may relate to population genetic structure. This presentation reveals first insights from combining remote sensing of near-complete populations and genomics of individual trees (thousands of whole genomes), focusing on widespread keystone tree genera in Central America (Handroanthus) and in Australia (Eucalyptus)

63 - Deep Reconstruction of the Migration History of the World

Eran Elhaik

Eran Elhaik
Lund University

Reconstructing the history of the world is one of the ultimate goals of paleogenomics, yet thus far, only a dozen major migration routes have been identified or speculated. Geographical inference based on anatomical or morphological information is highly complex and error-prone, particularly when the remains are physically damaged or fragmented. Utilizing ancient DNA for localization presents additional challenges due to the lack of intermediate samples over space or time, the small number of SNPs, and their spurious nature. Poorly dated samples pose an even greater challenge to the accuracy of the reconstruction. We developed ancient GPS (aGPS), which implements a dual machine-learning approach. Firstly, to predict the age of the ancient samples, and secondly, to predict their geographical site of origin. We demonstrate that aGPS age predictions align with known dates. We also show that aGPS accurately predicts the burial site of the ancient samples, comparable to when applied to modern samples. When set to identify ancient origins, aGPS maps historical migration routes, thereby unraveling the journey of humankind as far as aDNA data permit. aGPS is not limited to humans and enables the resolution of questions regarding origins and population movements. Addressing long-standing historical questions concerning the identity of Old World residents, aGPS stands as a powerful tool in the field of paleogenomics.

86 - Unlocking the Past: Ancient Protein Analysis Sheds Light on Early Human Evolution in Southern Africa

Imke Lankheet

Imke Lankheet, Carina Schlebusch
Uppsala University, Uppsala University, University of Johannesburg

Although aDNA has significantly improved our understanding of human evolution and human history, it also has its limitations. One of the biggest limitations is the survival and preservation of DNA. Proteins generally preserve longer than DNA and harbor crucial biological information, particularly relevant for early human evolution in Africa. Through the analysis of in-silico translated genomic data and ancient protein sequences, our study identifies several amino acid differences in bone-related proteins between southern African Khoe-San groups and the broader human population. In conclusion, this study marks an important step towards advancing the field of population genetics through the utilisation of ancient proteins. We demonstrate the potential of ancient protein analysis in unraveling the complexities of early human evolution. These findings not only underscore the importance of considering proteins alongside DNA in ancient genomics but also pave the way for future studies examining older samples. Ultimately, this research lays the foundation for a

new era in population genetics, where ancient proteins offer a unique perspective on population continuity and ancient population structure in sub-Saharan Africa.

147 - Reconstructing the history of archaic introgression in modern humans: Insights from whole genome sequences of worldwide populations

Laurits Skov

Laurits Skov, Priya Moorjani

UC Berkeley, UC Berkeley

Analysis of archaic and modern human genomes has revealed evidence of gene flow from archaic homininsâ€“â€“Neanderthal and Denisovansâ€“â€“into modern humans and highlighted its critical role in shaping the genetic and phenotypic variation in modern humans. However, most studies have focused on Europeans and East Asians, with very few genomes from other parts of the world, leaving our understanding of human evolutionary history incomplete. In this study, we integrate data from four different studies including ~31,000 previously published genomes and ~2700 newly sequenced genomes from South Asia. Using present-day genomes, we recover 1.7 Gb of Neanderthal genome and 1.1 Gb of Denisovan genome hidden in modern human genomes. This is the largest fraction of archaic ancestry recovered. By comparing the distribution of archaic segments across worldwide populations, we infer that the largest variation in Neanderthal ancestry is observed in South Asia. For Denisovan ancestry, we infer a large fraction of population-specific variation in Oceanians and South Asians. Finally, we develop a new framework for inferring the timing of archaic admixture using the mutation clock that leverages the new mutations that arose on the introgressed segments since the admixture. This new approach allows us to date Neanderthal admixture to around 50,000 years ago, consistent with the dates from previous studies. Together, these analyses provide a comprehensive picture of the archaic admixture in modern humans and the role of archaic ancestry in our evolution.

161 - Genomic signatures of archaic introgression in human populations across the Pacific

Sofía Vieyra-Sánchez

SofÃ-a Vieyra-SÃ¡nchez, Consuelo D Quinto-CortÃ©s, Stephen J Oppenheimer, Carmina Barberena-Jonas, Alexander Mentzer, Alexander G Ioannidis, AndrÃ©s Moreno-Estrada

National Laboratory of Genomics for Biodiversity (UGA-LANGEBIO), CINVESTAV, National Laboratory of Genomics for Biodiversity (UGA-LANGEBIO), CINVESTAV, School of Anthropology, University of Oxford, National Laboratory of Genomics for Biodiversity (UGA-LANGEBIO), CINVESTAV, The Centre for Human Genetics, University of Oxford, Department of Biomolecular Engineering, University of California Santa Cruz, Department of Biomedical Data Science, Stanford Medical School, National Laboratory of Genomics for Biodiversity (UGA-LANGEBIO), CINVESTAV

Numerous lines of evidence indicate that humans interbred with Neanderthals and Denisovans. While Neanderthal introgression appears to be evenly distributed among human populations, Denisovan introgression has been found by previous studies to be predominantly found in Pacific Islanders. Historically, most genetic studies in the Pacific have focused on Near Oceania, the western part of the region. The Oceanian Genome Variation Project addresses this gap by generating SNP array data from 981 individuals from East Asia and both Near and Remote Oceania. We used this novel dataset to explore the distribution of archaic introgression in a broader geographical range. Our results confirm that the highest global proportions of Denisovan introgression are found in Papuans and Negrito populations from the Philippines. We also found evidence that Polynesian populations have detectable Denisovan introgression, although at lower global proportions than Papuan populations. Ancestry-specific analyses were performed to assess differential segregation of Denisovan introgression in modern human populations. These results revealed that Denisovan introgression in Papua and Polynesia is exclusively

linked to their Papuan component, while Denisovan introgression in Negrito populations from the Philippines is recovered when looking at both the Papuan and East Asian components. This supports the idea that Denisovan introgression involved different human populations at different times, and unveils a wider distribution than previously thought. Finally, the multiple sources of Denisovan ancestry associated with different ancestry components in Negrito populations likely contributed to an unusual assimilation of archaic ancestors, explaining its ranking as the modern population with the highest Denisovan ancestry.

169 - Grave Matters: Discerning Ancient DNA Profiles from Grave Dirt vs. Skeletal Remains

Gözde Ata?

GÃ¶zde AtaÄŸ, Benjamin Vernot

Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology

Sedimentary ancient DNA (sedaDNA) has become a valuable tool for studying past populations in the absence of human remains. Here, we explore the use of sedaDNA as an alternative to ancient DNA from skeletal elements. We collected skeletal material, along with grave dirt from the skeletal surfaces, at the Late Bronze Age site of Didnauri (Georgia) and the Medieval site Alt-Inden (Germany). We generated mitochondrial and nuclear genome data using target enrichment, and analyzed the differences in sequence profiles from sediments versus bones. First, we compare the qualities, coverage distributions, variant properties and haplogroup inferences. Then, we test for the presence of non-human DNA in grave dirt, and evaluate the identified human DNA for sources other than the individual to which the skeletal material belongs. Further, we explore the differences between the two types of data in population genetics analyses, particularly assessing the relative affinities to the adjacent populations, and the consequent clustering patterns. Overall, by investigating ancient DNA sequences derived from the two sources, our study seeks to elucidate whether a comparable accuracy can be achieved in genetic studies by sampling sediments rather than the skeletal material itself.

231 - Reconstructing the genetic relationship between ancient and present-day Siberian populations

Haechan Gill

Haechan Gill, Juhyeon Lee, Choongwon Jeong

Seoul National University, Seoul National University, Seoul National University

Human populations across a vast area in northern Eurasia, from Fennoscandia to Chukotka, share a distinct genetic component often referred to as the Siberian ancestry. Most enriched in present-day Samoyedic-speaking populations such as Nganasans, its origins and history still remain elusive despite the growing list of ancient and present-day genomes from Siberia. Here we reanalyze published ancient and present-day Siberian genomes focusing on the Baikal and Yakutia, resolving key questions regarding their genetic history. First, we show a long-term presence of a unique genetic profile in southern Siberia, up to 6,000 years ago, which distinctly shares a deep ancestral connection with Native Americans. Second, we provide plausible historical models tracing genetic changes in West Baikal and Yakutia in fine resolution. Third, the Middle Neolithic individual from Yakutia, belonging to the Belkachi culture, serves as the best source so far available for the spread of the Siberian ancestry into Fennoscandia and Greenland. These findings shed light on the genetic legacy of the Siberian ancestry and provide insights into the complex interplay between different populations in northern Eurasia throughout history.

240 - Evaluating Local Ancestry Inference Analysis of Ancient DNA Data

Xiaowen Jia

Xiaowen Jia, Yilei Huang, Harald Ringbauer

Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Max Planck Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Institute for Pre- and Protohistoric Archaeology and Archaeology of the Roman Provinces, Ludwig Maximilian University, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Bioinformatics Group, Institute of Computer Science, UniversitÃ¤t Leipzig, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Max Planck Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM)

Local ancestry inference (LAI) is a computational method crucial for identifying the mosaic of ancestry segments within admixed chromosomes. Its application has recently extended to the rapidly growing field of human ancient DNA (aDNA). However, the distinctive characteristics of aDNA data, such as low sequencing coverage and high error rate, present challenges for existing LAI tools designed for modern genomes. Furthermore, the performance of these tools on a subcontinental level with ancient DNA remains unclear.
To explore these caveats, we conduct systematic benchmarking of one of the most widely used LAI software packages (RfMix), which is now also commonly applied to aDNA data. Our approach involves first constructing diploid chromosomes by integrating segments of farmer ancestry into one haploid of a phased Bronze Age Steppe-ancestry genome. Subsequently, we simulate typical aDNA fragments from these constructed ground-truth mosaics. Following the processing with a standard aDNA pipeline, we evaluate the accuracy of RfMix across different sets of reference panels, parameter settings, and aDNA quality.
Our findings highlight the critical importance of reference panel size and reference population selection as primary variables affecting the accuracy of LAI. Additionally, parameters such as window size and generation significantly influence LAI performance. Finally, we identify a concerning systematic error manifesting as a high false positive rate at positions of a true positive segment on the alternate haploid.

288 - Evaluating the applicability of imputation and kinship analyses for ancient sedimentary DNA datasets

Pnina Cohen

Pnina Cohen, Sarah Johnson, Elena I. Zavala, Priya Moorjany, Viviane Slon
Tel Aviv University, University of California, University of California, University of California, Tel Aviv University, Dan David Tel Aviv University

Ancient DNA (aDNA) analysesâ€” the study of genetic material from individuals that died hundreds or thousands of years agoâ€” have revolutionized the research in human evolutionary genetics. At most archaeological sites dated to the Middle or Late Pleistocene (780,000-12,000 years ago), no human remains have been found. However, recent studies have shown that aDNA can be recovered from archaeological sediments, even in the absence of any skeletal remains, providing an exciting new avenue to learn about our evolutionary past. Despite that, so far only a limited number of studies have successfully recovered, identified and authenticated ancient human DNA from prehistoric sediments. This is due to the complexity of ancient sedimentary DNA datasets, which tend to be composed of small quantities of short and degraded fragments, often representing multiple species and multiple individuals, within an overwhelming background of environmental DNA.
We focus on advancing and developing analytical frameworks to support the study of ancient human DNA from sediments. Here, we evaluate current methods for imputation, phasing and kinship analyses of genetic variants such as Glimpse, ancIBD, Beagle and LinkImpute - testing them for accuracy on both empirical data and on simulated datasets mimicking the characteristics of ancient sedimentary DNA, including very low genetic coverage, restricted relevant reference panels, mixture of individuals within a single sample, and the presence of non-human ancient DNA. Based on this, we propose appropriate filters and parameters for the tested software, when applied to the study of ancient human DNA from sedimentary samples.

332 - Elucidating the history of the European crow hybrid zone with paleogenomics

Chyi Yin Gwee

Chyi Yin Gwee, Verena Esther Kutschera, PatrÃ-cia PeÄ•nerovÃ¡, Laura Tassoni, Zlatozar Boev, Zbigniew M Bochenski, Teresa Tomek, Simon Trixl, John Stewart, Dilyara Shaymuratova, Quentin Goffette, Nadja PÃ¶llath, Joris Peters, Laurent Frantz, Love DÃ¡len, Sahra Talamo, Matthias Meyer, Johannes Krause, Jochen Wolf

Ludwig Maximilians-UniversitÃ¤t MÃ¼nchen, National Bioinformatics Infrastructure Sweden, University of Copenhagen , University of Bologna, Bulgarian Academy of Sciences, Institute of Systematics and Evolution of Animals, Polish Academy of Sciences, Institute of Systematics and Evolution of Animals, Polish Academy of Sciences, Landesamt fÃ¼r Denkmalpflege im RegierungsprÃ¤sidium Stuttgart, Bournemouth University, Institute of Problems in Ecology and Mineral Wealth, Tatarstan Academy of Science, Belgian Institute of Natural Sciences, State Collection of Palaeoanatomy Munich, Ludwig Maximilians-UniversitÃ¤t MÃ¼nchen, Ludwig Maximilians-UniversitÃ¤t MÃ¼nchen, Stockholm University, University of Bologna, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Ludwig Maximilians-UniversitÃ¤t MÃ¼nchen

Climatic oscillations influence the population dynamics of closely-related lineages. During the glacial period populations are forced into disparate refugia from which they emerge following the retreat of ice sheets. This results in secondary contact of gene pools that have experienced a period of reproductive isolation. All-black carrion and grey-coated hooded crows meeting along a narrow hybrid zone in central Europe are a textbook example. The hybrid zone is shared by distinct vertebrate groups, thus this contact zone is hypothesized to be shaped by the repeated events of isolation and recolonization during the glacial-interglacial cycles. This biogeographic model is inferred based on genetic signatures of remnant population samples, but accurate interpretation of the demography and historical admixture events are hindered by the presence of ghost population(s) and/ or anthropogenic activities, which may directly or indirectly influence migration patterns.

Using the carrion and hooded crows C. corone as a model, our study aims to directly reconstruct the migration route of a widespread European avian lineage after the Last Glacial Maximum with ancient genomic samples dating as far back as 20,000 years ago. The present study uses genomic data from 20,000 to 100 cal BP to reconstruct the formation of the hybrid zone and the movement of phenotypic genes through time and space. The data lends itself to study the coexistence of crows and humans by comparing their migration pathways as crows are synanthropes and have been culturally important to different civilization through time.

335 - When and where did our ancestor interbreed with Denisovans?

Yoko Satta

Yoko Satta, Sayaka Chiku

SOKENDAI, The Graduate University for Advanced Studies, SOKENDAI, The Graduate University for Advanced Studies

During the out-of-Africa dispersal of modern humans in the Upper Paleolithic, eastern non-Africans interbred multiple times with Denisovans after the Neanderthal admixture and as a result, currently harbor three distinct Denisovan genomic segments, (D0, D1, and D2) that differ both in the geographic distribution and in the extent of sequence divergences from the Altai Denisovan genome (D3). However, it is not known when and where these admixtures occurred. We developed a simple method that identifies Denisovan-derived segments in unphased genotype data. The method assumes a paraphyletic relationship among four genomes from a tested modern human and the reference D3 in addition to clock-like accumulation of nucleotide substitutions. We applied it to 3500-3800 year-old Jomon (F23), 31000 year-old Yana1, and extant Onge1, and obtained 500-900 segments per individual with each length being 30-100 kb. We first classified those segments into D0 and D2 using a two-component Gaussian mixture (GM) model and found a strong correlation between

the number of Denisovan-derived segments and DNA contents. The D0/D2 ratio for F23 and Onge1 is smaller than for East Asians, but much higher than for South Asians.This is consistent with the even higher D0/D2 ratio for Yana1, and strongly suggests that D0 admixture occurred in northern East Asia. The GM model also estimated the D0 and D2 divergence time from D3 as 150-160 kya and 300-340 kya, respectively. We currently date D0 and D2 admixture events.

364 - Exploring the genetic diversity in the Americas through ancient whole genomes

Judith Ballesteros Villascán

Judith Ballesteros VillascÃ¡n, Kay PrÃ¼fer, Kathrin NÃ¤gele
Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology

Whereas considerable strides have been made to characterize the present-day genetic diversity of the Americas, much less data is available for individuals who lived in the past. Fortunately, whole ancient genomes have been produced for approximately 150 individuals to investigate some rudimentary aspects of the human genetic history of the Americas. But whole genomes are rare since DNA is often not well preserved. To overcome this limitation, researchers have employed targeted enrichment methods. These allow for poorly preserved specimens to be analyzed. However, the design of some of the current oligonucleotide probes for enrichment was based on data available from two present-day American populations; therefore, it is possible that this does not represent the complete genetic variation of past American populations well.

In this project, we analyzed the published whole genomes from 30 ancient individuals from the Americas to test how well the enrichment reflects ancient diversity by current tools. We examined how much of the past genetic diversity would be retained after enrichment and found that it can introduce biases that, in some cases, can lead to wrong conclusions. Based on these insights, we propose an alternative strategy that will allow us to gain a more complete understanding of the Americas' genetic history, particularly in the tropical regions where DNA is poorly preserved.

444 - Genetic affinities and social structure analysis of individuals found at the Sa Galera Sanctuary inferred from paleogenomic data

Leonardo Yair Correa Mendoza

Leonardo Yair Correa Mendoza, Jose Jorge Arguello Menendez, Ramon Martin Gordon, Anders Gotherstrom, Cristina Valdiosera, Federico Andres Sanchez Quinto
Center of Genomic Sciences, International Human Genome Research Laboratory, AsociaciÃ³n Amics de na Galera, AsociaciÃ³n Amics de na Galera, Department of Archaeology and Classical Studies, University of Burgos, Department of History, International Human Genome Research Laboratory

Recent paleogenomic studies have investigated the demographic history and social structure dynamics across time and space in Europe. The Bronze Age stands out as a time period with significant cultural shifts and human mobility. In example, interesting findings have been obtained investigating this epoch across the Mediterranean. However, during this time period some regions remain unexplored. The sanctuary of Sa Galera is an archeological site on an islet of the Bay of Palma de Mallorca (Spain), archaeologically dated between 200 BCE and 100 CE. At this site, eight individuals have been found, as well as ceramic artifacts associated with a Phoenician archeological context.
In this project, we extracted, sequenced and analyzed DNA from the Sa Galera individuals to investigate their social structure and

demographic origin. We observed deamination patterns and contamination estimates consistent with an endogenous origin of the data. Regarding their social structure, we characterize the uniparental marker lineages and investigate kinship relations among Sa Galera individuals using autosomic data. Aiming to assess their genetic affinities to other ancient and present-day reference Mediterranean populations, we performed exploratory and hypothesis-driven demographic analyses. Results suggest kin connections among Sa Galera individuals and supported by archeological data. Moreover, uniparental and autosomic demographic analyses suggest at least three different demographic origins for the individuals at this site. Our study provides further evidence of the high mobility of individuals during the Bronze Age, where individuals which are almost contemporary to each other and buried on the same islet, had a different demographic history.

454 - How robust is phylogenetic inference from paleoproteomic data?

Johanna Krueger

Johanna Krueger, Ricardo Fong-Zazueta, David M Alba, XÃ¨nia Aymerich, Robin MD Beck, Enrico Cappellini, Guillermo Carrillo-MartÃ-n, Omar Cirilli, Nathan Clark, Omar E Cornejo, Kyle Farh, Luis FerrÃ¡ndez-Peral, David Juan, Joanna L Kelley, Lukas FK Kuderna, Jordan Little, Joseph D Orkin, Ryan S Paterson, Harvinder Pawar, Tomas Marques-Bonet, Esther Lizano
Institute of Evolutionary Biology (CSIC-Pompeu Fabra University), Barcelona Biomedical Research Park, DÃ©partement de Sciences Biologiques, UniversitÃ© de MontrÃ©al, Institute of Evolutionary Biology (CSIC-Pompeu Fabra University), Barcelona Biomedical Research Park, Institut CatalÃ de Paleontologia Miquel Crusafont (ICP-CERCA), Universitat AutÃ²noma de Barcelona, Cerdanyola del VallÃ¨s, Institut CatalÃ de Paleontologia Miquel Crusafont (ICP-CERCA), Universitat AutÃ²noma de Barcelona, Cerdanyola del VallÃ¨s, School of Science, Engineering and Environment, University of Salford, Geogenetics Section, Globe Institute, University of Copenhagen, Institute of Evolutionary Biology (CSIC-Pompeu Fabra University), Barcelona Biomedical Research Park, nstitut CatalÃ de Paleontologia Miquel Crusafont (ICP-CERCA), Universitat AutÃ²noma de Barcelona, Cerdanyola del VallÃ¨s, Department of Human Genetics, University of Utah, Ecology and Evolutionary Biology, University of California Santa Cruz, Illumina Artificial Intelligence Laboratory, Illumina, Institute of Evolutionary Biology (CSIC-Pompeu Fabra University), Barcelona Biomedical Research Park, Institute of Evolutionary Biology (CSIC-Pompeu Fabra University), Barcelona Biomedical Research Park, Ecology and Evolutionary Biology, University of California Santa Cruz, Illumina Artificial Intelligence Laboratory, Illumina, Department of Human Genetics, University of Utah, DÃ©partement dâ€™anthropologie, UniversitÃ© de MontrÃ©al, Institute of Evolutionary Biology (CSIC-Pompeu Fabra University), Barcelona Biomedical Research Park, Geogenetics Section, Globe Institute, University of Copenhagen, Institute of Evolutionary Biology (CSIC-Pompeu Fabra University), Barcelona Biomedical Research Park, Institute of Evolutionary Biology (CSIC-Pompeu Fabra University), Barcelona Biomedical Research Park, Catalan Institution of Research and Advanced Studies (ICREA), Centro Nacional de AnÃ¡lisis GenÃ³mico (CNAG), Institut CatalÃ de Paleontologia Miquel Crusafont, Universitat AutÃ²noma de Barcelona, Cerdanyola del VallÃ¨s, Institut CatalÃ de Paleontologia Miquel Crusafont (ICP-CERCA), Universitat AutÃ²noma de Barcelona, Cerdanyola del VallÃ¨s, Institute of Evolutionary Biology (CSIC-Pompeu Fabra University), Barcelona Biomedical Research Park

Ancient tooth enamel, and to some extent dentin and bone, contain characteristic peptides that persist for long periods of time. In particular, peptides from the enamel proteome (enamelome) have been used to reconstruct the phylogenetic relationships of fossil specimens and to estimate divergence times. However, the enamelome is based on only about 10 genes, whose protein products undergo fragmentation in vivo and post mortem. This raises the question as to whether the enamelome provides enough information for reliable phylogenetic inference. We address these considerations on a selection of enamel-associated proteins that has been computationally predicted from genomic data from 232 primate species. We created multiple sequence alignments (MSAs) for each protein and estimated the evolutionary rate for each site. We examined which sites overlap with the parts of the protein sequences that are typically isolated from fossils. Based on this, we simulated ancient data with different degrees of sequence fragmentation, followed by

phylogenetic analysis. We compared these trees to a reference species tree. Up to a degree of fragmentation that is similar to that of fossil samples from 1-2 million years ago, the phylogenetic placements of most nodes at family level are consistent with the reference species tree. We tested phylogenetic analysis on combinations of different enamel proteins and found that the composition of the proteome can influence deep splits in the phylogeny. With our methods, we provide guidance for researchers on how to evaluate the potential of paleoproteomics for phylogenetic studies before sampling valuable ancient specimens.

461 - A method for reconstructing ancient DNA methylation of additional archaic samples

Liran Carmel

Krystal White, Peleg Berkowicz, Eran Meshorer, Liran Carmel

The Hebrew University of Jerusalem, The Hebrew University of Jerusalem, The Hebrew University of Jerusalem, The Hebrew University of Jerusalem

Reconstructed DNA methylation maps of ancient samples serve as a proxy for gene activity patterns and provide a valuable tool for studying recent human evolution. Reconstruction of the premortem methylomes of a single Neanderthal, a single Denisovan, and a number of anatomically modern humans identified thousands of Differentially Methylated Regions (DMRs) separating these human groups. In addition to the limited availability of archaic genomes, the small number of archaic genomes was a result of the fact that the current DNA methylation reconstruction method is constrained to high-coverage genomes that have been treated with uracil DNA glycosylase and endonuclease VII (USER) to remove uracils prior to sequencing, a treatment that artificially creates large differences in deamination rates between methylated and unmethylated cytosines. However, this procedure was followed for only one of the three publicly available high-coverage Neanderthal genomes. Here, we develop methods to reconstruct DNA methylation from nonUSER-treated genomes by leveraging a difference in postmortem deamination rates between methylated and unmethylated cytosines. We show that the resulting DNA methylation maps of two additional Neanderthals capture meaningful biological information and can be used to identify lineage-specific DMRs. This triples the number of available Neanderthal methylomes, allowing for discrimination between DMRs that are specific to a single individual and DMRs that are likely to be common to all Neanderthals. We identify and discuss several interesting DMRs.

510 - Evaluating the reliability of f-statistics for making population genetics inferences using archaeological sediment samples

Sarah A Johnson

Sarah A Johnson, Elena Zavala, Pnina Cohen, Viviane Slon, Priya Moorjani

University of California, University of California, Tel Aviv University, Tel Aviv University, University of California

Sediment DNA extracted from archaeological sediment samplesÂ¬Â¬ is the next frontier in ancient DNA studies, as it provides the possibility of comprehensively sampling individuals present at a given archaeological site, even in the absence of skeletal remains. Sediment DNA samples tend to be composed of small quantities of degraded genomic fragments, often representing multiple species and individuals, within an overwhelming background of environmental DNA and present-day DNA contamination. The computational methods to analyze these data are still in their infancy and many widely used population genetics inference tools have not been tested for use with sediment DNA. In this study, we develop a novel simulator and assess the reliability of widely used f-statistics for application to sediment DNA. We systematically test the feasibility of answering two key questions: 1)Did admixture occur between a given set of populations? 2)Which known population is most closely related to a population of interest? Using a demographic model of human evolution, we assess the false discovery rate of f-statistics across various timescales. We further examine power as a function of SNP

number and quality, and sampling age. We find that false discovery rates are relatively low in most situations, but the power of all tested statistics significantly decreases with fewer than 10,000 SNPs and in the presence of extreme ancient DNA damage or contamination. Based on our analyses, we present best practices for conducting population genetic analysis using archaeological sediment DNA, and revisit analyses of ancient human DNA in published sediment samples.

524 - Medieval genomes from eastern Mongolia share a stable genetic profile over a millennium

Juhyeon Lee

Juhyeon Lee, Takehiro Sato, Atsushi Tajima, Tsend Amgalantugs, Batmunkh Batmunkh, Shigeki Nakagome, Toshihiko Miyake, Noriyuki Shiraishi, Choongwon Jeong, Takashi Gakuhari

School of Biological Sciences, Seoul National University, Graduate School of Medicine, University of the Ryukyus, Faculty of Medicine, Kanazawa University, Institute of Archaeology, Mongolian Academy of Sciences, Institute of Archaeology, Mongolian Academy of Sciences, School of Medicine, Trinity College Dublin, Institute for the study of Ancient Civilization and Cultural Resources, College of Human and Social Sciences, Kanazawa University, Institute for the study of Ancient Civilization and Cultural Resources, College of Human and Social Sciences, Kanazawa University, School of History, College of Humanities, Shukutoku University, Faculty of Humanities, Niigata University, School of Biological Sciences, Seoul National University, Institute for the study of Ancient Civilization and Cultural Resources, College of Human and Social Sciences, Kanazawa University

Recent archaeogenomic studies in Mongolia have elucidated the genetic origins of people from the Xiongnu and Mongol eras, but left the Medieval period between them only tangentially explored. Due to this dearth of ancient genomes, the dynamic history of Medieval Mongolia with the rise and fall of numerous polities still lacks a genomic perspective. To fill in this knowledge gap, here we report whole-genome sequences of nine ancient individuals from eastern Mongolia, who were excavated from two nearby cemeteries, Gurvan Dov and Tavan Khailaast. They are distributed from the Xiongnu-Xianbei period (ca. 200 CE) to the Mongol era (ca. 1,400 CE), forming a local time transect encompassing nearly 1,200 years. Remarkably, despite the long-time span, all nine individuals derive most of their ancestry (85â€“100%) from the eastern Eurasian lineages and show low heterogeneity in their genetic composition. This is in contrast to the general pattern observed in previously published Medieval genomes from central Mongolia, who showed higher heterogeneity and overall less eastern Eurasian ancestry, thus calling for a comprehensive archaeogenetic survey of Medieval Mongolia to fully capture the dynamic genetic history in this period.

555 - Paleogenomic analyses reveal a broad geographic range and ecological niche for the extinct American cheetah

Molly Cassatt-Johnstone

Molly Cassatt-Johnstone, Ellie E Armstrong, Megan Supple, Julie Meachen, Grant Zazula, Beth Shapiro

University of California, Santa Cruz, Washington State University, University of California, Santa Cruz, Des Moines University, Natural History Museums of Los Angeles County, Government of Yukon Paleontology Program, University of California, Santa Cruz, Genomics Institute, University of California, Santa Cruz

The American cheetah, Miracinonyx trumani, was an iconic member of North Americaâ€™s ice age megafauna community. However, little is known about its geographical range and ecological preferences, and its relationship to the only living large felid in North America, the puma, Puma concolor, remains contentious. Here, we generated paleogenomic data from three felid bones recovered from permafrost deposits in Yukon, Canada, all dating to around 30 thousand years ago (ka) and tentatively identified as puma. Analyses of mitochondrial genomes from each of these three felids and a ~39X coverage nuclear genome revealed that all three were American cheetahs,

suggesting a much broader range and ecological niche for the species than previously hypothesized. Additionally, we generated a ~23X coverage paleogenome from a M. trumani specimen from Natural Trap Caves in Wyoming, USA, dated to 23,500ka to compare demographic and evolutionary histories between populations. Our results confirm that American cheetahs are sister to pumas and estimate that the two species diverged around three million years ago. Analyses of the paleogenomes indicate that American cheetahs in Yukon and the lower 48 United States had low genome-wide diversity and likely maintained a small long-term effective population size. Scans for selection in the Yukon sample reveal a suite of potential cold-adapted genes, and we identify shared loss of function mutations in both individuals. Our results expand the known range and the inferred life history of American cheetahs, highlighting the potential of paleogenomic data to improve our interpretation of the fossil record.

592 - Reconstructing the Neolithic and Megalithic Demographic History of South Asia from Archaeological site of Burzahom, Kashmir using ancient DNA

Aparna Dwivedi

Aparna Dwivedi, Lomous Kumar, Richa Rajpal, Nagarjuna Pasupuleti, Niraj Rai
DST-Birbal Sahni Institute of Palaeosciences, Academy of Scientific and Innovative Research, DST-Birbal Sahni Institute of Palaeosciences, DST-Birbal Sahni Institute of Palaeosciences, Academy of Scientific and Innovative Research, CSIR- Centre for Cellular and Molecular Biology, DST-Birbal Sahni Institute of Palaeosciences, Academy of Scientific and Innovative Research

Burzahom, the Neolithic-Megalithic archaeological site in Kashmir demonstrates transitions from the Neolithic era, to the Megalithic era, to the early historic period. However, today five decades after the excavation have passed, question on the genetic identities and interactions of these individuals are still unanswered. This study utilizes ancient DNA analysis to analyse the genetic ancestry of these individuals. Petrous bones, teeth and long bones of ten individuals were used for DNA extraction, which were sequenced for shotgun sequencing. Five samples confirmed authentic ancient DNA using the damage pattern analysis. Further, two could be proceeded for downstream analyses. The genetic sex identified both samples as male individuals. The Neolithic sample was dated using Accelerator Mass Spectrometry to around 1927 BCE, while the historic one belonged to around 563 CE. One more Neolithic sample was dated to be around 2027 BCE. Two individuals were assigned the mitochondrial haplogroups M65a and U2b. Both these haplogroups are present in south Asia today. Haplogroup U2b was reported previously in an ancient individual from the Harappan site of Rakhigarhi, India. Autosomal data analyses for both samples will help us further understand the genetic ancestry and migratory history of residents as well as their genetic relationships to contemporary and later populations in South and Central Asia. This study represents one of the oldest ancient genomic datasets from South Asia thus far and is a milestone for ancient DNA analysis in the Indian as well as south Asian context.

598 - Optimising imputation and IBD segment retrieval in ancient genomes

Linda Ongaro

Linda Ongaro, Emily M Breslin, Marco Rosario Capodiferro, Emilia Huerta-Sanchez, Lara M Cassidy
Smurfit Institute of Genetics, Trinity College Dublin, Smurfit Institute of Genetics, Trinity College Dublin, Smurfit Institute of Genetics, Trinity College Dublin, Smurfit Institute of Genetics, Trinity College Dublin, Ecology and Evolutionary Biology and Center for Computational and Molecular Biology, Brown University, Smurfit Institute of Genetics, Trinity College Dublin

Ancient DNA has revolutionised the study of genetic variation and population movements over time. However, significant gaps in ancient genetic data, alongside geographic and temporal variations, hinder fine-scale characterisation of genomic structure and relatedness between individuals and/or archaeological sites. In this study, we implemented the imputation pipeline proposed by Hui et al. (2020), involving a two-step approach: (1) Conducting genotype imputation on individual ancient samples using the GLIMPSE software with the

1000 Genomes Project reference panel. (2) Constructing a multi-individual dataset of confident genotype calls from GLIMPSE for input into Beagle5, enabling a second round of imputation and phasing to improve genotype accuracy and minimise missing data. To benchmark our imputation pipeline, we downsampled high-coverage ancient genomes, including shotgun sequences and 1240k target enrichment data. We investigated each step of the pipeline testing different filter thresholds, including genotype posterior probability after the imputation steps. Besides the missingness, as heterozygote sites pose challenges for imputation, we prioritise heterozygote sensitivity as our primary metric for assessing imputation quality. Our findings highlight that a second round of imputation enhances the quality of imputed genotypes for both shotgun sequences and SNP capture samples. Moreover, we evaluate the performance of the second imputation with the identification of Identity-By-Descent (IBD) shared fragments between individuals using refinedIBD software, including long fragments, by examining known parent-offspring pairs. In conclusion, this study sheds light on improving ancient DNA analysis methodologies offering insights into population genetics and evolutionary dynamics.

645 - Detecting introgressed archaic haplotypes in ancient human genomes

Jiaqi Yang

Jiaqi Yang, StÃ©phane PeyrÃ©gne, Janet Kelso, Benjamin Peter

Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology

Interbreeding with archaic humans introduced Neanderthal DNA to all non-African populations and Denisovan DNA to Asian, American, and Oceanian populations. Investigating archaic ancestry in ancient genomes makes it possible to detect selection on introgressed haplotypes and study the dispersal and adaptation of both archaic and modern humans. However, applying haplotype-based methods to ancient genomes has been challenging, due to the poor data quality of these genomes. Here we compare two methods: 1. Admixfrog, a reference-based method that works with genotype likelihoods and requires high-quality sequenced archaic genomes as sources of introgression; and 2. Hmmix, a reference-free method, which takes genotypes from the target genome as input and needs an unadmixed outgroup to detect haplotypes introgressed from divergent lineages.

We first show how we can adapt the reference-free method to medium-to-low-coverage datasets without imputation. Based on both simulations and empirical data, we find that the reference-free method detects more introgressed haplotypes than the reference-based method when the sequenced archaic genomes are divergent from the introgressing populations, as observed for Denisovan haplotypes in present-day human genomes. However, the reference-based method remains the only solution when the coverage of the target genome is ultra-low. We then investigate Denisovan haplotypes in ancient genomes and show how Denisovan ancestry is distributed in modern humans through space and time.

703 - Unveiling Ancient Viral Communities: Insights from Coprolite and Dental Calculus Metagenomes

Remi Denise

Remi Denise, Alexander Hubner, Irina M Velsko, Colin Hill, Andrey N Shkoporov, Christina Warinner

APC Microbiome Ireland, University College Cork, Max Planck Institute for Evolutionary Anthropology, Leibniz Institute for Natural Product Research and Infection Biology - Hans KnÃ¶ll Institute , Max Planck Institute for Evolutionary Anthropology, Leibniz Institute for Natural Product Research and Infection Biology - Hans KnÃ¶ll Institute , APC Microbiome Ireland, University College Cork, APC Microbiome Ireland, University College Cork, Max Planck Institute for Evolutionary Anthropology, Leibniz Institute for Natural Product Research and Infection Biology - Hans KnÃ¶ll Institute , Friedrich Schiller University Jena, Harvard University

The study of ancient DNA (aDNA) has revolutionized our understanding of past civilizations, providing unprecedented insights into the genetic makeup of ancient populations. Within this realm, the exploration of viral communities preserved in coprolites and dental calculus presents a unique opportunity to delve into the microbial ecosystems of previous eras. In these ecosystems viruses have emerged as crucial players, yet their study remains challenging due to degradation and contamination risks.
These ancient viromes offer valuable insights into the diversity, evolution, and interactions of viruses in past populations. Moreover, they provide a unique lens through which to explore the coevolutionary relationships between viruses and their hosts. In this study we try to assess the importance of the choice of methods used for identification of viruses within aDNA metagenomes. How each methods can bring information at different level (reads, contigs, vMAGs), but also can miss some other informations.

713 - The impact of selection on human sex-specific genetic diversity: an ancient genomes perspective

Adamandia Kapopoulou

Stefan StrÃ¼tt, Adamandia Kapopoulou, Stephan Peischl, Laurent Excoffier
CMPG, Institute for Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics, CMPG, Institute for Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics, Interfaculty Bioinformatics Unit, University of Bern, Swiss Institute of Bioinformatics, CMPG, Institute for Ecology and Evolution, University of Bern, Swiss Institute of Bioinformatics

Differences in inferred demographic histories derived from human uniparental markers have frequently been linked to sex-specific transitions from migratory to sedimentary cultures during the Neolithic revolution. However, purifying selection against deleterious mutations can significantly distort genealogies at linked neutral sites. This poses a major challenge to accurately inferring demographic histories, particularly in regions with low or absent recombination, such as in human uniparental markers.
To address whether the presence and extent of putatively strong selection in mitochondrial DNA and weak selection in Y-chromosomes could account for the observed differences in sex-specific inferred human demography, we pursued two primary objectives. Firstly, we extended a structured coalescent model for strong purifying selection in regions without recombination to allow arbitrarily weak or strong selection. Secondly, we curated a high-quality dataset of newly sequenced ancient Mesolithic and Neolithic genomes in Europe. To achieve this, we developed a customized pipeline for variant calling tailored to the unique challenges presented by ancient haploid chromosomes.
Including our recently acquired ancient dataset and using our extended structured coalescent, we compare patterns of diversity in ancient and modern Y and mtDNA chromosomes to see when differences emerged and whether these differences can be attributed to the sedentarisation of male early farmers, as previously postulated.

714 - The genetic history of Neolithic to Bronze Age Aegean at the crossroad between Anatolia and Europe

Angeliki Papadopoulou

Angeliki Papadopoulou, Nikolaos Psonis, Despoina Vassou, Eugenia Tabakaki, Dilek Koptekin, Samuel Neuenschwander, Ricardo RodrÃ-guez-Varela, Lucas Anchieri, Lukas HÃ¼bner, Stefanos Papadantonakis, Argyro Nafplioti, Sevasti Koursioti, Eleni Stravopodi, Chrysanthi Kontaxi, Vassilis L. Aravantinos, Nina Kyparissi-Apostolika, Yasemin YÄ±lmaz, Sevgi Yorulmaz, Evangelia Daskalaki, Morten E. Allentoft, Anders GÃ¶therstrÃ¶m, FÃ¼sun Ã–zer, Mehmet Somel, Pavlos Pavlidis, Alexandros Stamatakis, Anna-Sapfo Malaspinas, Nikos Poulakakis
Department of Computational Biology, University of Lausanne, Institute of Computer Science, Foundation for Research and

Technology-Hellas, Ancient DNA Lab, Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Department of Biology, School of Sciences and Engineering, University of Crete, Ancient DNA Lab, Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Ancient DNA Lab, Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Ancient DNA Lab, Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Department of Computational Biology, University of Lausanne, Department of Computational Biology, University of Lausanne, Vital-IT, Swiss Institute of Bioinformatics, Centre for Palaeogenetics, Department of Archaeology and Classical Studies, Department of Computational Biology, University of Lausanne, Swiss Institute of Bioinformatics, Computational Molecular Evolution Group, Heidelberg Institute for Theoretical Studies, Institute for Theoretical Informatics, Karlsruhe Institute of Technology, Institute of Computer Science, Foundation for Research and Technology-Hellas, Ancient DNA Lab, Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Department of Biology, School of Sciences and Engineering, University of Crete, Ancient DNA Lab, Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Ancient DNA Lab, Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Department of Biology, School of Sciences and Engineering, University of Crete, Ephorate of Palaeoanthropology and Speleology, Ministry of Culture and Sports, Ephorate of Palaeoanthropology and Speleology, Ministry of Culture and Sports, Ephor Emeritus of Boeotian Antiquities, Ephorate of Palaeoanthropology and Speleology, Ministry of Culture and Sports, Department of Archaeology, DÃ¼zce University, Department of Biological Sciences, Middle East Technical University, Department of Archaeology and Classical Studies, Stockholm University, Lundbeck Foundation GeoGenetics Centre, Globe Institute, University of Copenhagen, Centre for Palaeogenetics, Department of Archaeology and Classical Studies, Stockholm University, Human-G Laboratory, Department of Anthropology, Hacettepe University, Department of Biological Sciences, Middle East Technical University, Department of Biology, School of Sciences and Engineering, University of Crete, Institute of Computer Science, Foundation for Research and Technology-Hellas, Institute of Computer Science, Foundation for Research and Technology-Hellas, Computational Molecular Evolution Group, Heidelberg Institute for Theoretical Studies, Institute for Theoretical Informatics, Karlsruhe Institute of Technology, Department of Computational Biology, University of Lausanne, Swiss Institute of Bioinformatics, Department of Biology, School of Sciences and Engineering, University of Crete, Ancient DNA Lab, Institute of Molecular Biology and Biotechnology (IMBB), Foundation for Research and Technology-Hellas, Natural History Museum of Crete, School of Sciences and Engineering, University of Crete

The Neolithic and Bronze Age (BA) periods in the Aegean are characterized by genetic transitions, extensive population interactions, and migrations. Questions pertaining to the Neolithisation process, the contribution of local hunter-gatherer (HG) populations and the extent and timing of a Steppe-like migration during the BA, are still unanswered. To characterize the population structure across time, we generated whole-genome sequence data for 14 Neolithic to Late BA individuals from the Aegean and Anatolia, complemented them with published ancient genomes, and conducted population genetic analyses including admixture modeling, f-statistics and IBD analyses using both pseudo-haploid and imputed data. By examining the effect of imputation on allele frequencies and Linkage Disequilibrium (LD), we observe that imputation distorts the LD decay and we discuss the potential impact on downstream analyses. Our results show that the Neolithic Aegean population was predominantly derived from Western Anatolian farmers, confirming previous studies and suggesting a genetic turnover from the Mesolithic. By the Late Neolithic and onwards there was an influx of a Caucasus HG/Iran Neolithic-like component, possibly introduced by migrants from the east. During the BA, population heterogeneity in the Aegean increased. An Eastern HG-related component already appeared during the Early BA, suggesting that a Steppe-like ancestry may have reached the Aegean earlier than previously described. The EHG-related component is maximized in Middle BA Aegean populations, but compared to other European BA populations its proportion remains low, suggesting that Steppe-like ancestry reached Greece later than mainland Europe.

759 - An Algorithm for Optimally Partitioning Spatiotemporal Archaeological and Genetic Data

Victor Yan Kin Lee

Victor Yan Kin Lee, Adrian Timpson, Jan Kolar, Robert Staniuk, Stephen Shennan, Mark G. Thomas, Fernando Racimo
Globe Institute, University of Copenhagen, UCL Genetics Institute, Department of Genetics, Evolution and Environment, University College London, UCL Institute of Archaeology, University College London, UCL Institute of Archaeology, University College London, UCL Institute of Archaeology, University College London, UCL Genetics Institute, Department of Genetics, Evolution and Environment, University College London, Globe Institute, University of Copenhagen

Human history is characterized by intertwined biological, cultural and ecological processes. Discretizing space and time is often necessary for comparing these processes across different periods and/or regions of the world. However, partitioning archaeological and biological data into analytically comparable spatiotemporal groupings is far from straightforward. The general approach in archaeology is to group observations into archaeological cultures based on the similarity of discovery circumstances, such as proximity, material culture and stratigraphy. Depending on the type of discretization used, archaeological cultures could have various definitions which do not lend themselves naturally to a comparative analysis with other types of datasets (e.g. genetics).

We propose a novel spatiotemporal partitioning algorithm designed for local time-series analysis that aims to aggregate data points in an automated manner, driven by the distribution of data itself. The algorithm is characterized by an approach that searches for an optimal division of space-time based on user requirements. In our specific case, we sought to compare time-series data in different geographical regions. Thus, our implementation works by minimizing within-group variance in space, while simultaneously maximizing data coverage across time. We applied this to the Big Interdisciplinary Archaeological Database (BIAD) with records spanning the European Neolithic and Bronze Age, enabling us to jointly compare evolutionary and cultural processes throughout Western Eurasia. We expect our method to be widely applicable to archaeo-scientists aiming to model different types of datasets in a statistically explicit and reproducible manner, allowing for more rigorous interdisciplinary studies of the past.

762 - Evaluating the bias affecting population genetics analyses when co-analyzing Whole Genome Shotgun sequencing and 1240K capture genomes

Lucas Anchieri

Lucas Anchieri, Samuel Neuenschwander, Anna-Sapfo Malaspinas, JazmÃ-n Ramos-Madrigal
University of Lausanne, Swiss Institute of Bioinformatics, University of Lausanne, Vital-IT, Swiss Institute of Bioinformatics, University of Lausanne, Swiss Institute of Bioinformatics, Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen

Over the past decade, the amount of publicly available ancient human genome-wide data has increased exponentially. Most of this data has been generated either by whole genome shotgun sequencing (WGS) or in-solution enrichment targeting ~1,24 million SNPs (1240K capture). Previous work has suggested that different sequencing methods may introduce bias into downstream population genetic analyses. Yet, the extent and the nature of this bias remain to be thoroughly investigated and WGS and 1240K are routinely analyzed together in ancient genomic studies. Here, we compiled a dataset consisting of publicly available genome-wide data including WGS and 1240K capture data (with some individuals sequenced by both strategies) resembling a reference panel used in an ancient DNA study. We carry out standard population genetics analyses such as MDS, admixture clustering, and f-statistics, showing the potential biases that can arise from having a mix of WGS and 1240K data in state-of-the-art population genetic analysis. Preliminary results show clustering analyses such as MDS and Admixture are substantially affected by having a mixture of data from different sources, with individuals sequenced similarly appearing more related, and clustering together on early dimensions of an MDS. Our results suggest that this bias could have important implications for the interpretation of ancient DNA data. We also show how some precautions, such as restricting the

analyses to a set of SNPs as suggested by Rohland et al. (2022), can reduce the impact of this bias.

806 - Quantifying differential host adaptation during the viral life cycle

Mary Reed-Weston

Mary R Reed-Weston, David Enard
University of Arizona, University of Arizona

Due to their notorious ability to asymptomatically host highly virulent pathogens, bats are an intriguing study system to examine adaptations that contribute to viral tolerance. Bats have unique life history traits, such as their longevity, high metabolism, and ability to fly, which have been suspected to contribute to their high viral tolerance. While some recent studies have shown that bats lack certain genes related to immunity and inflammation, molecular mechanisms underlying bat viral tolerance are still unknown. Previous research has also made it clear that there is significant adaptation in host genomes in response to viral infection, but little is known about what host proteins related to specific steps of viral infection are under the strongest selection. Therefore, we intend to characterize adaptation by looking at viral-interacting proteins (VIPs) at different stages of viral infection (e.g. entry into the cell, intracellular transport, viral release from the cell). Using human data as proof of concept, we quantified adaptation using an extension of the Mcdonald-Kreitman test. We identified the viral release category specifically as a hotspot of adaptation in humans, as well as data indicating a broad range of adaptation patterns across the different categories of host proteins. We will conduct the same analysis in Myotis bats, where we expect higher power to quantify adaptation due to significantly higher levels of genetic variation than in humans. This will ultimately reveal what viruses bats can and cannot tolerate, and the ways in which they adapt in response to viral infection.

847 - Evaluating allele frequency trajectory and selection coefficient estimates from genealogies including ancient DNA

Aina Colomer i Vilaplana

Aina Colomer-Vilaplana, Antonio Barbadilla, SÃ²nia Casillas, Leo Speidel
Department of Genetics and Microbiology, UAB, Institute of Biotechnology and Biomedicine, UAB, Department of Genetics and Microbiology, UAB, Institute of Biotechnology and Biomedicine, UAB, Department of Genetics and Microbiology, UAB, Institute of Biotechnology and Biomedicine, UAB, Genetics Institute, University College de Londres, Francis Crick Institute

During the dispersal across continents, humans have faced vastly different environments, pathogenic exposure, and technological innovations. Yet, a question still unsolved is the extent to which selection has played a role in shaping our genomes across different time periods of human evolution. The sequencing of increasingly large ancient DNA cohorts from single populations is now starting to make selection acting on our genomes over time directly observable, with several recent techniques enabling the inference of allele frequency trajectories and the associated selection coefficients. However ancient DNA preservation is typically best in colder climates and becomes progressively more difficult to obtain in the deeper past. These dense ancient DNA time series are therefore still only available in limited settings. Recently, new methods for inferring genealogies jointly for modern and ancient genomes - such as Relate, tsinfer, and ARGneedle -, have made it possible to fully leverage haplotype information for the inference of selection, including to infer allele frequency trajectories and selection coefficients. This presentation will explore the benefits of combining genealogical and ancient DNA data to observe genomic adaptation through time. We present a framework to simulate genomes undergoing positive selection that allows the sampling of ancient DNA. We benchmark several new techniques for inferring selection on this data, with and without genealogies and with varying densities of ancient genomes, under different selection regimes, showing that genealogy inference using ancient DNA allows us to extrapolate beyond time periods where aDNA is readily available and improves selection estimates.

885 - Preservation of ancient DNA in archaeological sediment from Stone Age Mongolia

Michael James Boyle

Michael James Boyle, Arina M Khatsenovich, Byambaa Gunchinsuren, Evgeny P Rybin, Dashzeveg Bazargur, Tserendagva Yadmaa, Bolorbat Tsedendorj, John W Olsen, Diyendo Massilani

Yale University School of Medicine, Department of Genetics, Institute of Archaeology and Ethnography of the Siberian Branch of the Russian Academy of Sciences, Institute of Archaeology, Mongolian academy of Sciences, Institute of Archaeology and Ethnography of the Siberian Branch of the Russian Academy of Sciences, Institute of Archaeology, Mongolian academy of Sciences, Institute of Archaeology, Mongolian academy of Sciences, Institute of Archaeology, Mongolian academy of Sciences, University of Arizona, School of Anthropology, Yale University School of Medicine, Department of Genetics, Max Planck Institute for Evolutionary Anthropology, Department of Evolutionary Genetics

Reconstructing the history of human occupation in Eastern Eurasia is hindered by gaps in the fossil record and unreliable dates. How early modern humans colonized this region, and to what extent they interacted with local archaic human groups such as the Denisovans, are unresolved questions that can be investigated through the retrieval of ancient DNA from archaic and early modern humans in Eastern Eurasia. However, the scarcity of human fossils from this region hinders ancient DNA analyses. To overcome this challenge, we developed an approach utilizing sediment from archeological sites with evidence of human occupation as a source of ancient DNA. We collected sediment samples from four archaeological sites in Mongolia: Tolbor-4, Tolbor-21, Tsagaan Agui cave, and Kharganyn Gol 5. We investigated ancient DNA preservation in these samples with shotgun sequencing and by capturing ancient mitochondrial DNA from the most abundant mammalian families found in Eurasia during the Stone Age. We found evidence of ancient DNA preservation in the sediment of Tolbor-4, Tolbor-21 and Tsagaan Agui cave. These results suggest that ancient environmental DNA approaches and higher density sampling of sediment in archaeological sites in Mongolia has the potential to provide insights into the genomics of past human populations in Eastern Eurasia.

958 - Ice Age Interactions: Insights from Late Upper Palaeolithic genomes using haplotype-based analyses

Emily M Breslin

Emily M Breslin, Valeria Mattiangeli, William E Banks, Mathieu Langlais, Sylvain Ducasse, Lara M Cassidy, Emilia Huerta-SÃ¡nchez, Andrea Manica, Daniel G Bradley

Smurfit Institute of Genetics, Trinity College Dublin, Smurfit Institute of Genetics, Trinity College Dublin, University of Bordeaux, CNRS, MCC, PACEA, University of Bordeaux, CNRS, MCC, PACEA, University of Bordeaux, CNRS, MCC, PACEA, Smurfit Institute of Genetics, Trinity College Dublin, Smurfit Institute of Genetics, Trinity College Dublin, Department of Zoology, University of Cambridge, Smurfit Institute of Genetics, Trinity College Dublin

The extreme scarcity of human remains from the Palaeolithic period means the demographic processes which shaped the genomic makeup of Europe after the end of the Last Glacial Maximum (LGM) are still largely unknown. Several distinct genetic ancestries have been reported; the first associated with individuals from Solutrean and Magdalenian archaeological contexts, and the second, named for the Villabruna individual, represents the genetic ancestry which formed the majority ancestry fraction of later European Mesolithic individuals. To investigate the processes by which these major genetic lineages of post-glacial Europe migrated and admixed, we generated six new shotgun sequenced genomes spanning 10,000 years, from the end of the LGM to the beginning of the Holocene. These genomes originate from the south-west of France, a key region in decoding LGM population demography as it may have been continuously occupied since the Aurignacian period. Through the co-analysis of these genomes with existing genetic data, using

imputation and cutting-edge haplotype-based methods, we present a framework for population movements and contact during the late Upper Palaeolithic.

971 - Genomic capture and typing of HLA in individuals from the pre-hispanic and colonial period of Mexico

Walter Nicolas Ortega

Walter Nicolas-Ortega, Viridiana Villa-Islas, Federica Pierini, Alejandra Castillo Carvajal, Ernesto Garfias Morales, Laura Carrillo Oliva, Miriam J. Bravo-LÃ³pez, Marcela Sandoval-Velasco, Julie Wesp, Elizabeth M. PÃ©rez-Campos, Alejandro Meraz Moreno, Lourdes MÃ¡rquez MorfÃ-n, Flora Jay, Emilia Huerta-Sanchez, MarÃ-a C. Ã•vila-Arcos
Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, UNAM, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, UNAM, Max Planck Institute for Evolutionary Anthropology, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, UNAM, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, UNAM, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, UNAM, Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, UNAM, Centro de Ciencias GenÃ³micas, UNAM, Department of Sociology & Anthropology, Instituto Nacional de AntropologÃ-a e Historia, DirecciÃ³n de Salvamento ArqueolÃ³gico, INAH, Escuela Nacional de AntropologÃ-a e Historia, INAH, UniversitÃ© Paris-Saclay, CNRS, INRIA, Laboratoire Interdisciplinaire des Sciences du NumÃ©rique, Department of Ecology, Evolution, and Organismal Biology, Brown University , Laboratorio Internacional de InvestigaciÃ³n sobre el Genoma Humano, UNAM

A factor that possibly influenced the decline of the Indigenous population following European colonization â€‹could be a poor response of the adaptive immune system against pathogens brought from other regions of the world. In this project, we analyze the genetic diversity of the HLA system, one of the most important region related to the adaptive immune response, in archaeological samples from pre-Hispanic and colonial periods. Since ancient DNA (aDNA) genomic libraries usually have low concentrations of endogenous DNA, we performed in-solution genomic capture of the HLA regions with a myBaits kit for 39 ancient individuals (34 from colonial and 5 from pre-Hispanic period, respectively), followed by NGS. We observed an increase in the coverage and depth in 26 genes in the HLA region, amounting to â€‹â€‹100% coverage and depths close to 25x in our best libraries. HLA genotyping is currently being carried out with the tool Optitype, which has been successfully used before in aDNA data. The preliminary findings involve the identification of HLA class I alleles A*02; A*24; B*35; B*40; C*04 and C*03 as the most frequent in the dataset, which coincides with those reported today in Mexico.

983 - Determining the prevalence of adaptive evolution in genomic regulatory sequences of Myotis bats

Addison B. Lander

Addison B. Lander, David Enard, Elise Lauterbur, JesÃºs Murga-Moreno
University of Arizona, University of Arizona, University of Arizona, University of Arizona

It has been demonstrated that viruses act as strong selective agents in host populations, although much of this research is based on identifying adaptation in proteins (e.g., specialized antiviral proteins). However, it is thought that most adaptation in mammalian genomes occurs in sequences that regulate gene expression. In this study, we seek to address the disconnect between research in virus-driven adaptation and adaptation that occurs at the regulatory level. Bats are of particular interest in characterizing virus-induced selection, given their capacity to carry several families of zoonotic pathogens stemming from an extensive history of co-evolution alongside those pathogens. Here, we quantify virus-driven adaptation in gene expression regulatory sequences from both human genomes (as a proof-of-concept) and Myotis bat genomes. Because we do not have experimental annotations of regulatory sequences in the bat

genomes, we use regulatory sequences identified by the neural network Enformer. We modify the Approximate Bayesian Computation extension of the McDonald-Kreitman test (ABC M-K) to apply it to regulatory regions; this test quantifies possible adaptive evolution. The ABC M-K test demonstrates possible signals of adaptive evolution in the DNA sequences that regulate the expression of virus-interacting proteins. These signals of regulatory adaptation allow us to use a greater portion of the genome to identify the pathogens that caused ancient epidemics than would be possible in just relying on coding sequences.

1060 - Benchmarking imputation accuracy in Ancient DNA datasets of pre-Columbian individuals from Mexico

VIRIDIANA VILLA ISLAS

VIRIDIANA VILLA ISLAS , Emily M Breslin, Flora Jay, Emilia Huerta-Sanchez, Maria C Avila-Arcos

International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), Smurfit Institute of Genetics. Trinity College Dublin, UniversitÃ© Paris-Saclay, CNRS, INRIA, Laboratoire Interdisciplinaire des Sciences du NumÃ©rique, Smurfit Institute of Genetics. Trinity College Dublin, Department of Ecology and Evolutionary Biology, Brown University, International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM)

Imputation of ancient genomes from the Americas has been assessed using genomes with relatively high coverage (>10x) and downsampling to lower depths, with results suggesting that imputation is reliable for aDNA datasets with > 0.5x genome coverage, for variants with MAF >5%. Here, we tested the reliability of imputation on lower depth genomes (<4.7X) from pre-Columbian individuals from Mexico. We evaluated the influence of biases introduced during imputation of these low coverage pre-Columbian genomes for which there is not a good representation of Native American individuals in the reference panel. First, we used one genomic dataset of a pre-Columbian (671-867 CE) individual from Mexico (2417Q) with a coverage of 4.7x , and down-sampled it to 2x, 1x, 0.5x, 0.2x, and 0.1x. We then, imputed each down-sampled genome at 77.8 million biallelic SNPs using the 1000 Genomes Project haplotype dataset as the reference panel. The evaluation of concordance between the variants observed in the original dataset and the imputed down-sampled genomes was low (0.67-0.72). Next steps of the project involve testing the effect of different parameters that could be influencing the imputation quality (e.g. usage of genotype likelihoods, improvement of the reference panel by adding local genomes) and evaluation of concordance (e.g. in function of the minor allele frequency or ascertained SNPs). Also, imputed down-sampled genomes from these and other pre-Columbian individuals from Mexico will be assessed to observe whether they conserve the same population structure as the original data set.

74 - Evolution of sex-biased gene expression (SBGE) during transitions to separate sexes in the Silene genus

Aline Muyle

Djivan Prentout, Aline Muyle, Niklaus Zemp, Adil el Filali, Bastien Boussau, Pascal Touzet, Alex Widmer, Jos KÃ¤fer, Gabriel Marais

Department of Biological Sciences, Columbia University, USA, Centre dâ€™Ecologie Fonctionnelle et Evolutive (CEFE), University of Montpellier, CNRS, 34 EPHE, IRD, Genetic Diversity Centre (GDC), ETH, UniversitÃ© de Lyon, UniversitÃ© Lyon 1, CNRS, Laboratoire de BiomÃ©trie et Biologie Evolutive UMR 5558, 69622, UniversitÃ© de Lyon, UniversitÃ© Lyon 1, CNRS, Laboratoire de BiomÃ©trie et Biologie Evolutive UMR 5558, 69622, CNRS, UMR 8198 â€“ Evo-Eco-Paleo, Univ. Lille, F-59000, Institute of Integrative Biology, ETH Zurich, ISEM, Univ Montpellier, CNRS, IRD, CIBIO, Centro de InvestigaÃ§Ã£o em Biodiversidade e Recursos GenÃ©ticos, InBIO, LaboratÃ³rio

Associado, Campus de VairÃ£o, Universidade do Porto, 4485-661

Sexual dimorphism is widespread among species with separate sexes and its extent is thought to be governed by the differential expression of thousands of genes between males and females (known as Sex-Biased Genes, hereafter SBGs). SBGs have been studied in numerous species, but rarely in a comparative way, which curtails our understanding of their evolution, especially during multiple independent transitions to separate sexes. We sequenced the transcriptomes of nine dioecious species, two gynodioecious species (separate females and hermaphrodites) and two hermaphrodite species from the Silene genus. Our dataset provides access to three independent transitions to dioecy (dating from less than 1 Myo to about 11 Myo). We demonstrated that male-biased expression emerges first during a transition to separate sexes, later followed by female-biased genes. Furthermore, we showed that, despite a mixture of selective regimes, positive selection significantly affects the evolution of some SBGs. Overall, this study provides new insights on the causes of SBG evolution during transitions to separate sexes.

114 - Genome and Sex Chromosome Analyses of Japanese Frogs Carrying Both XY and ZW Chromosomes within the Same Species

Yukako Katsura

Yukako Katsura
Kyoto Univeristy

In the evolution of sex chromosomes, particular attention is given to the poorly understood phenomenon of "sex chromosome turnover." Sex chromosome turnover is defined as the phenomenon where the sex chromosome system changes from XY type to ZW type (or vice versa), or even when different evolutionarily originated sex chromosomes emerge within the same type (e.g., from XY to another XY). The Japanese frog (Glandirana rugosa) inhabiting Japan is a rare species that possesses both XY and ZW sex chromosomes within the same species, and the turnover of sex chromosomes from XY to ZW has been observed. Our research group aims to elucidate at the molecular level how the turnover of sex chromosomes occurs in the frogs. Previously, we determined the nuclear genome of the frog using short read sequencing (Katsura et al., LSA 2021). Furthermore, it has been suggested that sex chromosomes originating from at least three different chromosomes independently emerged within the population (Miura et al., Mol. Ecol. 2022). The frogs have 13 chromosomes, and in two populations (Tokai region and Hokuriku Tohoku region), the chromosome 7 has morphologically differentiated into ZW and XY sex chromosomes, as reported in previous studies. However, in other populations, sex chromosomes do not show morphological differentiation. In addition to presenting the results of sequence comparisons of morphologically differentiated XYZW sex chromosomes, I would like to introduce the findings from our analyses of populations and genome studies in the frogs.

119 - Faster Z under the sea: drift drives evolution of sex chromosomes in the salmon louse

Andrew J. Mongue

Andrew J Mongue, Robert B Baird
University of Florida, University of Edinburgh

One key limitation to existing knowledge of sex chromosome evolution is an intense focus on only a handful of taxa, resulting in uncertainty about whether observed patterns reflect general processes or are idiosyncratic to the more widely-studied clades. In particular, the Z chromosomes of female heterogametic (ZW) systems tend to be quickly but not adaptively evolving in birds, while in butterflies and moths Z chromosomes tend to be evolving adaptively, but not always faster than autosomes. To understand how these two observations fit into broader evolutionary patterns, we explore, for the first time, patterns of Z chromosome evolution outside of these

two well-studied clades. We utilize a publicly available high-quality genome, gene expression, population, and outgroup data for the salmon louse Lepeophtheirus salmonis, an important aquacultural pest copepod. We find that the Z chromosome is faster evolving than the autosomes, but that this effect is driven by increased drift rather than adaptive evolution. Due to high rates of female reproductive failure, the Z chromosome exhibits only a slightly lower effective population size than the autosomes which is nonetheless sufficient to decrease efficiency of hemizygous selection acting on the Z. These results highlight the usefulness of organismal life history in calibrating population genetic expectations and demonstrate the usefulness of the ever-expanding wealth of modern publicly available genomic data to help resolve outstanding evolutionary questions.

121 - Direct selection on recombination rates alters the recombination landscape in fission yeast

Cristina Berenguer Millanes

Cristina Berenguer Millanes, Bart P.S. Nieuwenhuis
Ludwig-Maximilians UniversitÃ¤t MÃ¼nchen, Ludwig-Maximilians UniversitÃ¤t MÃ¼nchen

The canonical theory for sex chromosome evolution says that sex chromosomes evolved from a pair of homologous chromosomes with a sex-determining locus, followed by incorporation of sexually antagonistic alleles through recombination arrest. Inversions are assumed to be the main mechanism of recombination suppression and successive incorporation of loci with sexual antagonistic alleles will lead to divergence of the sex chromosomes over time. However, little experimental evidence for this process exists. We used the fission yeast Schizosaccharomyces pombe as a model species to study the effects of selection for recombination rate changes on the first steps of sex chromosome evolution, simulating sexual antagonism or negative epistasis. We performed an artificial evolution experiment for 36 sexual generations, sorting by Fluorescent Activated Cell Sorting (FACS) combinations of fluorescent proteins encoded at specific linked loci in the genome. Each generation the haploid offspring were selected for specific allele combinations, and used as parents in a successive generation. The selected loci were either located on the mating-type chromosome or an autosome. Selection was expected to increase linkage disequilibrium and lead to reduced recombination rates between the loci. Additionally two selection regimes for increased recombination, or hotspots were performed. While reduction in recombination rates was not observed in the autosomal regions, due to reduced selection pressure by homozygosity after mating, reduced recombination around the mating type was observed. This reduction was likely caused by antagonistic selection, with inversions or other chromosomal rearrangements being selected when reduced recombination increased fitness.

381 - The sexual system of the dioecious species Laurus azorica

Catarina Branco

Catarina Branco, Ruben Rego, Gabriel Marais, Raquel Linheiro, Raquel Tavares, Sylvain Santoni, Alessio Iannucci, Claudio Ciofi, Maria A. Diroma, Chiara Natali, Marco Morselli, Guilherme Roxo, Roberto Resendes, Manuela Sim-Sim, Helena Trindade, Sara Vicente, JoÃ£o P. Marques, Vitor Sousa, JosÃ© Melo-Ferreira, Ann M Mc Cartney, Alice Mouton, Giulio Formenti, Jos KÃ¤fer, Monica Moura
Centro de InvestigaÃ§Ã£o em Biodiversidade e Recursos GenÃ©ticos, InBIO LaboratÃ³rio Associado, Universidade do Porto, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, Faculdade de Ciencias, Universidade do Porto, Centro de InvestigaÃ§Ã£o em Biodiversidade e Recursos GenÃ©ticos, InBIO LaboratÃ³rio Associado, Universidade do Porto, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, UNESCO Chair â€“ Land Within Sea: Biodiversity & Sustainability in Atlantic Islands, University of the Azores, Faculty of Sciences and Technology, University of the Azores, Centro de InvestigaÃ§Ã£o em Biodiversidade e Recursos GenÃ©ticos, InBIO LaboratÃ³rio Associado, Universidade do Porto, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, Faculdade de Ciencias, Universidade do Porto, Centro de InvestigaÃ§Ã£o em Biodiversidade e Recursos GenÃ©ticos, InBIO LaboratÃ³rio

Associado, Universidade do Porto, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, Faculdade de Ciencias, Universidade do Porto, Centro de InvestigaÃ§Ã£o em Biodiversidade e Recursos GenÃ©ticos, InBIO LaboratÃ³rio Associado, Universidade do Porto, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, Faculdade de Ciencias, Universidade do Porto, UniversitÃ© de Montpellier, University of Florence, National Biodiversity Future Center , University of Florence, University of Florence, University of Florence, Okinawa Institute of Science and Technology, University of Parma, Centro de InvestigaÃ§Ã£o em Biodiversidade e Recursos GenÃ©ticos, InBIO LaboratÃ³rio Associado, Universidade do Porto, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, UNESCO Chair â€“ Land Within Sea: Biodiversity & Sustainability in Atlantic Islands, University of the Azores, Faculty of Sciences and Technology, University of the Azores, Centro de InvestigaÃ§Ã£o em Biodiversidade e Recursos GenÃ©ticos, InBIO LaboratÃ³rio Associado, Universidade do Porto, Faculdade de CiÃªncias, Universidade de Lisboa, Museu Nacional de HistÃ³ria Natural e da CiÃªncia , Universidade, Faculdade de CiÃªncias, Universidade de Lisboa, Escola Superior de SaÃºde Ribeiro Sanches (ERISA), Instituto PolitÃ©cnico da Lusofonia (IPLUSO), Centro de InvestigaÃ§Ã£o em Biodiversidade e Recursos GenÃ©ticos, InBIO LaboratÃ³rio Associado, Universidade do Porto, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, Faculdade de CiÃªncias, Universidade de Lisboa, Centro de InvestigaÃ§Ã£o em Biodiversidade e Recursos GenÃ©ticos, InBIO LaboratÃ³rio Associado, Universidade do Porto, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, Faculdade de CiÃªncias, Universidade do Porto, Genomics Institute, University of California, University of Florence, University of Liege, The Rockefeller University, Institute of Evolutionary Science of Montpellier, University of Montpellier, Centro de InvestigaÃ§Ã£o em Biodiversidade e Recursos GenÃ©ticos, InBIO LaboratÃ³rio Associado, Universidade do Porto, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, UNESCO Chair â€“ Land Within Sea: Biodiversity & Sustainability in Atlantic Islands, University of the Azores, Faculty of Sciences and Technology, University of the Azores

The Laurus genus spans the southern Black Sea region, the Mediterranean Basin, Morocco, and the Macaronesian islands. Despite numerous attempts, previous studies faced challenges in elucidating this genus's taxonomy, ploidy, and phylogenetic relationships. This study delves into L. azorica, a dioecious tree endemic to the Azores archipelago that falls within the European Reference Genome Atlas initiative to ensure high-quality reference genomes of the European biodiversity. Despite its well-established dioecy, its sex-determination system remains elusive. We intend to clarify its sex system to understand the role of sexual systems and polyploidy in island colonization. Here, we aim (i) to obtain a chromosome-level assembly and (ii) to identify the sex determination system. Firstly, long-read sequencing data from Pacific Biosciences (PacBio) were acquired. Additionally, chromatin conformation capture data (Hi-C) were incorporated to facilitate the construction of a chromosome-level assembly. Then, we obtained Illumina short reads for 32 individuals. SDPOP (KÃ¤fer et al. 2021) analysis will be applied to evaluate alternative sex-determination systems (i.e., presence or absence of sex chromosomes, distinguishing between XY and ZW systems) and to infer the sex-linked genes. Our preliminary results suggest that L. azorica is tetraploid (n=18). Ongoing research is expected to yield additional insights into the sex-determination system and contribute valuable information to broader studies on the role of dioecy and sex-determination mechanisms in plant species and their expansion in islands. Moreover, this study is a cornerstone for exploring the evolutionary dynamics within the Laurus genus.

446 - Evolution of anisogamy in early diverging fungus Allomyces

Sujal Phadke

Sujal Phadke, Melissa Sayres

California State University, Arizona State University

Why do many organisms have large eggs and tiny sperm? Addressing this fundamental question in evolutionary biology requires investigation into molecular pathways that regulate gamete dimorphism (anisogamy). It also requires investigation into selection pressures that may explain the prevalence of anisogamy. The fungus Allomyces belongs to the phylum Blastocladiomycota, which is an

early diverging branch in fungi. Closely related species of Allomyces show anisogamy with dimorphism in gamete size, pigment, motility and pheromone production. Allomyces carries homologues of CatSper genes that are critical for mammalian sperm motility. Our preliminary RNA-seq analysis showed that male-biased samples of Allomyces overexpress CatSper homologs compared to the female-biased samples, which is reminiscent of the patterns of CatSper expression in mammals including humans. Furthermore, other groups have shown that synthetic analogue of Sirenin- the female pheromone of Allomyces, mimics the action of human hormone progesterone and can cause cross-species activation of human sperm motility through CatSper complexes in the membrane of human sperm. We hypothesized that genes responsible for gamete dimorphism and interaction are conserved between mammals and early diverging fungi including Allomyces. We identified homologs of additional gamete-specific proteins and investigated their patterns of expression in male-biased vs female-biased mutants in Allomyces. Our results suggest that the role of some of these proteins in gamete-interaction may have evolved at the root of Ophisthokonts.

450 - The fast and slow of X chromosome evolution under paternal genome elimination

Robert B Baird

Robert B Baird, Andrew J Mongue, Laura Ross

Massachusetts Institute of Technology, University of Edinburgh, University of Florida Gainesville, University of Edinburgh

Comparisons between sex chromosomes and autosomes can provide fundamental insights into how the genome evolves and the forces that shape it. Specifically, differences in their transmission and ploidy drive their divergent evolutionary trajectories, with sex chromosomes being predicted to adapt and consequently evolve faster. However, in the well-studied model systems these effects are confounded: sex-linked genes are under less efficient selection due to their reduced effective population size, as well as more efficient selection because recessive mutations are exposed to selection in the heterogametic sex. Fungus gnats and gall midges are two fly families that present unique opportunities to disentangle these effects. They have XO sex determination, but males transmit only maternally inherited chromosomes. This phenomenon, known as paternal genome elimination (PGE), results in equal transmission of the X and autosomes, allowing the effect of hemizygous selection to be studied in isolation. We examine rates of divergence under PGE and find that, surprisingly, X chromosomes diverge more slowly than autosomes. Using population resequencing and gene expression data, we find that stronger purifying selection on the X drives its slower divergence. Our findings demonstrate the potential of systems with unusual inheritance for understanding fundamental evolutionary processes.

535 - Hidden sex-chromosomes in Middle-Earth - Revealing sex-chromosome diversity within the geckos of Aotearoa New Zealand

Ludovic Dutoit

Phoebe Webster, Sarah Walters, Alison Cree, Jo Monks, Tony Gamble, Mary Morgan-Richards, Samantha King, Tania King, Ludovic Dutoit

University of Otago, University of Otago, University of Otago, University of Otago, University of Otago, Marquette University, Massey University, Wildlands , University of Otago, University of Otago

Sex-determination systems are of critical importance to a range of fundamental evolutionary processes. The dynamics of speciation, local adaptation, and genetic conflict are all affected by sex chromosome systems. Yet, testing hypotheses empirically is challenging because of the typical lack of diversity in sex-determination systems over short evolutionary time scales.

Within vertebrates, geckos exhibit an extraordinary diversity of sex-determination systems. Male and female heterogamety have evolved several times from an ancestral temperature sex-determination system that can still be found in a few species today. Aotearoa New

Zealand is often considered a land of odd birds, but it is truly a land of wonderful lizards, inhabited by at least 48 gecko species in a monophyletic radiation. We will present work highlighting several sex-chromosome transitions within New Zealand geckos alone, suggesting that these species present an attractive model system to study extremely young sex-chromosome systems in vertebrates.

809 - Using the Y-Chromosome to Resolve Controversial Mammalian Phylogenetic Relationships

Emmarie P Alexander

Emmarie P Alexander, William J Murphy

Interdisciplinary Program in Genetics & Genomics, Texas A&M University, Veterinary Integrative Biosciences, Texas A&M University, Interdisciplinary Program in Genetics & Genomics, Texas A&M University, Veterinary Integrative Biosciences, Texas A&M University

Despite significant progress made towards defining the relationships of placental mammals (Eutheria), some branches deep within the phylogeny remain controversial. This discordance is largely driven by several biological factors â€" including incomplete lineage sorting and gene flow â€" which make the elucidation of the true species tree challenging. Within Eukarya, meiotic recombination considerably influences the pattern and distribution of phylogenomic signals across the genome. Regions with historically low recombination rates are often enriched for the species tree. The majority of placental mammals possess a XX/XY sex chromosome system which are extensively differentiated in structure and sequence â€" with only the pseudoautosomal regions still recombining. Therefore, the male-specific portion of the Y (MSY) lacks recombination, has elevated rates of mutation, and is subject to background selection. Yet, the Y has rarely been used to address deep-level phylogenetic questions due to sequencing, assembly, and alignment difficulties across diverse clades. To expand our knowledge of mammalian Y-chromosome evolution and function, we present the first comprehensive phylogenetic analysis of placental mammal Y chromosomes across a diverse sampling of ordinal clades using a hierarchical alignment approach from whole genome sequence assemblies. We present a Y chromosome phylogeny for Eutheria and compare it to phylogenies constructed using the X and autosomes. Moreover, we discuss the utility of sex chromosomes in delineating controversial phylogenetic relationships and their molecular signatures of selection based on patterns of MSY gene loss across different lineages.

952 - Evolutionary Insights into Ape Sex Chromosomes from Telomere-To-Telomere (T2T) Genome Assemblies: Palindromes and Multi-Copy Genes.

Karol Pal

Karol Pal, Robert Harris, Monika Cechova, Huiqing Zeng, Brandon D Picket, Prajna N Hebbar, Sergei K Pond, Kateryna D Makova, T2T Primates Consortium

Department of Biology, The Pennsylvania State University, Department of Biology, The Pennsylvania State University, Department of Biomolecular Engineering, University of California, Santa Cruz, Department of Biology, The Pennsylvania State University, National Human Genome Research Institute Home, National Institutes of Health, Genomics Institute, gineering, University of California, Santa Cruz, Department of Biology, Temple University, Department of Biology, The Pennsylvania State University, Department of Biology, The Pennsylvania State University, National Human Genome Research Institute Home, National Institutes of Health, Department of Genome Sciences, University of Washington

The Telomere-To-Telomere (T2T) Primates Consortium is building complete genome assemblies for multiple primate species, leveraging methods recently developed for constructing gapless assemblies of the human genome. Here, we, for the first time, assembled T2T sequences of the X and Y chromosomes of five great apes and one lesser apeâ€"all endangered species. These assemblies have enabled the first comprehensive analysis of palindromes and multi-copy genes evolution, the most challenging regions but also the hallmarks of ape sex chromosomes. Palindromesâ€"long inverted repeats undergoing intrachromosomal recombination and gene conversion between

their armsâ€”have evolved on the Y to counteract the accumulation of mutations due to lack of recombination (outside of the pseudoautosomal regions), but they are also present on the X. We found Y palindromes to be 2â€“3 times longer and constitute a higher proportion of the chromosome as compared to X palindromes. Y palindromes were frequently species-specific, whereas X palindromes were largely shared among species. Interestingly, X palindromes had high gene density and harbored housekeeping genes. In contrast, Y chromosome palindromes carried multi-copy gene families with lineage-specific expansions (CDY in Pongo, and RBMY in Panâ€”both gene families involved in spermatogenesis). Besides confirming previously described multi-copy genes, we identified new ones, and searched for their autosomal homologs. We discuss evolution of these genes in light of our analyses of their selection and copy number variation. The potential role of these genes in determining mating patterns and levels of sperm competition should be evaluated in further functional studies.

1061 - DNA methylation in female and male X chromosomes in human neurons and oligodendrocytes

Yong Hwee Eddie Loh

Robert Morgan, Yong Hwee Eddie Loh, Devika Singh, Isabel Mendizabal, Soojin V Yi

Georgia Institute of Technology, Arbor Biotechnologies, University of California Santa Barbara, Georgia Institute of Technology, Foundation Medicine, Inc., Center for Cooperative Research in Biosciences (CIC bioGUNE), Ikerbasque, Basque Foundation for Science, Biocruces Bizkaia Health Research Institute, University of California Santa Barbara

The presence of heterogametic sex chromosomes often leads to dosage imbalance between sexes. For example, if all human X-linked genes were expressed equally, females, with two X chromosomes, would have two times the RNA products of the X-linked genes compared to males. X chromosome inactivation (XCI) is considered to have evolved to silence one of the two female X chromosomes to balance the gene dosage, and one of the key mechanisms in this transcriptional silencing process is DNA methylation, whereby a methyl group (-CH3) is added to cytosine bases in DNA. Most DNA methylation studies to date often exclude the X chromosomes in their analyses. Moreover, many of these studies used array-based methods that examine only subsets of genomic positions, rather than the whole X chromosome. Consequently, our understanding of X chromosome DNA methylation lags behind that of autosomes. In this study, we investigated X chromosome DNA methylation using 89 neuronal and oligodendrocyte methylomes generated using whole genome bisulfite sequencing (WGBS), which provides an unbiased measurement of DNA methylation at single base resolution across the entire genome. This study furthers our understanding of gene dosage regulation by DNA methylation on the chromosomal level as well as on individual gene levels.

1091 - An evolutionary transcriptomics approach to understanding the origin of the HPG axis in sea lamprey, a basal vertebrate

Sara Victoria Good

Sara V Good, Avery Keen, Tamanna Yasmin

University of Manitoba , The University of Winnipeg, The University of Winnipeg, University of Manitoba, University of British Columbia

Lampreys are jawless fish that belong to the oldest extant lineage of vertebrates and are the oldest lineage to harbour a pituitary. They have a complex life cycle that involves larval, juvenile and adult stages. Lampreys are considered models for the origin of the hypothalamus-pituitary-gonadal (HPG) axis. The objective of this research project is to understand the origin, mode of evolution, and expression of genes related to the HPG axis in gonadogenesis throughout development. First, we used phylogenetics and syntenic analyses to identify and classify putative orthologs of core vertebrate genes involved in the HPG axis in lampreys, based on whether the genes originated in i) the pre-vertebrate genome (amphioxus); ii) the post 1R genome; or iii) were cyclostome specific gene duplicates.

Next, we employed an evolutionary transcriptomics analysis by further grouping genes into evolutionary strata based on their rate of evolution as estimated by the number of nonsynonymous to synonymous substitutions of genes relative to the outgroup hagfish. We then tested the hypothesis that slower evolving genes and/or those present in the pre-vertebrate ancestor, were more likely to be expressed during earlier developmental stages and exhibit less variation in gene expression over developmental stages, while lineage specific gene duplicates present within jawless fish or those exhibiting a higher rate of evolution exhibit higher variation in gene expression across sex and/or stage. The results of this research shed light on the core and peripheral genes involved in the early origin of the vertebrate-specific HPG axis.

76 - Tying population genomics to chemical pollutants in an anthropogenic hybrid zone

Benjamin M Moran

Benjamin M Moran, Wilson F RamÃ-rez Duarte, Daniel L Powell, Terrance TL Yang, Chelsea Rochman, Molly Schumer
Stanford University, University of Toronto, Stanford University, Stanford University, University of Toronto, Stanford University

As humans modify natural environments, it can cause new overlaps between species, remove behavioral barriers to interspecific mating, and increase the relative fitness of hybrids, all of which may increase the observed frequency of hybridization. As such, the growing evidence for widespread hybridization in nature may not only be because of improved methodology, but because anthropogenic effects on the environment are actively increasing the frequency of hybridization. Here, we connect anthropogenic changes in water chemistry to recent hybridization in the swordtail fishes Xiphophorus birchmanni and X. malinche. These sister species overlap along an elevational gradient in multiple independent rivers of Hidalgo, MÃ©xico, and the genetic structure within these areas of sympatry vary from low levels of gene flow to a complete collapse into a hybrid swarm. Prior work has shown that olfactory mate choice in these species can be disrupted by changes in chemical parameters, but no attempt has been made to directly tie hybridization to water chemistry changes observed in nature. We have collected an extensive dataset of water chemistry and low-coverage genomes across four X. birchmanni / X. malinche hybrid zones, and observe that the locations of highest pollution are those in which sympatric populations collapse into a hybrid swarm. We use olfactory histology to test for cellular damage in polluted and unpolluted waters, and test for the same histological effects in fish exposed to relevant pollutants in the laboratory. In doing so, we develop a case study of how anthropogenic change is impacting biodiversity through hybridization.

82 - Genomic insights into a forest reserve reveal ecological stratification of R genes and CYP450 family genes

Jia Jun Ngiam

Jia Jun Ngiam, Matti Asser Niissalo, Sitaram Rajaraman, Joan Jong, Shawn Lum, Gillian Khew, Jarkko Tapani SalojÃ¤rvi
Nanyang Technological University, National Parks Board, University of Helsinki, Nanyang Technological University, Nanyang Technological University, National Parks Board, Nanyang Technological University

The tropical rainforests in Southeast Asia are amongst the world's most hyper-biodiverse and species-rich regions. One prominent hypothesis for the maintenance of species-rich assemblages in the tropics is the Janzen-Connell hypothesis, where plant diversity is driven and maintained by plant-pathogen coevolution. However, the eco-evolutionary dynamics of plant immunity with respect to forest ecological succession is relatively understudied, due to the paucity of baseline genomic information. Here, we sequenced and assembled draft genomes of 499 angiosperms from a conserved forest reserve in Singapore. Our analyses reveal that plant investment to defenses differs across varying plant lifeforms with a strong phylogenetic signal, suggesting shared evolutionary histories of immune responses. Interestingly, we also found a strong stratification of resistance genes and defense-related secondary metabolism across different forest

successional zones. Our findings contribute novel genomic perspectives into ecological sorting in relation to the diversification of plant immunity and secondary metabolism in tropical forest ecosystems.

91 - The influence of gene flow on population viability in an isolated urban caracal population

Christopher C Kyriazis

CHRISTOPHER C KYRIAZIS, Laurel Serieys, Jacqueline M Bishop, Marine Drouilly, Storme Viljoen, Robert K Wayne, Kirk E Lohmueller
San Diego Zoo Wildlife Alliance, Panthera, University of Cape Town, Panthera, University of Cape Town, UCLA, UCLA

Wildlife populations are becoming increasingly fragmented by anthropogenic development. Small and isolated populations often face an elevated risk of extinction, in part due to inbreeding depression. Here, we examine the genomic consequences of urbanization in a caracal (Caracal caracal) population that has become isolated in the Cape Peninsula region of the City of Cape Town, South Africa and is thought to number ~50 individuals. We document low levels of migration into the population over the past ~75 years, with an estimated rate of 1.3 effective migrants per generation. As a consequence of this isolation and small population size, levels of inbreeding are elevated in the contemporary Cape Peninsula population (mean FROH>1Mb=0.20). Inbreeding primarily manifests as long runs of homozygosity >10Mb, consistent with the effects of isolation due to the rapid recent growth of Cape Town. To explore how reduced migration and elevated inbreeding may impact future population dynamics, we parameterized an eco-evolutionary simulation model. We find that if migration rates do not change in the future, the population is expected to decline, though with a low projected risk of extinction. However, if migration rates decline or anthropogenic mortality rates increase, the potential risk of extinction is greatly elevated. To avert a population decline, we suggest that translocating migrants into the Cape Peninsula to initiate a genetic rescue may be warranted in the near future. Our analysis highlights the utility of genomic datasets coupled with computational simulation models for investigating the influence of gene flow on population viability.

101 - Genomic Consequences of Strong Bottlenecks in a Previously Endangered Bird

Anna Maria Calderon

Anna Maria Calderon, Andrew W Wood, Zachary A Szpiech, David Toews
The Pennsylvania State University, -, -, The Pennsylvania State University, The Pennsylvania State University, The Pennsylvania State University

Runs of homozygosity (ROH) are homozygous chromosomal segments that arise from population bottlenecks, isolation, and inbreeding, when two related individuals mate and pass on identical by descent (IBD) haplotypes to an offspring. In addition to providing insights into levels of inbreeding and genetic diseases, quantifying the distribution and prevalence of ROH can also illuminate population history such as bottlenecks and population expansions. In the early 1970â€™s Kirtlandâ€™s warblers (Setophaga kirtlandii) experienced population declines as a result of jack pine habitat degradation and nest parasitism by the brown-headed cowbird. The population crash caused it to be listed under the 1973 Endangered Species Act. Even though the population is above the recovery goal, the existence of this species relies heavily on conservation management. Using whole genome sequencing of contemporary and historical samples of this previously endangered bird, we compare the distribution of runs of homozygosity as well as allele frequencies pre-bottleneck and post-bottleneck recovery to understand the long-term genomic consequences of reduced population size.

103 - Leveraging synteny to generate reference genomes for conservation: Assembling the genomes of Hector's and M?ui dolphins

Sebastian Alvarez-Costes

Sebastian Alvarez-Costes, C. Scott Baker, Rochelle Constantine, Emma L Carroll, Joseph Guhlin, Ludovic Dutoit, Sara Ferreira, Dorothea Heimeier, Neil J Gemmell, Joanne Gillum, Rebecca M Hamner, Will Rayment, Wendi D Roe, Lara Urban, Alana Alexander
University of Otago, University of Otago, Oregon State University, University of Auckland, University of Auckland, University of Otago, University of Otago, University of Otago, University of Auckland, University of Otago, University of Otago, Oregon State University, University of Otago, Massey University, Technical University of Munich, University of Otago

Escalating concern regarding the impacts of reduced genetic diversity on the conservation of endangered species has spurred efforts to obtain chromosome-level genomes through consortia such as the Vertebrate Genomes Project. Assembling reference genomes for many threatened species remains challenging due to the difficulty in obtaining optimal input samples, that can characterize long-term conservation collections. Here, we present a pipeline that leverages genome synteny to construct high-quality genomes for species of conservation concern despite less-than-optimal samples, demonstrating its use on Hector's and MÄ•ui dolphins. These endemic New Zealand dolphins are threatened by human activities. Hectorâ€™s dolphins are classified as endangered by the IUCN, while the MÄ•ui dolphin is among the most critically endangered marine mammals. To assemble reference genomes for these dolphins, we created a pipeline combining de novo assembly tools with reference-guided techniques, utilizing chromosome-level genomes of closely related species. The pipeline assembled highly contiguous chromosome-level genomes (scaffold N50: 110 MB, scaffold L50: 9, miniBUSCO completeness scores >96.35%), despite non-optimal input tissue samples and sequencing data. We demonstrate that these genomes can provide insights relevant for conservation, including historical demography revealing long-term small population sizes, with subspecies divergence occurring ~20kya, linked to the Last Glacial Maximum. MÄ•ui dolphin heterozygosity was 40% lower than Hector's and comparable to other cetacean species noted for reduced genetic diversity. Through these exemplar genomes, we demonstrate that our pipeline can provide high-quality genomic resources to facilitate ongoing conservation genomics research.

105 - Genetic management for an endangered carnivore: the spatiotemporal impacts of supplementation.

Andrea L Schraven

Andrea L Schraven, Catherine E Grueber, Elspeth A McLennan, Katherine A Farquharson, Andrew V Lee, Samantha Fox, Carolyn J Hogg
The University of Sydney, The University of Sydney, The University of Sydney, The University of Sydney, Department of Natural Resources and Environment, Department of Natural Resources and Environment, Toledo Zoo and Aquarium, The University of Sydney

Management of threatened species must be planned and executed, especially where prevailing threats, like disease, remain within the speciesâ€™ range. Supplementations (or population reinforcements) involve releasing conspecifics into an existing population, to alleviate small population pressures, i.e. inbreeding and genetic drift. The Tasmanian devil (Sarcophilus harrisii) has suffered dramatic declines since the emergence of a contagious cancer, devil facial tumour disease, which is unmitigated in the wild. The Save the Tasmanian Devil Program (STDP) is responsible for managing devils in the wild and have been trialling releases of individuals into wild and diseased sites since 2015. The aim of our study was to quantify changes in genetic differentiation among devil sites before and after supplementation and assess the impact of this action on 1) genetic diversity and 2) relatedness within each site over time. We used genome-wide SNPs from nine years (2014 to 2022) of sampling across eight devil sites, four supplemented and four not supplemented. We found a statistically significant decline in genetic differentiation between supplemented sites compared to between not supplemented sites. On average, genetic differentiation decreased by 11.99% among not supplemented sites while genetic differentiation among supplemented sites decreased by 45.29% between 2014 and 2022. We compared changes in genetic diversity and relatedness overtime between not supplemented and supplemented sites and found substantial variability among all sites. Our study sheds light on both the spatial and temporal genetic ramifications of reinforcing wild populations - a management strategy previously avoided where disease

presence cannot be mitigated.

183 - Genetic consequences of recent population mixing in an endangered marsupial, the Eastern barred bandicoot (Perameles gunnii)

John George Black

John G Black, Ary A Hoffmann, Tom L Schmidt, Andrew R Weeks
University of Melbourne, University of Melbourne, University of Melbourne, University of Melbourne

Population mixing aims to assist conservation of isolated or captive populations by increasing genetic variation. This genetic variation will be reduced by drift in the generations after mixing, but the extent of reduction is not well studied and may be species specific. Here, I examine the genetic outcomes of population mixing in a large, captive population of a specialist marsupial â€“ the Eastern barred bandicoot (Perameles gunnii). This population was established using animals from Victoria, Australia, shortly before the species became locally extinct. Declining genetic diversity in this population led to the implementation of assisted gene flow from a Tasmanian population, a divergence of ~10,000 generations. First, I present work on genetic structure and diversity in the Tasmanian source populations, finding widespread population structure over several hierarchical levels. Next, I reevaluate estimates of genetic diversity in the Tasmanian and Victorian populations using an improved pipeline for assessing autosomal heterozygosity. Finally, I compare changes in genetic diversity and morphometrics of the captive population over approximately eight generations of hybrid offspring, and I further contextualise results with comparison to another captive site that only contains Victorian animals.

193 - Inbreeding reLoaded: A genomic approach to characterize the genetic load in a small population of captive tigers

Vinay Sagar

Vinay Sagar, Shweta Ramdas, Uma Ramakrishnan
National Centre for Biological Sciences, Tata Fundamental of Fundamental Research, Centre for Brain Research, Indian Institute of Science, National Centre for Biological Sciences, Tata Fundamental of Fundamental Research

Small and isolated populations face extinction threats due to demographic, environmental, and genetic stochasticity. Additionally, inbreeding and subsequent inbreeding depression can harshly impact such populations. Inbreeding increases homozygosity in the genome, causing the accumulation and expression of recessive deleterious alleles. When expressed, accumulated deleterious variants (genetic load) can cause inbreeding depression. Historical studies have utilized captive animal populations to study inbreeding and inbreeding depression using pedigrees (FPED). Recent studies, primarily in wild populations, estimate inbreeding by examining the proportion of genome in long runs of homozygosity (FROH) and quantify inbreeding-associated genetic load using genomic data.

Captive tigers of Nandankanan Biological Park, Bhubaneswar (NKB), India, present an inbred population with only two founders and frequent mating between relatives. Using genomic data from NKB tigers, we show a correlation between FPED and FROH and observe that FROH likely accurately measures inbreeding. We also show that individuals with higher values of FROH carry more realized genetic load of putative loss of function (pLOF) alleles. Finally, utilizing gene loss of function databases in humans, mice, and other organisms, we identify several pLOF mutations that could have strong deleterious effects and contribute to inbreeding depression in NKB tigers. Our unique approach towards characterizing genetic load can help understand the fitness effects of mutations in endangered species despite small sample sizes. Such genomic analyses can be utilised for effective, well-informed management of captive populations. We also suggest future studies inventory frequencies of such curated pLOF mutations in wild populations for their conservation.

232 - Insights into Disease Control in Marine Fish Farms: Digital Droplet PCR-Based eDNA Tool for Monitoring Cryptocaryonosis

Hin Hung Tsang

Hin Hung Tsang, Jose A Domingos, Jacob A F. Westaway, Maximilian H Y. Kam, Roger Huerlimann, Giana Bastos Gomes
Department of Infectious Diseases and Public Health, Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong,, Simon F.S. Li Marine Science Laboratory, School of Life Sciences, The Chinese University of Hong Kong, Tropical Futures Institute, James Cook University Singapore, College of Public Health, Medical and Veterinary Science, James Cook University, 1/14-88 7 McGregor Road, Smithfield, Department of Infectious Diseases and Public Health, Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Centre for Tropical Bioinformatics and Molecular Biology, James Cook University, James Cook Dr, Douglas, College of Science and Engineering, James Cook University, Marine Climate Change Unit, Okinawa Institute of Science and Technology (OIST), Department of Infectious Diseases and Public Health, Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Temasek Life Sciences Laboratory, 1 Research Link, National University of Singapore

Ciliate parasites are of particular concern to the global fish production due to its complex life cycle and fast propagation in the culture environment without warning. The growth and proliferation of ciliate parasites was shown to be closely linked to bacterial pathogens, and influenced by environmental and physical-chemical parameters. Therefore, the development of an early detection system in correlation with water quality, environmental monitoring, and bacterial microbiome could help to prevent disease outbreaks in aquaculture farms.

We used Cryptocaryon irritans as a model to understand the parasite abundance and potential drivers of its presence in marine fish farms. We developed a novel droplet digital PCR (ddPCR) assay to quantify the absolute abundance of C. irritans from water samples using environmental DNA (eDNA) approach. This method allows for accurate and early quantification of parasite loads without its visible presence of infections. Environmental and physical-chemical data were analyzed together with C. irritans abundance and 16S metagenomics data to determine associations and triggers between parasites and specific bacterial groups. Rainfall and temperature demonstrated positive associations with C. irritans abundance as detected by ddPCR. Coraliomargarita was the bacterial genus with the most significant relationship with low abundance of C. irritans in water.

Understanding the environmental triggers of ciliate parasites propagation and associated bacterial microbiome can provide new insights into environmental control, microbial management, and promote the reduction of chemical use in marine fish farms. This important information can be adopted by environmental and governmental agencies for of new aquaculture sites and environmental protection.

238 - Genetic structure of Shorea curtisii and S. leprosula (Dipterocarpaceae) in Southeast Asia using MIG-seq

Misato Ogasahara

Misato Ogasahara, Rahayu Sukmaria Sukri, Faizah Metali, Kenzo Tanaka, Reiji Yoneda, Mohamad Alias Azani, Eyen Khoo, Melvin Gumal, Bibian Diway, Iskandar Z. Siregar, Koichi Kamiya
Asian School of the Environment, Nanyang Technological University, Institute for Biodiversity and Environmental Research, Universiti Brunei Darussalam, Faculty of Science, Universiti Brunei Darussalam, Japan International Research Center for Agricultural Sciences, Shikoku Research Center, Forestry and Forest Products Research Institute, Faculty of Forestry and Environment, Universiti Putra Malaysia, Sabah Forestry Department, Sarawak Forestry Corporation, Research, Development and Innovation Division (RDID), Forest Department Sarawak, Department of Silviculture, Faculty of Forestry and Environment, IPB University, Graduate School of Agriculture, Ehime University,

Matsuyama

In the last glacial period, it is thought that tropical rainforests were distributed widely on the exposed Sunda Shelf by lower sea level, but the savanna corridor interfered with east-west dispersal of rainforest species. Population genetic studies revealed a genetic divergence between â€œSumatra and the Malay Peninsulaâ€• and â€œBorneoâ€• for the Dipterocarpaceae, a major tree family in Southeast Asia. Divergence time and presence or absence of migration after divergence can differ between species, possibly reflecting restrictions on gene flow in the past. To understand details of vegetation in Southeast Asia in the past, population genetic analysis of NGS data for several species widely distributed in this region were utilized. In this study, MIG-seq was used to assess the geographic distribution of genetic variation in S. curtisii and S. leprosula. The analysis showed that a clear pattern of genetic structure was found in Bornean populations, but such a clear structure was not shown among populations in the Malay Peninsula. This suggests that the extent and distribution of refugial populations and population size changes differed between the two regions.

249 - Recent increase in species-wide diversity in the highly endangered Iberian lynx following introgression from Eurasian lynx

Jose A. Godoy

Maria Lucena-Perez, Johanna L. A. Paijmans, Francisco Nocete, Jordi Nadal, Cleia Detry, Love Dalen, Michael Hofreiter, Axel Barlow, Jose A. Godoy

EstaciÃ³n BiolÃ³gica de DoÃ±ana, CSIC, Department of Zoology, University of Cambridge, Departamento Historia I (Prehistoria), Universidad de Huelva, Departament de Prehistoria, Historia Antiga i Arqueologia, Universitat de Barcelona, UNIARQ â€“ Centro de Arqueologia da Faculdade de Letras da Universidade de Lisboa, Centre for Palaeogenetics, Evolutionary Adaptive Genomics, University of Potsdam, School of Environmental and Natural Sciences, Bangor University, School of Environmental and Natural Sciences, Bangor University, EstaciÃ³n BiolÃ³gica de DoÃ±ana, CSIC

Interspecific admixture events can increase genetic variation in the recipient speciesâ€™ gene pool, but empirical examples of species-wide restoration of genetic diversity by admixture are lacking. The Iberian lynx (Lynx pardinus) is one of the most endangered felids and one of the most genetically eroded species, due to small long-term population sizes and serial bottlenecks. The analyses of multi-fold coverage whole-genome data from three ancient Iberian lynx specimens dated ca. 2,000â€“4,000 ybp show lower diversity and lower introgression levels from Eurasian lynx (Lynx lynx) than contemporary samples. This suggests a continuous or recurrent process of interspecies admixture that increased modern Iberian lynx genetic diversity above that occurring millennia ago. Our results add to the accumulating evidence for widespread natural admixture and introgression, and show that this can result in a substantial increase in the standing genetic diversity in highly eroded species. The strict avoidance of interspecific sources in genetic restoration measures may need to be carefully reconsidered, particularly in cases where no additional conspecific source population exists, so a closely related species might be the only source of novel genetic diversity.

289 - Critically endangered Rice’s whale exhibits an unexpectedly complex demographic history

Diana Aguilar Gomez

Diana Aguilar Gomez, Jacqueline Robinson, Sergio Nigenda, Chris Kyriazis, Nicole Vollmer, Lynsey Wilcox Talbot, Patricia Rosel, Phillip Morin, Kirk Lohmueller

University of California Los Angeles, University of California San Francisco, California State University San Marcos, San Diego Zoo Wildlife Alliance, Cooperative Institute for Marine and Atmospheric Studies, Southeast Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Southeast Fisheries Science Center, National Marine Fisheries Service, National

Oceanic and Atmospheric Administration, Southeast Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, University of California Los Angeles

The Riceâ€™s whale (Balaenoptera ricei) was not formally recognized as a distinct species until 2021. They are the sole resident baleen whale species in the Gulf of Mexico, an area heavily impacted by industrialization. Riceâ€™s whales are highly endangered with an estimated population of fewer than 100 individuals. Principal threats to affect their population recovery include ship strikes, oil spills, climate change, noise, fisheries, and other human disturbances.
To support conservation efforts, we are conducting a population genetic study to assess the genomic health of these whales. We sequenced whole genomes of 20 individuals at ~30x coverage. Our findings reveal a highly inbred population, with most sequenced individuals showing significant relatedness and extensive runs of homozygosity, comprising approximately 40% of the genome. Furthermore, we also found that their demographic history is complex, suggesting drastic population fluctuations and/or genetic admixture to account for observed diversity patterns.
Specifically, we inferred the existence of a ghost population from which the Riceâ€™s whale diverged approximately 100,000 years ago, followed by a recent bottleneck. This bottleneck likely stems from anthropogenic causes, such as historical whaling. While cetaceans commonly exhibit introgression, no genetically close species near the Gulf of Mexico serve as viable candidates to validate this hypothesis, deepening the mystery. Information gathered from this genomic analysis will be integral for better understanding this rare and endangered species, and ensuring effective conservation and management planning.

294 - Genomics of Brazilian howler monkeys reveals adaptation to malaria

Katherine McVay

Katherine McVay, Erick Figueroa-Ildefonso, Katharine Korunes, Luana Portela, Victor Y GuimarÃ£es, Ligia S. L. Silveira da Mota, CauÃª Monticelli, JoÃ£o Valsecchi, Fernando A Perini, Alcides Pissinatti, Patricia Domingues de Freitas, Amy Goldberg
Duke University, , Duke University, Duke University, Federal University of SÃ£o Carlos, SÃ£o Paulo State University, SÃ£o Paulo State University, SÃ£o Paulo Wildlife Coordination, MamirauÃ¡ Sustainable Development Institute, Federal University of Minas Gerais, Primatology Center of Rio de Janeiro, Federal University of SÃ£o Carlos, Duke University

Malaria is one of the strongest selective pressures in human evolution, with dozens of genes implicated in adaptation to different Plasmodium parasites. Over 30 Plasmodium species infect non-human primates, including regular parasite sharing with humans. Plasmodium is hypothesized to have arrived in South America ~500 years ago through European Colonization and the Trans-Atlantic Slave Trade. Thus, South American primates provide an opportunity to understand adaptation in a malaria-naÃ¯ve primate system. In particular, howler monkeys (genus Alouatta) are the primary reservoirs for P. simium, a vivax-like zoonotic malaria infecting humans in southern Brazil, and regularly share P. malariae/brasilianum with humans. We conducted whole-genome sequencing on 88 howler monkeys (average coverage ~11X) to understand broad patterns of shared or unique adaptations to malaria between humans and multiple howler species. Based on 43 host genes known to interact with malaria (adapted from Ebel et al. 2017, PLoS Genetics), we find preliminary evidence for shared and species-specific variation and adaptation in genes associated with red blood cell structure and invasion using a combination of allele-frequency- and haplotype-based summary statistics. Additionally, we detect species-level infections with Plasmodium parasites in 14 howler individuals using KrakenUniq, with a higher proportion of infections in the Legal Amazon than other regions. We further reconstruct broader population structure and demographic history of 4 howler monkey species from across Brazil, finding differences in rates of inbreeding and population size by species.

343 - Cryptic diversity in Eurasian minnows: Implications for the conservation of German freshwater fishes

Temitope Opeyemi Oriowo

Temitope O Oriowo, Ioannis Chrysostomakis, Nils Sternberg, Astrid BÃ¶hne, Madlen Stange
Leibniz Institute for the Analysis of Biodiversity Change, Museum Koenig, Leibniz Institute for the Analysis of Biodiversity Change, Museum Koenig,, Leibniz Institute for the Analysis of Biodiversity Change, Museum Koenig, Leibniz Institute for the Analysis of Biodiversity Change, Museum Koenig,, Leibniz Institute for the Analysis of Biodiversity Change, Museum Koenig

Recent studies have revealed previously unknown cryptic diversity within Eurasian minnows (Leuciscidae â€“ Phoxinus phoxinus), challenging its monotypic classification in Europe (PalandaÄ•iÄ‡ et al., 2015, 2017). These findings have profound implications for conservation and taxonomy, particularly in Germany, where multiple Phoxinus species coexist. The recognition of multiple genetically distinct lineages within Phoxinus led to a revision of German freshwater conservation laws to reflect the conservation needs of each recognised species (Bundesamt fÃ¼r Naturschutz, 2023).
The existence of several species in Germany and the historical stocking and reintroduction programmes prompted us to understand the degree of differentiation and hybridisation between different Phoxinus species and populations, with a focus on the Middle Rhine. We conducted population structure analyses using our annotated, chromosome-resolved reference genome (Oriowo et al. 2024) and whole-genome datasets of over 300 Phoxinus individuals from various catchments in and around Germany. The analysis of more than 6 million SNPs revealed highly differentiated populations in various river catchments, such as the Meuse, Seine, and RhÃ´ne, with minimal to no hybridization with non-native species. In the Upper and Middle Rhine, we identified a distinct lineage within lineage '5b' (P. csikii), genetically distinct from the 'Danube lineage 5b'. In the Lower Sieg, a tributary to the Middle Rhine near Bonn that experienced high degrees of stocking, an analysis of hybridisation revealed a high degree of interbreeding between lineages '10' (P. phoxinus) and '5b of the Rhine type'. These findings underscore the importance of whole-genome analyses in conservation relevant species.

379 - Exploring the effect s of human overexploitation in the demographic history of a tropical sardine

Julia Tovar Verba

Julia Tovar Verba, Jessica Coelho, Sergio Lima, Ricardo J Pereira
LMU, UFRN, UFRN, Stuttgart State Museum of Natural History

Anthropic effects, such as resource exploitation, can impact populationsâ€™ abundance, possibly affecting its genetic diversity. Past demographic variability can leave genetic signs, allowing for the investigation of populationsâ€™ history through the analysis of modern samples. Tropical species are thought to be under more recent exploitation when compared to, for example, European species. Stronger exploitation started in the past two centuries, possibly not for long enough to leave an impact on demographic markers. However, short-living, fast-growing species can present signs of recent exploitation due to their short generation time. The small scaled-sardine Harengula sp. is a heavily exploited species endemic to Brazil. Previously, we demonstrated that the scaled-sardine shows signs of overexploitation. Additionally, we revealed that there are at least two isolated lineages, one occurring in a small archipelago, with a decreased population size after the split with the larger coastal population. We are currently using a whole-genome sequencing of individuals covering the whole species distribution, and linkage disequilibrium-based analysis to investigate the recent populationsâ€™ demographic history and test for the effect of human exploitation since early-colonization times. Because oceanic islands in the region were not inhabited by human populations until recently, we hypothesize that the island populations will present a smaller, but stable, effective population size, while the coastal population will be larger but showing signs of recent size contraction linked to fishery. These results will shed light on how tropical species are affected by overexploitation and which populations need stronger management

measures.

403 - Nearly-neutral theory in Asian Elephants and the conservation implications of deleterious mutation accumulation.

Gabe D O'Reilly

Gabe D O'Reilly, Rebekah Rogers, Michael Campana, Jesus Maldonado, Natalia A Prado-Oviedo, Jeroen Kappelhof
Department of Bioinformatics, UNC Charlotte, Department of Bioinformatics, UNC Charlotte, Center for Conservation Genomics, Smithsonian's National Zoo and Conservation Biology Institute, Center for Conservation Genomics, Smithsonian's National Zoo and Conservation Biology Institute, Biology Department, Adelphi University, Center for Conservation Genomics, Smithsonian's National Zoo and Conservation Biology Institute, Wageningen University & Research - Animal Breeding and Genomics

Nearly Neutral Theory states that a population's ability to purge a deleterious mutation is a function of the strength of selection and the effective population size of a population. The larger a population, the more readily it can purge deleterious mutations. However, under intense bottle necks populations can lose the ability to purge those mutations. In bottle-necked and fragmented populations, this could allow for slightly deleterious mutations to accumulate in the genome. These dynamics of nearly-neutral evolution have previously been documented in extinct elephantids. Modern populations of Asian Elephants have experienced parallel population declines and may be at similar risk of detrimental mutations. Deforestation, large infrastructure projects and regional conflicts have fragmented the Asian elephant population, leaving them endangered. Using short read genomic data from 100 Asian elephant genomes, we have investigated the population structure of Asian Elephants in South-to-South East Asia and are cataloguing their structural variation. With the aim to identify any accumulations of deleterious mutations and identify exactly how key genes are being altered by those mutations. These results will give us a much clearer idea of how the Asian Elephant population has been evolving during the Anthropocene and help clarify the circumstances under which genomic decay may occur, offering empirical bounds on evolutionary theory. Additionally, these genetic results can help steer conservation efforts for Asian elephants, a species of cultural and ecological importance in Southeast Asia.

414 - Population structure of two butterfly rays from Gymnura genus in the California Gulf using ddRADseq

Alexis Omar Tarula-Marin

Alexis O. Tarula-Marin, Nancy C Saavedra-Sotelo, Fausto Valenzuela-QuiÃ±onez, Ofelia Escobar-Sanchez
Facultad de Ciencias del Mar (FACIMAR) - Universidad Autonoma de Sinaloa (UAS), Facultad de Ciencias del Mar (FACIMAR) - Universidad Autonoma de Sinaloa (UAS), Investigadoras e Investigadores por MÃ©xico - CONAHCYT, Centro de Investigaciones BiolÃ³gicas del Noroeste S.C., Facultad de Ciencias del Mar (FACIMAR) - Universidad Autonoma de Sinaloa (UAS), Investigadoras e Investigadores por MÃ©xico - CONAHCYT

The California Gulf (CG) is a closed sea with a high environmental variation, thus is a region candidate to present ecologically divergent natural selection in many marine species. Two species of butterfly rays (Gymnura marmorata and Gymnura crebripunctata), cohabit in the GC, both present morphology similitud, thus are usually confused. Accordingly, we use population genomics to assess the phylogeographic break between species. Add to this, the intraspecific genomic variation was evaluated. We use the technique ddRADseq to obtain SNPs from samples of both species obtained across the CG and the Pacific Ocean. To assess the phylogeographic break, we obtained 768 SNPs for both species, which showed a phylogeographic break in the south coast of Sonora. This pattern can reflect an adaptive divergence associated with bathymetry, chlorophyll, and temperature in the GC. The preliminary intra-species genomic structure suggests a genomic structure in G. marmorata between localities from the CG and the Pacific, due to the Baja California peninsula acting

as a geographic barrier; while G. crebripunctata shows divergence between the mouth and center of the CG, this could be due to changes in the oceanographic conditions between zones and/or fidelity to reproductive sites. These results can be used as basic biological information for the design of fisheries management and conservation strategies in the GC.

435 - Structural genomic variation in the inbred Scandinavian wolf population contributes to the realized genetic load but is positively affected by immigration

Linnéa Smeds

LinnÃ©a Smeds, Lars SA Huson, Hans Ellegren

Uppsala University, Penn State University, Uppsala University, Uppsala University

The Scandinavian wolf population consists of 500 individuals, and is threatened by large hunting pressure and governmental aims to further decrease the population size. Due to a small number of founders and geographical isolation, Scandinavian wolves are extremely inbred, and a previous study using SNP data have shown that the realized genetic load has increased over generations since the founding event. In this study we expand our work to structural variants (SVs), an under-explored type of mutations potentially affecting a larger fraction of the genome than SNPs.

We analyzed whole-genome, short-read sequences from 212 wolves sampled in Scandinavia and from neighboring populations in Finland and Russia. More than 26,000 high-confidence variants were detected after stringent filtering and manual curation. The majority of variants were shorter than 1â€‰kb, with a distinct peak in the length distribution at 190â€‰bp, corresponding to insertion events of SINE/tRNA-Lys elements. The site frequency spectrum of SVs in protein-coding regions was significantly shifted towards rare alleles compared to putatively neutral variants, consistent with purifying selection. The realized genetic load of SVs in protein-coding regions increased with inbreeding levels in the Scandinavian population, while immigration provided a genetic rescue effect by lowering the load and reintroducing ancestral alleles at loci fixed for derived SVs. Our study shows that structural variation comprises a common type of in part deleterious mutations in endangered species and that establishing gene flow is necessary to mitigate the negative consequences of loss of diversity.

490 - Counting invisible elephants. Using non-invasive DNA sequencing approaches to monitor wild elephants.

Andrew J Tighe

Andrew J Tighe

Pwani University, University College Dublin

Earthâ€™s remaining elephants are keystone species in their ecosystems, and are a crucial element in maintaining biodiversity and even restoring degraded landscapes. By selectively browsing certain plant species and spreading the seeds of others, elephants play an outsized role in shaping forest ecosystems. Effective conservation and management strategies for wild elephants require accurate information on population numbers and dynamics, inbreeding potential, disease risks, parasite loads and diet. Increasingly there is also a push to be able to obtain such information non-invasively to avoid unnecessary stress to the target animals. This study highlights how eDNA and scat sampling can provide such information, with a specific focus on a population of savannah elephants (Loxodonta africana) living in Arabuko Sokoke forest in Kenya. Arabuko Sokoke forest is the largest remaining fragment of East African dry coastal forest and contains an under studied population of elephants with an unknown population number. The very dense nature of the forest makes traditional counting methods impossible, and so this project will use a novel SNP based approach to generate a population size estimate for the elephants living in the forest, using DNA collected non-invasively from scat samples. By individually genotyping every elephant

which visits the one remaining waterhole during the dry season, this project will then be able to extrapolate the number of elephants surviving in the forest. Further the project will utilize DNA metabarcoding methods to uncover both the diet of the elephants and what parasites they carry.

497 - Past, present & future sea lamprey (Petromyzon marinus) conservation genomics

Eleana Karachaliou

Eleana Karachaliou, Evelien de Greef, Jaanus SuurvÃ¤li, Claudio MÃ¼ller, Colin J. Garroway, Margaret F. Docker
University of Manitoba, University of Manitoba, University of Manitoba, University of Manitoba, University of Manitoba, University of Manitoba

Sea lamprey, an ancient vertebrate, is of conservation concern across its native range despite it being an invasive pest in the Laurentian Great Lakes. As an anadromous species, it is subject to habitat degradation, barriers to migration, and the increasing effects of climate change. We aim to inform conservation efforts by using genome-wide information to account for past and present effective population size and future vulnerability under climate change. We re-sequenced whole genomes for 155 sea lamprey from locations across their entire native range. We estimated the trajectory of effective population size in the recent past for each genetically distinct population and per-site contemporary effective population size. We then calculated genetic offsets to predict maladaptation to future climate change. All sea lamprey populations were found to have experienced severe effective population size declines in the recent past. Contemporary effective population size estimates identified locations of conservation concern. Genetic offsets forecasted locations at which sea lamprey are most likely to be maladapted and therefore more vulnerable under future climate change. Using these complementary methods, we make recommendations regarding conservation priorities and the potential for implementing genetic rescue as a conservation measure. This study shows the strength of applying population genomic methodology across different evolutionary timescales for the entirety of a speciesâ€™ range.

512 - Whole genome sequencing of Greenland muskox (Ovibos moschatus) individuals using surface-deposited skeletal elements

Mijin Park

Mijin Park, Juhyeon Lee, Donghee Kim, Raffaela Bianco, Johannes Krause, Tae-Yoon S. Park, Won Young Lee, Choongwon Jeong
School of Biological Sciences, Seoul National University, School of Biological Sciences, Seoul National University, School of Biological Sciences, Seoul National University, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Division of Glacier & Earth Sciences, Korea Polar Research Institute, Division of Glacier & Earth Sciences, Korea Polar Research Institute, School of Biological Sciences, Seoul National University, Institute for Data Innovation in Science, Seoul National University

The muskox (Ovibos moschatus) is a large-bodied ungulate species adapted to the tundra climate in Canada and Greenland. The current natural population in Greenland is assumed to have expanded from the Canadian Arctic Archipelago to north and northeast Greenland approximately 2,000 years ago through a serial founder event. However, their initial expansion process to Greenland as well as subsequent gene flow has not been thoroughly studied due to limited accessibility to muskox samples suitable for genome sequencing. In this study, we collected surface-deposited muskox skeletal elements from Ella Island in northeast Greenland and assessed their suitability for the non-disruptive source of genome information. Firstly, by comparing 13 libraries from 10 individuals, our analysis revealed that muskox horn shafts exhibit superb level of DNA preservation, making them easily accessible and high-quality targets for sampling in the field. Secondly, genomes of the newly sequenced Ella island individuals (n=9; 0.1-43.8x coverage) showed that they form

a clade with previously published individuals from nearby sites, supporting that the surface-deposited horn samples can yield high-quality genomic data. Finally, using the newly produced high-coverage genomes, we quantified the extreme reduction in genetic diversity over time in the Greenland populations through the Multiple Sequentially Markovian coalescent (MSMC) analysis. In summary, our study presents an accessible non-disruptive sampling strategy for muskox conservation genomics, along with high-quality whole genome sequences from eastern Greenland.

525 - Exploring the Bobcat Holobiome: Applications of Conservation Multiomics at the Wildland Urban Interface

Natalie Payne

Natalie Payne, Desiree Andersen, Robert Davis, Cheryl Mollohan, Kerry Baldwin, Albert L. LeCount, Koenraad Van Doorslaer, Leigh Combrink, Melanie Culver

School of Natural Resources and the Environment, University of Arizona, School of Natural Resources and the Environment, University of Arizona, Arizona GIS, LLC, Bobcats in Tucson Research Project Â©, Bobcats in Tucson Research Project Â©, Bobcats in Tucson Research Project Â©, Bobcats in Tucson Research Project Â©, Department of Immunobiology, University of Arizona, The BIO5 Institute, Cancer Biology Graduate Interdisciplinary Program, Genetics Graduate Interdisciplinary Program, UA Cancer Center, University of Arizona, School of Natural Resources and the Environment, University of Arizona, U.S. Geological Survey, Arizona Cooperative Fish and Wildlife Research Unit, University of Arizona, School of Natural Resources and the Environment, University of Arizona

Rapid advances in genomics in the 21st century provide higher resolution for answering questions related to wildlife conservation. However, conservation genomic studies focusing on host organisms alone provide incomplete pictures of biological processes affecting population viability. Characterization of interactions between individuals and within ecosystems of wildlife holobionts is crucial for understanding organismal health and can yield important insights into disease spread and environmental change.

Here, we describe the use of a multiomics dataset from a population of radiocollared bobcats (Lynx rufus) straddling the wildland urban interface (WUI) in the Tucson Mountains, Arizona, USA. We aimed to investigate how kinship and urbanization influence bobcat social dynamics and microbial communities.

Integrating paired host genomic and longitudinal GPS data uncovered extensive home range sharing of adult mother-daughter pairs. As with other felids, female bobcats typically maintain exclusive home ranges. Our results provide evidence of a more complex matrilineal social structure in this solitary felid.

Virome characterization has revealed numerous feline and canine pathogens with the potential to impact wild and domestic carnivore health at the WUI. Models of virome and microbiome (prokaryotic and eukaryotic) composition along with host location and genomic datasets reveal the roles of kinship and urbanization in shaping communities of host-associated taxa.

We highlight the importance of incorporating genomic research into field studies to investigate ecological hypotheses. With the current rate of anthropogenic ecosystem alteration, investigating population changes at a microbial level reveals early indicators of threats to wildlife population viability and the health of humans and domestic animals.

526 - Genetic variation, adaptation, and conservation of guenons across multiple natural environments

Adeola Oluwakemi Ayoola

Adeola Oluwakemi Ayoola, Amy Goldberg

Duke University, Duke University

Guenons (Cercopithecus), are the most widely dispersed non-human primates across Africa's tropical forests, exhibiting diverse taxonomic, phenotypic, and ecological behaviors. Here, we focus on guenons of Nigeria at the edge of the diverse Upper Guinea-Congolian forest. Multiple of these species are on the IUCN Red List resulting from bushmeat hunting, habitat loss, and fragmentation. Across three field seasons, almost one-third of bushmeat we tested belonged to guenon species, including Mona (31), Sclaterâ€™s (10), Red-eared (1), and White-throated guenon (4), emphasizing a role for hunting pressures in conservation. We next considered the parasite pressures shaping guenon genetic variation and conservation risk. As our close phylogenetic relatives, multiple monkeys act as a reservoir of and model for, malaria that can infect human populations. Using genomic data, we previously found a signature of adaptation and selection at the malaria-interacting gene G6PD, notably, only in a subset of populations. To understand the selective landscape and genomics of adaptation to malaria, we collected and sequenced ~400 fecal samples from multiple ecologiesâ€”tropical rainforest, mangrove, and Guinea savannaâ€”using a newly designed genus-specific capture. We tested for population-specific and shared signatures of selection and identified novel and shared variants across guenons and humans. Further, we tested for multiple malaria species to reconstruct the geographic range and natural occurrence of parasites compared to local human populations informing the risk for zoonoses. Importantly, this work is synergistic with educational components and community engagement towards building long-term conservation and public health initiatives.

553 - Genome connectivity of juvenile sharks Sphyrna lewini in potential nursery areas of the southern Gulf of California

Karla L. Velarde-Mota

Karla L. Velarde-Mota, Daniela G. FÃ©lix-LÃ³pez, Nancy C. Saavedra-Sotelo

Facultad de Ciencias del Mar, Universidad AutÃ³noma de Sinaloa, Division of Cardiovascular Sciences, Faculty of Biology Medicine and Health, The University of Manchester, Facultad de Ciencias del Mar, Universidad AutÃ³noma de Sinaloa, Investigadoras e Investigadores por MÃ©xico, CONAHCYT

The scalloped hammerhead shark Sphyrna lewini is a highly migratory top predator distributed in tropical and subtropical oceans worldwide. Unlike males, females are mainly distributed in coastal zones using mangrove estuaries as nursery areas (providing protection and food for pups). In many cases, females exhibit philopatry to these nursery grounds promoting population genetic structure. In the Gulf of California (GC) S. lewini is an important fishery resource, particularly there is a high abundance of juveniles and young-of-the-year (YOY) individuals. This work aims to evaluate the use of the GC as nursery areas for S. lewini. In this study, we collected muscle tissue samples from thirty-seven juvenile and YOYs from artisanal fisheries in four localities in the GC. We will assess the sibling relationships and adaptive variation through single nucleotide polymorphisms (SNPs) from the ddRADseq technique. Sibship analysis will be performed to evaluate the degree of connectivity between potential nursery areas. We will also evaluate the adaptive genetic structure as a proxy for residency from sharks within the GC. The results obtained may have implications for the conservation and management of S. lewini in the GC, due to a high degree of residency in nursery areas could indicate that the GC is a critical region for the survival of this species.

Key words: Population genomics, Philopatry, Residence, Local adaptation, Shark fishery.

585 - Prey in Peril: Impact of Habitat Alterations on two Large Herbivore species in Central Indian Landscape

Abhinav Tyagi

Abhinav Tyagi, Uma Ramakrishnan

National Centre for Biological Sciences, National Centre for Biological Sciences

In current times, habitat fragmentation and loss are the major threats to biodiversity worldwide. Reduction of habitat to smaller patches and increased distance between patches, lead to small isolated populations that pose an increased risk of loss of genetic diversity, inbreeding and genetic load. Maintaining gene-flow among these fragmented habitat patches is critical for long-term species persistence. Several natural and anthropogenic landscape features impede animal movement. Identifying these features is crucial for effective conservation planning. Genetic approaches aid in quantifying connectivity but most studies have focussed on a single species, usually, a large carnivore. Despite playing important role in ecosystem functioning and habitat maintenance, herbivores have been largely neglected. Here, we address genetic connectivity of two large herbivores, Gaur (Bos gaurus) and Sambar (Rusa unicolor) in central India. We collected fecal samples to generate genome-wide SNP data using ddRAD sequencing for 124 Gaur and 99 Sambar individuals. We demonstrate that gaur population in central India is fragmented and exhibit high genetic differentiation, especially in small populations like Umred Karhandala WLS. Although Sambar shows low genetic structure, small population in Bor Tiger Reserve exhibits slight differentiation. Our results suggest that although forest degradation and roads restrict animal movement, the extent of the impact varies with species' ecology. Our findings reveal that different species exhibit varied responses to various landscape features. We identify small and isolated populations that need conservation intervention. We opine a shift from large and charismatic species-focused conservation to a multi-species landscape conservation approach.

604 - Pedigree reconstruction informs conservation genomics: Inbreeding and mating patterns of a wild tiger population

BV Aditi Prasad

BV Aditi, Anubhab Khan, Kaushal Patel, Megan Aylward, Sampurna Roychoudhury, Uma Ramakrishnan

National Centre For Biological Sciences-TIFR, National Centre for Biological Sciences-TIFR, National Centre for Biological Sciences-TIFR, National Centre for Biological Sciences-TIFR, Indian Institute of Science Education and Research, National Centre for Biological Sciences, NA, National Centre for Biological Sciences-TIFR

Inbreeding and its fitness consequences could pose a threat to endangered species that exist in small and isolated populations. Inbred populations show varying temporal trajectories, with some populations declining or going extinct, whereas other populations survive over time. From a conservation genetic perspective, it is crucial to monitor the genetics of such populations on a temporal scale. Tigers have undergone extensive decline, and individuals in Northwestern India show genomic signatures of inbreeding and isolation. We conducted individual-based whole genome resequencing and analyses from 32 invasive and 32 non-invasive shed hair samples to i) understand the matrilineal contributions of founding females, ii) track inbreeding and relatedness over five generations; and iii) establish paternity to estimate the relatedness between mate pairs. Whole mitochondrial genomes reveal closely related matrilineal founders, with three matrilines surviving in current individuals. We find a small (but insignificant) increase in inbreeding over these five generations. Paternity assignments suggest that no mate pairs are first-degree relatives. Based on the genomics of non-invasive (and invasive) samples, we synthesize an unprecedented wild pedigree over five tiger generations and highlight potential inbreeding avoidance in an endangered species.

614 - Metagenomic analysis in tropical fresh and wastewater

Eugenia Zarza

Eugenia Zarza, Alexander LÃ³pez-Roblero, David J MartÃ-nez-Cano, Elia Diego-GarcÃ-a, Karina GuillÃ©n-Navarro, Pavel IÅ¡a

El Colegio de la Frontera Sur, Investigadora CONAHCYT, El Colegio de la Frontera Sur, El Colegio de la Frontera Sur, El Colegio de la Frontera Sur, Investigadora CONAHCYT, El Colegio de la Frontera Sur, Instituto de BiotecnologÃ-a, UNAM

Highly biodiverse tropical environments are threatened due to habitat transformation and climate change among other threats. Environmental DNA and RNA methods are promissing tools to uncover and monitor tropical ecosystems, but their implementation is still in early stages in these crucial environments. Here we performed a metagenomic analysis of fresh and wastewater from southern Mexico where we compared viral community composition. We detected different taxonomic profiles for viruses that reflect water use. We identified 38 virus families infecting a wide taxon range. We detected 15 different phytopathogenic viruses that belonged to the family Virgaviridae, genus Tobamovirus. Pepper mild mottle virus and tomato brown rugose fruit virus had the largest relative abundance. Applying phylogenetic methods on assembled sequences of the coat protein (CP) we inferred the origin of plant infecting viruses. We observed a close relationship between the solanaceous infecting viruses and the reference sequences found in GenBank. On the other hand, Fabaceae and Cucurbitaceae-infecting viruses did not show such a close relationship when compared with GenBank reference sequences. These results suggest that the detected viral sequences probably originated from new strains or species of local viruses likely from the consumption of locally cultivated produce, whereas those infecting solanaceous plants might come from consumption of processed or imported food to southern Mexico. This study highlights that global food and plant trade can influence local viral communities, especially in areas altered by human activities with the potential of affecting interactions within the ecosystems and the introduction of pathogens.

649 - The double-edged sword of genetic diversity in caribou

Rebecca S. Taylor

Rebecca S Taylor, Micheline Manseau, Sonesinh Keobouasone, Peng Liu, Paul J Wilson
Environment and Climate Change Canada, Environment and Climate Change Canada, Environment and Climate Change Canada, Environment and Climate Change Canada, Trent University

A major goal of conservation biology is to maintain high genetic diversity to ensure the resilience and adaptive potential of species given rapid global changes. However, high variation may also support deleterious alleles, or genetic load, increasing the risk of inbreeding depression if population sizes decrease. New sequencing technologies and analytical methods allow the measurement of genetic load in non-model organisms, and the purging of deleterious variation has been demonstrated in some threatened species. However, less is known about the costs of declines and inbreeding in species with large population sizes and high genetic diversity even though this encompasses many species globally that are expected to undergo declines. Caribou is a species of ecological and cultural significance in North America with a wide distribution and extensive phenotypic variation but with some populations undergoing significant declines resulting in their at-risk status in Canada. We assessed intra-specific genetic variation, adaptive divergence, inbreeding, and genetic load across populations with different demographic histories using a chromosome-scale reference genome and 66 whole-genome sequences. We found caribou to have high genetic diversity, comparable to the most diverse mammal species, and adaptive diversification of genes, but also high genetic load among lineages. We found highly divergent levels of inbreeding across individuals, including genomic erosion and the loss of alleles by drift but not increased purging in inbred individuals. Our results highlight the â€˜â€˜double-edged swordâ€™â€™ of genetic diversity that may be representative of other species at risk affected by anthropogenic activities.

665 - Mitochondrial Genetic Diversity and Phylogeny of Orangutans (Pongo sp.)

Viviani Mantovani Amador Cordon

Viviani Mantovani, Ana Agapito Vera, Irune Ruiz-Gartzia, Tomas Marques-Bonet, Michael KrÃ¼tzen
University of ZÃ¼rich, University of ZÃ¼rich, Universitat Pompeu Fabra, Universitat Pompeu Fabra, Universitat Pompeu Fabra, University of ZÃ¼rich

Orangutans (Pongo sp.) exhibit significant geographic variation in habitat and behavior, offering valuable insights into primate evolution. However, the impact of environmental and anthropogenic factors on their genetic diversity and differentiation remains underexplored. We investigated the mitochondrial genetic diversity and phylogeny of extant orangutans across the islands of Sumatra and Borneo. A total of 142 mitochondrial DNA genomes were analyzed, including 88 newly acquired ones, providing a comprehensive representation of orangutan distribution.

The analysis revealed four Sumatran and five Bornean clusters, with a clearly delineated separation between Bornean and Sumatran populations. The recently identified Sumatran species, P. tapanuliensis, forming a reciprocally monophyletic group with Bornean orangutans. Within Borneo, our analyses revealed a mismatch between the three recognised subspecies and the number of genetic clusters, warranting a future close examination of the demographic history of these groups based on autosomal data.

Our results contribute to identifying important taxonomic units that will aid in conservation efforts on preserving dispersal corridors, as well as protecting isolated populations to preserve genetic diversity and ensure the long-term survival of orangutans.

668 - Survival and extinction in small isolated populations

Anubhab Khan

Anubhab Khan
University of Copenhagen

Most of population genetics theory is based on the assumption of large panmictic populations. However, with the increase in study of population genomic datasets it is becoming clear that most populations either are or were small at some point in their history. Additionally, population structure leads to populations that get isolated from the main population that are often small, this is often the basis for allopatric speciation. Thus, small populations are the raw material for speciation. However, small populations are prone to impacts of drifts and inbreeding which makes them vulnerable to extinction. With whole genome sequences of several endangered populations of tigers, elephants and rhinoceros we show that small isolated populations often purge some of their deleterious allele loads. However, whatever deleterious alleles remain in the population, are in high frequency and homozygosity. This might lead the population to extinction. We further observe with simulations and emperical observations that migrations from a different population although increases the deleterious allele load, also leads to decrease in frequency of deleterious alleles and their homozygosity. Does this indicate that allopatric species also needs mild gene flow from other populations?

687 - The Role of Population History in Shaping the Mutational Load of Structural Variants Relative to SNPs, in Distinct Island versus Continental Lagopus Lineages

Gabriel David

Gabriel David, Patrik RÃ¶din MÃ¶rch, Jacob HÃ¶glund, Theodore Squires, Kristinn PÃ©tur MagnÃºsson
Uppsala universitet, Uppsala universitet, Uppsala universitet, Uppsala universitet, University of Akureyri

While the importance of specific demographic histories in shaping patterns of mutational load conferred by deleterious single nucleotide polymorphisms (SNPs) has received considerable attention in the recent past, few studies have investigated the corresponding fitness consequences of structural variation in distinct evolutionary lineages. We performed heuristic-based filtering and rapid automated curation of short-read-discovered SVs callsets from 99 re-sequenced individuals across two recently (~2 million years) diverged ptarmigan (Lagopus) species. High-confidence SV callsets reveal that the relative proportion of deleterious structural variants is

consistently greater in small effective population sizes, but that the relative frequency of deleterious variants differs between populations having experienced temporary bottlenecks versus longer-term low Ne. Despite the Svalbard rock ptarmigan population exhibiting the lowest coalescent Ne (Watterson's theta) estimate for all populations, it did not carry the highest masked and realised load for both SVs and SNPs. Crucially, these differences in genetic load may reflect differences in the type of bottleneck: the Svalbard population has likely experienced a founder event during post-glacial colonisation of the islands, while the Pyrenean rock ptarmigan population likely experienced a longer-term gradual decline with recent inbreeding. Our findings demonstrate the importance of considering the nuances of population history when interpreting the potential effects of small effective population size on mutation load. Furthermore, similar to SNPs, we find that many canonical SV classes (deletions, duplications and inversions) may largely conform to nearly-neutral expectations.

699 - Heavily fished Malawi cichlids rapidly adapt to anthropogenic pressures.

Alexander Hooft van Huysduynen

Alexander Hooft van Huysduynen, Hannes Svardal
University of Antwerp, University of Antwerp

In some remarkable cases, natural populations are able to persist through rapid evolutionary adaptation or phenotypic plasticity despite significant anthropogenic pressures. While ecological consequences are prioritized in conservation, the genomic impact of human-induced environmental change is often under-prioritised.
The Malawi cichlid radiation is renowned for its species richness and ecological diversity, but it is also a critical source of dietary protein and livelihood for millions of people. Over the past 40 years, exploitative fishing of some Malawi cichlid populations has led to local extinctions, dramatically reduced population size, and has also induced shifts in life history traits, such as in the Lake Malombe, where Copadichromis mloto has undergone a rapid reduction in size at maturity of 50% compared with only 20 generations ago.
To comprehensively explore the genomic signatures of this anthropogenically induced evolutionary adaptation, we used whole-genome resequencing data from 223 Copadichromis mloto samples. Our analysis identified several selective sweeps in the genomes of heavily fished populations and differential allele frequency patterns between heavily and weakly fished populations consistent with rapid evolutionary adaptation. Candidate genes were found to be related to life history traits such as gonad development and whole organism growth.
Our results demonstrate the genomic consequences of fisheries-induced evolution in a non-model study system and present signals of very recent adaptation to localized anthropogenic pressures. As WGS datasets grow, we will become more capable of investigating the genomic scars of anthropogenic pressures and their consequences for the future resilience of natural populations.

708 - Diversity patterns of estuarine fish communities across an urban gradient in Okinawa

Katerina Zapfe

Katerina L Zapfe, Alex Dornburg
UNCC, UNCC

Estuaries serve a wide-range of functions including natural environmental buffers and other ecosystem services. These ecosystems are also critical habitats for numerous marine species, including the juvenile stages of taxa critical to marine food webs and multi-billion dollar fishing and tourism industries. However, the persistence of taxa that utilize estuarine habitats is increasingly threatened globally by rapid environmental changes resulting from widespread coastal development and urbanization. Here we use environmental DNA (eDNA) to quantify the presence of marine fishes utilizing estuarine habitats along an environmental gradient in Okinawa, Japan that spans dense

human population centers to a UNESCO world heritage site. Specifically, we integrate data on species occurrence, water quality parameters, oceanographic data and human population data to assess trends in marine fish community and phylogenetic diversity. Our results reveal spatiotemporal trends across this landscape with critical implications for estuarine stewardship and the management of living marine resources.

722 - Regulatory divergences and ecological adaptation in Gopherus morafkai and G. agassizii

N. Jade Mellor

N. Jade Mellor, Kenro Kusumi, Greer A Dolby

University of Alabama at Birmingham, Arizona State University, University of Alabama at Birmingham

Mutations in different parts of genes can have very different effects on phenotypes and can be important for understanding how species adapt to local environments. Changes in regulatory regions of genes can shape time-dependent cellular and behavioral processes that underly pre and post-zygotic reinforcement mechanisms. Gopherus agassizii and G. morafkai are two recently speciated desert tortoise lineages that reside in adjoining but seasonally distinct deserts. The species are differentially adapted to rainfall and brumation patterns and have different disease instances. We used whole genome sequencing of 21 individuals to characterize differences between the two species. Eighty percent of mutations found in the 1% most diverged portion of the genome mapped to intergenic regions. Mutations within genic regions were largely found in promoter regions. Gene enrichment of promoter changes revealed UV nucleotide excision repair and the circadian/circannual regulation of gene transcription to be enriched. Changes to the regulation of these genes could affect the timing of biological functions related to key reproductive behaviors like brumation, mating, and laying, reinforcing isolation. Genomic regions with high exonic divergence included mucosal genes. Further investigation found extensive expansion of Mucin 5 genes with differential duplication, pseudogenization, and retention between the two lineages. Mucosal genes are important in osmoregulation and host defense against bacterial pathogens. Overall, we show regulatory divergence could underly transcriptional cascades that underlie the offset of reproductive behaviors between the species, and gene expansion that could play a role in the differential disease response among these species of conservation concern.

745 - Developing new strategies to characterise genomic variation for biodiversity conservation

J. Camilo Chacón-Duque

J. Camilo ChacÃ³n-Duque, Karin NorÃ©n, Tomas Marques-Bonet, Love DalÃ©n

Centre for Palaeogenetics, Department of Zoology, Stockholm University, Department of Zoology, Stockholm University, Institut de Biologia Evolutiva, CSIC-Universitat Pompeu Fabra, InstituciÃ³ Catalana de Recerca i Estudis AvanÃ§ats (ICREA), Centre for Palaeogenetics, Department of Zoology, Stockholm University

Genetic diversity is one of the three basic elements of biodiversity and understanding the full extent of this diversity within species is fundamental for their conservation. However, despite the ongoing advances in DNA sequencing technologies, conservation genomics is lagging behind in using whole-genome data efficiently. In contrast, other fields, such as human genomics, have developed theoretical and computational approaches to tackle this. If applied to endangered species, these approaches could provide a better understanding of their genetic diversity, particularly regarding the structure of this diversity across populations and its effect on adaptation to changing environments.

By leveraging two types of genetic data currently overlooked in conservation genomics â€“ haplotype information and structural genomic variation, we are developing new analysis strategies to characterise genomic variation in endangered species. Using the Arctic fox as a

model species, we are using whole-genome data from both present-day and historical samples to: 1) infer subtle patterns of population structure and gene flow, 2) quantify temporal changes in genomic structural variation and 3) assess the ability of these types of genetic data to inform on adaptation to local environments.

With these strategies, we expect to provide not only a comprehensive characterisation of genomic diversity in Arctic foxes â€“ with a direct impact on the Swedish Arctic fox conservation programme, but also new analytical tools for conservation genomics, unlocking the full potential of whole-genome data for biodiversity conservation.

778 - Contrasting levels of genetic diversity in two endemic axolotl species from Michoacan Mexico.

Esther Denice de la Cruz Santos

Denice de la Cruz-Santos, Eria A Rebollar, Yurixhi Maldonado-LÃ³pez, Maria D Basanta, Anahi Canedo-TÃ©xon, Eugenia Zarza, Gabriela Castellanos-Morales

El Colegio de la Frontera Sur, unidad Villahermosa, Centro de Ciencias GenÃ³micas, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de Investigaciones sobre los Recursos Naturales, Universidad Michoacana de San NicolÃ¡s de Hidalgo, Biology Integration Institute, University of Nevada Reno, El Colegio de la Frontera Sur, unidad Villahermosa, El Colegio de la Frontera Sur, unidad Tapachula, Consejo Nacional de Humanidades, Ciencias y Tecnologias, El Colegio de la Frontera Sur, unidad Villahermosa

Mexico is a center of diversity for axolotls (Ambystoma sp.) and most of the species are endemic to our country and are considered endangered species. In particular, obligate neotenic species, which never transition through metamorphosis, are highly threatened since their distributions are restricted to only one lagoon per species. Thus, exploring their genetic diversity is relevant to propose conservation measures. In the present study, we used a genotyping by sequencing (GBS) approach to compare the levels of genetic diversity and genetic differentiation between two critically endangered species, Ambystoma andersoni (N = 30) and Ambystoma dumerilii (N = 15), which are endemic to the Lago de Patzcuaro and Laguna Zacapu, respectively. We performed a reference-based SNP calling, using the available genome of A. mexicanum to map our reads, and retained 28,816 SNPs after filtering. We found that each species constitutes a well-defined genetic cluster without evidence of gene flow. Ambystoma dumerilli showed higher levels of genetic diversity (HE = 0.278) than A. andersoni (HE = 0.097). Lack of genetic diversity in A. andersoni raises awareness about the urgent need to conduct in situ and ex situ conservation action to raise levels of genetic variation and protect this species. For A. dumerilii habitat protection and restoration is needed to promote its permanence in the mid to long term.

844 - How many babies in the bush? Reevaluating species diversity and evolutionary history of western dwarf galagos (Galagoides) in Africa’s Guineo-Congolian forests.

Anna Penna

Anna Penna, Logan Kistler, Luca Pozzi

National Museum of Natural History, Smithsonian Institution, University of Texas at San Antonio, Lund University, National Museum of Natural History, Smithsonian Institution, University of Texas at San Antonio

Understanding evolutionary processes is critical to elucidate biodiversity patterns in species-rich regions like Africa's Guineo-Congolian forests. The discontinuities between lowland rainforest habitats in western Africa by the open and dry Dahomey Gap, raises questions about how historical environmental changes have influenced regional species richness and endemism. However, uncertainties in species boundaries and distributions hinder efforts to assess the impact of these changes. Western dwarf galagos (genus Galagoides), a clade of broadly distributed nocturnal primates, exemplify this challenge due to unclear morphological diagnoses.

We present a comprehensive molecular phylogenetic analysis of western dwarf galagos, leveraging genomic data from over 130 museum dry skin samples covering the entire geographic range of the genus. Genotypic clustering analyses based on tens of thousands of nuclear loci reveal two broadly distributed lineages occurring sympatrically across the sub-Saharan rainforests. These lineages exhibit minimal genetic admixture and remarkably similar patterns of population substructure that potentially correspond to cryptic species. Spatial transitions between these clusters occur around the Sanaga River Basin, in eastern Nigeria and northern Cameroon, suggesting that shared historical landscape changes have shaped similar patterns of population structure across species. We also found differentiation in cranial traits between major genetic lineages, emphasizing the need for multiple lines of evidence to delimit evolutionary lineages.

Our study highlights the value of historical DNA in clarifying species boundaries and distributions, particularly for species listed as data deficient in the IUCN red list. Moreover, our approach provides insights into biogeographic processes shaping biodiversity in Africa's lowland rainforests and beyond.

880 - A paleogenomic approach to reconstruct historical responses of coral reefs to anthropogenic change

Raul A Gonzalez-Pech

Raul A Gonzalez-Pech, Stephanie Marciniak, Colin Howe, Raphael A Eisenhofer-Philipona, Sofia Roitman, Julia E Cole, Laura Weyrich, Mateo Lopez-Victoria, George Perry, Monica Medina

The Pennsylvania State University, McMaster University, The Pennsylvania State University, University of Copenhagen, New England Biolabs, University of Michigan, The Pennsylvania State University, Pontificia Universidad Javeriana, The Pennsylvania State University, The Pennsylvania State University

Coral reefs are declining worldwide due to anthropogenic-driven environmental change. The foundation and health of these ecosystems rely on the harmonic functioning of all members of coral holobionts, i.e., cnidarian host, symbiotic microalgae (families Symbiodiniaceae and Ostreobiaceae), and associated microbiome. Coral stress responses often involve shifts in the taxonomic identity of their symbionts and microbiomes. Tracing back changes in coral holobiont composition over prolonged time periods can help us reconstruct health history of reefs and gain a better understanding of coral response to current stressors. Here, we focused on a major Caribbean reef-builder coral, Orbicella faveolata, from the Varadero Reef, Colombia. This reef has undergone extensive freshwater sediment discharge for decades as result of urbanization. We show, for the first time, that a paleogenomic approach can be used to reconstruct historical, microbial shifts of coral holobionts, potentially associated with stress.

936 - Using comparative transcriptomics to understand response to heat stress in polar fishes

Tina Cai

Tina Cai, Samuel N Bogan, Joanna L Kelley

University of California, Santa Cruz, University of California, Santa Cruz, University of California, Santa Cruz

Anthropogenic climate change is one of the most pressing issues globally, increasing risks of species extinction and biodiversity loss. While the polar and sub-polar regions are experiencing an accelerated impact due to declining sea-ice cover and warming, species that have adapted to these extremely cold and stable environments might be the most susceptible to the drastic changes. Many fish species have adapted to live in the subfreezing water temperatures, with adaptations including the ability to produce antifreeze proteins and the loss of an inducible heat shock response. There is currently a limited understanding of how polar fish species will respond to increased

climatic disturbances. Here, we conduct a reanalysis of existing transcriptome studies to understand how polar fishes and these specialized functions respond to elevated heat stress. We compare gene expression among different fish species under varying levels of heat exposure across life stages, tissues, and types of antifreeze proteins. Transcriptomic data provides an opportunity to investigate whether gene expression responses to heat stress are conserved across polar fish lineages. We expect to observe a difference in sensitivity to heat stress across tissue types depending on the magnitude and duration of thermal exposure. Activation of genes linked to DNA repair and metabolism response may be essential for the preservation of cellular functions across lineages. Our findings will elucidate the mechanisms by which polar fishes respond to heat stress, contributing to a better understanding of their capacity to adapt to rising global temperatures.

1044 - Decoding the population structure and history of the world's deadliest cat: (Felis nigripes) the black-footed cat

Victoria B Grant

Victoria B Grant, Shreya Banerjee, Jonas Oppenheimer, Katherine A Solari, John J Baczenas, Beth Shapiro, Dmitri Petrov, Ellie E Armstrong, Molly Schumer

Stanford University , University of California, Davis, University of California, Santa Cruz, Stanford University , Stanford University, University of California, Santa Cruz, Stanford University, Washington State University, Stanford University

Advances in genomic sequencing have opened new opportunities for studying speciesâ€™ genetic diversity and evolutionary past to address conservation concerns. Listed as vulnerable to extinction by the IUCN Red List, black-footed cats (Felis nigripes) are one of Africaâ€™s rarest cats. The only cat residing in the arid and semi-arid parts of southern Africa, very little is understood about black-footed cats because of their small size and nocturnal behavior. Moreover, black-footed cats live in isolation and their intraspecific interactions and demographics are not well described. In such cases where ecological knowledge is lacking, range-wide population genomics can improve our understanding of population dynamics.

Here, we sequenced whole genomes of black-footed cat individuals from across their speciesâ€™ range in the first genomic study of free-roaming individuals. To do so, we incorporated whole genomes generated from modern biological samples and century-old museum specimens. We assembled a highly contiguous reference genome using a combination of PacBio HiFi reads and publicly available Hi-C data for population analyses. We found evidence of larger populations in the past with a prolonged population decline in the last 10,000 years. Furthermore, we found remarkably low genome-wide diversity (Ï€ â‰ˆ 0.0004) suggesting a recent bottleneck. We compared individuals from across the range to evaluate patterns of population structure and found higher genetic similarity between individuals in close geographic proximity. Overall, these results provide a range-wide and long-term baseline informing the conservation of this enigmatic species.

1070 - North African fox genomes show signatures of repeated introgression and adaptation to life in deserts

Joana L Rocha

Joana L. Rocha, Pedro Silva, Nuno Santos, Monia Nakamura, Sandra Afonso, Abdeljebbar Qninba, Zbyszek Boratynski, Peter H Sudmant, Jose C Brito, Rasmus Nielsen, Raquel Godinho

Department of Integrative Biology and Department of Statistics, University of California Berkeley, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de VairÃ£o, Departamento de Biologia, Faculdade de CiÃªncias, Universidade do Porto,

BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de VairÃ£o, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de VairÃ£o, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de VairÃ£o, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de VairÃ£o, Institut Scientifique, Mohammed V University of Rabat, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de VairÃ£o, 4Department of Integrative Biology and Department of Statistics, University of California Berkeley, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de VairÃ£o, Department of Integrative Biology and Department of Statistics, University of California Berkeley,, 7Globe Institute, University of Copenhagen, BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de VairÃ£o

Elucidating the evolutionary process of animal adaptation to deserts is key to understanding adaptive responses to climate change. Here we generated 82 individual whole genomes of four fox species (genus Vulpes) inhabiting the Sahara Desert at different evolutionary times. We show that adaptation of new colonizing species to a hot arid environment has probably been facilitated by introgression and trans-species polymorphisms shared with older desert resident species, including a putatively adaptive 25â€‰Mb genomic region. Scans for signatures of selection implicated genes affecting temperature perception, non-renal water loss and heat production in the recent adaptation of North African red foxes (Vulpes vulpes), after divergence from Eurasian populations approximately 78â€‰thousand years ago. In the extreme desert specialists, Rueppellâ€™s fox (V. rueppellii) and fennec (V. zerda), we identified repeated signatures of selection in genes affecting renal water homeostasis supported by gene expression and physiological differences. Our study provides insights into the mechanisms and genetic underpinnings of a natural experiment of repeated adaptation to extreme conditions.

1083 - New model of hybrid incompatibility by metabolic control reveals the origin of overdominance and inbreeding depression: antecedent of Darwin's corollary.

Francisco García Franco

Francisco GarcÃ-a Franco, Lilian Milena Barandica CaÃ±on, Ezel Jacome Galindo-PÃ©rez, Blanca Estela ChÃ¡vez Sandoval
Universidad AutÃ³noma Metropolitana, Universidad de Cartagena, Universidad AutÃ³noma Metropolitana, Instituto PolitÃ©cnico Nacional

With the purpose of contributing to the understanding of the mystery of mysteries, how species arise, for more than 100 years, hybrid crossings between pairs of flies from different sister species of the Drosophila genus have been the fundamental strategy of several studies. evolutionary studies to identify speciation genes and their processes in the genetic architecture of reproductive isolation.
To interpret the evolutionary process based on the results of these studies, the fundamental theoretical model, Bateson-Dobzhansky-Muller or BDM, of hybrid genetic incompatibility between pairs of complementary genes under the control of Mendelian dominance has been developed.
Although the BMD model has been the dominant model for more than 100 years to interpret the genetic architecture of reproductive isolation and the speciation process.
Alternatively, models of genetic dominance under metabolic control have been proposed, which allow a more direct and mechanistic interpretation of the mechanisms underlying reproductive isolation.
In this work, we develop a model of the genetic mutation buffering system (SAGM), which reveals how recessive and loss-of-function allele mutations are integrated into the genome, leading to an increase in the genetic variability of populations and how New genotypes with particular dominance and their corresponding new phenotypes are generated, which reveals the origin of overdominance and inbreeding depression.

93 - Active Learning Experience using an analogy with the lac Operon model (PAPIME PE216224)

Marco Antonio Carballo-Ontiveros

Marco Antonio Carballo-Ontiveros, AmÃ©rica Nitxin CastaÃ±eda-SortibrÃ¡n, Marcos Nahmad-Bensusan
Facultad de Ciencias, UNAM, Facultad de Ciencias, UNAM, Departamento de FisiologÃ-a, BiofÃ-sica y Neurociencias, CINVESTAV

The analogy of the lac operon with a vending machine is intriguing. Analogies can be powerful tools for teaching complex concepts by relating them to something more familiar. The analogy to a vending machine might help students understand the concept of gene regulation and expression more concretely. To test this analogy, we designed an active learning class, which included a Learn-Before-Lecture (LBL) activity, an in-depth study of some aspects taught by the subject teacher, and, as a central strategy, the analogy of the lac operon with a coffee vending machine, which was accompanied by activities in which students submitted their answers digitally. In this regard, our students showed astonishment at the change in class dynamics, presenting resistance to answering a questionnaire associated with the LBL activity. On the other hand, a favorable response was observed to the analogy of the lac operon and a coffee vending machine. The activities of this analogy were solved quickly and favorably by the students. In our experience, implementing an analogy to an abstract topic, such as the lac operon, within an active learning class allowed students to assimilate knowledge quickly and enthusiastically.

132 - From evolution to solution: a solutions manual and e-Learning portal for Ziheng Yang's books "Computational Molecular Evolution", and "Molecular Evolution: A Statistical Approach"

Sishuo Wang

Sishuo Wang, Jianhao Lv
The Chinese University of Hong Kong, Nanfang College of Sun Yat-Sen University

Molecular evolution and phylogenetics have rapidly become crucial domains in biology in the past decades, yet pose a high entry barrier due to their interdisciplinary nature, encompassing molecular biology, evolutionary biology, statistics, computer science, etc. Ziheng Yang's books, "Computational Molecular Evolution" (http://abacus.gene.ucl.ac.uk/CME/) and "Molecular Evolution: A Statistical Approach" (http://abacus.gene.ucl.ac.uk/MESA/), stand as cornerstone references in this interdisciplinary domain, offering not only comprehensive content but exceptional exercises covering proofs/computations, programming, and software operation. However, no solutions to the problems are available, posing significant challenges for biologists to further understand the problems and the books.

We present an over 140-page comprehensive solutions manual for all the 75 exercises included in Ziheng Yang's books. This manual provides detailed, step-by-step derivations, along with extensive code examples and datasets for hands-on practice. Additionally, it provides alternative solutions/derivations where applicable, incorporating state-of-the-art methods to tackle some of the "classic" problems, enabling readers to experience the evolution of phylogenetic methods. To facilitate engagement and collaboration, we have established an online platform hosted on GitHub (https://github.com/sishuowang/Solutions_Manual_CME2006_MESA2014), inviting feedback, alternative solutions, and error reporting from the community. Embracing e-learning principles and a flipped classroom approach, this platform empowers learners to drive their learning journey actively. This work has never received any financial support, but we believe that our initiative will catalyze a community-driven effort to help researchers and students from various disciplines interested not only in Yang's seminal works but also in computational molecular evolution and phylogenetics.

482 - NoRCEL – Advancing Interdisciplinary Research and Global Engagement in Life Sciences

Miryam Palacios-Pérez

Miryam Palacios-PÃ©rez, Sohan Jheeta
Universidad Nacional AutÃ³noma de MÃ©xico, Network of Researchers on the Chemical Emergence of Life, Network of Researchers on

the Chemical Emergence of Life

The Network of Researchers on the Chemical Emergence of Life (NoRCEL) is an interdisciplinary collaborative network focused on understanding the origins and evolution of life from a chemical perspective. Founded by Sohan Jheeta, NoRCEL aims to bring together scientists from various disciplines, including chemistry, biology, astronomy, and Earth sciences, to investigate one of the most fundamental questions in science: how life began. In recent times, NoRCEL has successfully organised annual debate forums on the other side of human evolution, its own extinction and that of other species due to overexploitation of natural resources. Finally, looking for preservation of human health, a few NoRCEL scientists have examined the microbiome and its evolution. In short, like the community's motto, NoRCEL is about: Emergence, Evolution, Extinction.
NoRCEL fosters collaboration irrespective of geographical boundaries. A cornerstone of NoRCEL's mission is supporting early-career scientists and students, particularly from underrepresented regions, in astro-sciences and life emergence studies. By connecting emerging researchers with established scientists, the network aims to nurture the next generation of interdisciplinary scholars. Overall, NoRCEL endeavours to advance understanding of life's origins while promoting collaboration, inclusivity, and public outreach in the scientific community. Through colloquiums, workshops, and conferences, it facilitates the exchange of ideas and findings, providing platforms for researchers and students to present work and forge new partnerships. Moreover, NoRCEL prioritizes public engagement, making the study of life's chemical evolution accessible through lectures, publications, and online resources.
Gradually, NoRCEL will be transformed from only an academic community into a scientific institute for education.

547 - Climate change and ESKAPEE bacteria in AR and VR

Katia Pamela Pamela Villavicencio Sánchez

Katia Pamela Villavicencio SÃ¡nchez, Jonathan E. GÃ³mez GonzÃ¡lez, Diana Corina CeapÄƒ, Daniel Finkelstein-Shapiro
UNAM, UNAM, UNAM, UNAM

Antimicrobial resistance (AMR) is a serious global public health problem that is responsible for around 5 million deaths per year. The growth of AMR has accelerated in the last few decades due to several factors, including the inappropriate use of antibiotics in both human and animal health, the transfer of resistant bacteria from animals to humans, and the contamination of water sources with antimicrobial drugs and their residues. This issue is growing silently, and poor infection prevention and control practices further promote its spread.

Climate change (CC) has often been compared to AMR because it indirectly contributes to AMR. Higher temperatures and moisture are associated with higher rates of bacterial growth and horizontal gene transfer, which means bacteria reproduce and transmit faster, making them more dangerous and increasing resistance. The loss of efficacy of antibiotics makes the treatment of infections such as pneumonia, tuberculosis, gonorrhea, and salmonellosis harder.

It is, therefore, essential to develop a broader awareness of the serious problem of antimicrobial multidrug resistance and how it is affected by CC, caused by the improper disposal of drugs contaminating the ecosystem. It is also necessary to understand the various morphologies of the bacterial genera with the greatest resistance to antibiotics currently known. Creating interactive models of cellular and molecular structures of various multidrug-resistant bacterial pathogens, to be observed in AR and VR, along with information cards to explain the diseases they cause, will help to generate a conscious education, and prepare us for the future against this silent pandemic.

873 - PhyloGenome: a gamified and participatory approach to genomic education

Sònia Casillas

SÃ²nia Casillas, Noah Jornet-Calomarde
Department of Genetics and Microbiology, Universitat AutÃ²noma de Barcelona, Institute of Biotechnology and Biomedicine, Universitat AutÃ²noma de Barcelona, Department of Genetics and Microbiology, Universitat AutÃ²noma de Barcelona, Institute of Biotechnology and Biomedicine, Universitat AutÃ²noma de Barcelona

PhyloGenome is an innovative educational platform that integrates genomic science with the interactive elements of a collectible card game, fostering active learning and collaboration among students. It allows learners to become protagonists in their educational journey, applying theoretical concepts to real-world genomic sequencing, assembly, and annotation.

As part of the course, which complements traditional lectures, students choose a genome to investigate, culminating in the creation of an informative species card. This card is then added to the PhyloGenome deck, contributing to a growing educational resource. The rich game mechanics, inspired by the open-access Phylo game, combine scientific concepts with strategic gameplay, immersing players in their very own narrative of the genome sequencing era through a diverse array of species, event, and generation progress cards.

The PhyloGenome project is supported by a dedicated website (phylogenome.omicsuab.org), which acts as a central hub for educators and learners. It provides free access to the complete card set, game rules, glossaries, and teaching materials. The website also encourages a collaborative community by inviting educators to involve their students in contributing new cards to the project.

By distributing all materials under a Creative Commons license, PhyloGenome stands at the forefront of inclusive, participatory science education. It innovates genomic education, merging academic learning with gameplay to spark scientific curiosity and collaboration. Beyond educational outcomes, PhyloGenome aims to impact broader societal issues, such as environmental awareness, open education access, and community development, paving the way for a more sustainable future.

887 - PopLife: Population Genomics across the tree of life

Sònia Casillas

Cristina Amor-JimÃ©nez, Noah Jornet-Calomarde
Department of Genetics and Microbiology, Universitat AutÃ²noma de Barcelona, Institute of Biotechnology and Biomedicine, Universitat AutÃ²noma de Barcelona, Department of Genetics and Microbiology, Universitat AutÃ²noma de Barcelona, Institute of Biotechnology and Biomedicine, Universitat AutÃ²noma de Barcelona

A vast amount of population genomic data in a growing number of species is currently being generated by the newest sequencing technologies. This data contains precious information about the evolutionary history of life on our planet. To help in the evolutionary interpretation of this information, we have launched the PopLife project, a free online resource for Population Genomics to describe and explain genetic variation within and between populations across any species in the Tree of Life.

PopLife focuses on extracting knowledge from genomic data, applying both â€œclassicalâ€• population genetics methodologies and techniques based on artificial intelligence. PopLife provides a highly performing pipeline specially designed to compute population genetics statistics for large-scale genomic datasets, such as variation and divergence metrics, linkage disequilibrium parameters, and neutrality tests. Its novel user-friendly genome browser interface allows interactive visualization and data retrieval. The available information is not only applicable to evolutionary genomics, but also to fields such as functional genomics, conservation genetics, genetic

breeding, pandemics risk, and countless applications in other research fields.

In addition to computing summary population genetics statistics genome-wide, PopLife aims to estimate ancestral recombination graphs (ARGs), a powerful approach that infers complete cis genealogies along genome sequences. This information can be used to evaluate the allele frequency trajectory and selection coefficient estimates from genealogies whether or not it includes ancient DNA data. We are building a collaborative network of researchers with complementary knowledge and skills to make PopLife an international reference resource in population genomics.

48 - Geo-genomic predictors of phylogeographic and introgression history in an Amazonian bird

Lukas J Musher

Lukas J Musher, Glaucia Del-Rio, Rafael Marcondes, Robb T Brumfield, Gustavo Bravo, Gregory Thom
The Academy of Natural Sciences of Drexel University, American Museum of Natural History, University of Florida, Rice University, Lousiana State University, Instituto de InvestigaciÃ³n de Recursos BiolÃ³gicos Alexander von Humboldt, Louisiana State University

Selection, drift, recombination, and gene flow interact to produce a genomic landscape of divergence wherein patterns of differentiation and genealogy vary nonrandomly across the genomes of diverging populations. For instance, genealogical patterns that arise due to gene flow should be more likely to occur on smaller chromosomes, which experience high recombination, whereas those tracking histories of geographic isolation (reduced gene flow caused by a barrier) and divergence should be more likely to occur on larger and sex chromosomes. In Amazonia, populations of many bird species diverge and introgress across rivers, resulting in reticulated genomic signals. Herein, we used reduced representation genomic data to disentangle the evolutionary history of four populations of an Amazonian antbird whose biogeographic history was associated with the dynamic evolution of a large river basin. Specifically, we evaluate whether a river capture event gave rise to reticulated genealogies in the genome by making spatially explicit predictions about isolation and gene flow based on knowledge about genomic processes. We modeled levels of divergence, introgression, and genealogical patterns as functions of chromosome length and used this information to reconstruct the complex spatial history of this species. Our results suggest that a significant portion of genomic heterogeneity arises due to extrinsic biogeographic processes such as river capture interacting with intrinsic processes associated with genome architecture. Future studies would benefit from accounting for genomic processes, as different parts of the genome reveal contrasting, albeit complementary histories, all of which are relevant for disentangling the intricate geogenomic mechanisms of biotic diversification.

127 - Phylogeographic analysis of the invasive Asian hornet, Vespa velutina, to reconstruct its colonization of Europe from Asia

Xuhua Xia

Xia C Xia
University of Ottawa

The Asian hornet, Vespa velutina, is an invasive species that has established successfully in multiple European countries and incurred financial costs by damaging local apiary. Did the colonization occur only once or multiple times from Asia to Europe? How did the population expand and differentiate after the initial colonization? Can ecological models successfully predict its habitat suitability in Europe and its biogeographic expansion? Is it possible to estimate the genetic variation of the founding queens? How can a geophylogeny help us visualize genetic variation over time and space and identify geographic barriers to gene flow? How would human activities facilitate the dispersal of the hornets? I present preliminary results to address these questions.

195 - The power of dispersal driven by ocean currents in explaining spatial genetic patterns of marine invertebrates around Aotearoa New Zealand

Charles Michie

Charles Michie, Ravinder Kanda, Alice Della Penna, Carolyn J Lundquist, Shane D Lavery

Oxford Brookes University, Oxford Brookes University, University of Auckland, University of Auckland, National Institute of Water and Atmospheric Research , University of Auckland

Larval dispersal and connectivity in the marine environment are important for biodiversity, ecosystem function and population well-being. Invertebrate marine species around Aotearoa New Zealand vary tremendously in factors that influence larval dispersal, including the amount of time larvae spend in the ocean and the time of year in which larvae are released. We have previously shown both of these to significantly impact larval dispersal. Here we examine the output of dispersal simulations depending on these factors with a genetic dataset of multiple species. We expect that genetic patterns should reflect those of larval dispersal if larval dispersal drives gene flow between populations. By comparing the results of larval dispersal simulations from the Lagrangian tracking software â€œOpenDriftâ€• with a nationwide circulation model with population genetic metrics, we assessed the power of dispersal driven by ocean currents in predicting genetic patterns. Across a broad range of species, we found limited linear correlation between dispersal and genetic metrics, which suggests that processes outside of oceanic currents alone impact genetic patterns. Using a variety of dispersal metrics, we estimated how many generations would be required to connect distant locations by dispersal only and which locations are sinks and sources. Locations on the east coast of the South Island were sources, while those in central locations were sinks. These findings highlight the use of a multidisciplinary approach in disentangling what factors are involved in causing contemporary genetic patterns and the general understanding of marine connectivity and dispersal.

265 - Serial founder effects slow range expansion in an invasive social insect

Thomas J Hagan

Thomas J Hagan, Guiling Ding, Gabriele Buchmann, Benjamin P Oldroyd, Rosalyn Gloag

University of Missouri St Louis, University of Sydney, Chinese Academy of Agricultural Sciences, University of Sydney, University of Sydney, University of Sydney

Invasive populations often experience founder effects: a loss of genetic diversity relative to the source population, due to a small number of founders. Even where these founder effects do not impact colonization success, theory predicts they might affect the rate at which invasive populations expand. This is because secondary founder effects are generated at advancing population edges, further reducing local genetic diversity and elevating genetic load. We show that in an expanding invasive population of the Asian honey bee (Apis cerana), genetic diversity is indeed lowest at range edges, including at the complementary sex determiner, csd, a locus that is homozygous-lethal. Consistent with lower local csd diversity, range edge colonies had lower brood viability than colonies in the range centre. Further, simulations of a newly-founded and expanding honey bee population corroborate the spatial patterns in mean colony fitness observed in our empirical data, and show that such genetic load at range edges will slow the rate of population expansion.

531 - Evolution of phenotypic plasticity owing to migration

Davorka Gulisija

Davorka Gulisija

University of New Mexico

Phenotypic plasticity is one of the pivotal mechanisms that allow populations to persist in the face of environmental change. For plasticity to evolve and persist, populations need to continuously experience selection either due to changing environments or an influx of maladaptive phenotypes. Continuously changing environments are limited to specific eco-zones and organisms, but not universal, making immigration a likely contributor to the wide prevalence of phenotypic plasticity in nature. However, the conditions under which plasticity arises due to migration are not clear, particularly if plasticity is costly. Here, we propose that plasticity will arise in populations across a broad range of conditions, including costs of plasticity, if migration rates are sufficiently large. We support our argument with analytic derivation and extensive simulations assuming a two-locus two-habitat population genetics model. We examine the properties of the evolution of the plasticity modifier allele across adapted Wright-Fisher populations assuming a range of migration rates, plasticity effects with and without the cost of plasticity, recombination rates, mutation rates, and genetic architectures. We find the evolution of plasticity due to migration is robust to the genetic architecture of the plasticity effect and recombination rates, and occurs across a wide range of migration scenarios.

634 - The genomic footprint of social structure

Barbara Parreira

Barbara Parreira, Ida Moltke, Shyam Gopalakrishnan
Center for Evolutionary Hologenomics, Globe Institute, University of Copenhagen, Copenhagen, Denmark., Department of Biology, University of Copenhagen, Copenhagen, Denmark, Center for Evolutionary Hologenomics, Globe Institute, University of Copenhagen, Copenhagen, Denmark

A classical population genetics approach adopts the hierarchical view of nature that populations are networks of discrete random-mating entities (demes) placed in space. But real populations are rarely structured in clearly identifiable demes. In a large number of taxa such as insects, birds and mammals, species are subdivided beyond space. Individuals are organized in units where they live and reproduce according to more or less complex strategies. Genetically, fine-scale population structure creates profound violations to the standard coalescence imposing, particularly affecting coalescent events in the recent past. Therefore, genetic summaries providing information on processes occurring at short evolutionary time-scales, such as identical-by-descent tracts, are potentially very powerful to characterize fine-scale population structure. In this study, we are investigating the properties of gene trees in socially structured populations. We study gene genealogies in highly structured models underlying species with long life-spans and high reproductive skew. Our aim is to characterize the genomic footprint of fine-scale population structure. We show that the subdivision of species in social groups affects commonly used statistics, such as the site frequency spectrum, which can form a basis for inference of fine-scale structure from genomic data.

683 - Repeatability of Introgression in a Grasshopper Hybrid Zone: Insights into the Maintenance of Species Boundaries by Haldane’s Rule

Ricardo J. Pereira

Ricardo J. Pereira, Linda Hagberg, DÃ¶rte Neumeister, Edgardo Ortiz-Valencia, David Buckley, Jose L Bella
State Museum of Natural History in Stuttgart, LMU-Munich, LMU-Munich, TUM, UAM, UAM

The evolution of hybrid male sterility, known as Haldaneâ€™s rule, is a widely observed phenomenon in species formation across animals and plants. However, the persistence of hybrid male sterility in the face of long-term hybridization and its implications for introgression patterns remain unclear. Here, we investigate these questions using target capture data from two replicated transects across the hybrid zone of the meadow grasshopper, where hybrid male sterility has repeatedly evolved within a species. Ancestry analyses reveal

consistent patterns of introgression across replicated transects, indicative of hundreds of generations of gene flow since secondary contact. Notably, we observe higher rates of gene flow at lower altitudes. Demographic modeling of pure populations indicates that the hybrid zone originated from the expansion of a northern taxon over a stable taxon that remained in the southern slope of the Pyrenees. Geographic clines demonstrate asymmetric neutral gene flow due to demographic imbalances, while genes associated with barriers to gene flow exhibit cline centers in two distinct geographic clusters, coinciding with phenotypic clines for hybrid male sterility. This genomic analysis of replicated transects reveals repeatable patterns in locus movement and selection, suggesting that Haldaneâ€™s rule can be resolved in old hybrid zones, leading to predictable patterns of introgression.

704 - Geographic distance prevails: predicting genetic diversity distribution of highly mobile marine species

Jéssica Fernanda Ramos Coelho

JÃ©ssica FR Coelho, Sergio MQ Lima, Julia T Verba, Ricardo J Pereira

Universidade Federal do Rio Grande do Norte, Universidade Federal do Rio Grande do Norte, Ludwig Maximilians UniversitÃ¤t MÃ¼nchen, Stuttgart State Museum of Natural History

Habitat suitability is fundamental for a species survival in a particular area. Regions characterized by environmental stability over time have been recognized as biodiversity hotspots, harboring elevated genetic diversity, thus holding considerable conservation value. Uncovering the connection between climatic oscillations and genetic changes in marine ecosystems may offer insights into whether abiotic dynamics can act as predictors for the genetic diversity of populations. Using genomic and ecological data of two populations (coastal and island) of tropical scaled-sardines Harengula sp. in the Atlantic Southwest, our study investigates the correlation between environmental suitability stability between current and past climates (Last Glacial Maximum) versus patterns of genetic diversity and demography. In addition, we tested the role of environmental resistance on the distribution of genetic diversity. We used nucleotide diversity (Ï€) and the neutrality test Tajimaâ€™s D as proxies of genetic diversity and demography, respectively. Environmental suitability differences between current and past climates were estimated from ecological niche models. Our analysis revealed that the genetic patterns of coastal and island populations did not reflect the extent of environmental suitability stability of the Atlantic Southwest (Spearman correlation coefficient rs=-0.12, p>0.05). A Maximum Likelihood Population Analysis showed that genomic data is better explained by geographic distance rather than environmental resistance (AIC= -355.3, weight=0.724), which reinforces the isolation by distance previously detected in the coastal population. Together, our results indicate that geographic distance, rather than environmental suitability stability or resistance, is a good predictor of genetic diversity distribution in highly mobile species, like sardines.

861 - Application of coalescence-based inference of gene flow rates

Elizabeth T. Chevy

Elizabeth T. Chevy, Peter L. Ralph, Sohini Ramachandran

Brown University, University of Oregon, Brown University

Sexual organisms exist in geographic space, a fact that restricts the available sexual partners with whom an individual can reproduce. Over generations, genes travel over the landscape, creating meaningful patterns of genetic diversity across space. Inferring rates of this spatial gene flow has applications across many fields; for example, gene flow rates can explain how ecological features contribute to genetic differentiation, or reveal cryptic subpopulation structure. Here we implement and extend Lundgren and Ralph's 2019 gene flow rate inference method, and demonstrate its use in an empirical scenario. The method uses a flexible discrete-deme model of subpopulations, and uses a coalescence-time framework to allow for anisotropic gene flow surfaces. We implement Bayesian inference

of the gene flow rates in the open-source computing platform Stan, which is well-suited to high- dimensional inference tasks, and allows for greater reproducibility of our analyses. We apply our method to Anopheline data from the Ag1000G project.

901 - Exploring the behavior of two-locus statistics in continuous space

Lloyd Kirk

Lloyd Kirk, Simon Gravel, Peter L Ralph, Jerome Kelleher, Aaron P Ragsdale
Department of Integrative Biology, University of Wisconsin-Madison, Department of Human Genetics, McGill University, Institute of Ecology & Evolution, University of Oregon, Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Department of Integrative Biology, University of Wisconsin-Madison

Succinct tree sequences have transformed the way we store and analyze genomic variation in both simulated and empirical datasets. Tskit is the widely-used reference implementation of tree sequences and the leading toolset for analyzing tree sequence data. We have developed a generalized framework for computing two-locus statistics in tskit, implementing a broad range of two-locus statistics. This work enables the flexible computation of LD matrices (as well as LD scores and LD decay) from observed mutations and expectations for those statistics from the underlying genealogical structure. Our framework extends to any single or multi-population statistic between sites with arbitrary numbers of alleles.
We use the robust ecosystem of simulation tools that integrate with tskit to study how two-locus statistics behave under realistic demographic scenarios. Most population genetic models of geographic dispersal are based on assumptions of discrete, homogeneous populations. However, individuals in fact live across continuous geography, and recent advances enable simulation of population genetic processes in continuous space. We use these new methods, along with our framework for computing two-locus statistics to investigate the impacts of neighborhood size on local LD structure. We aim to uncover the effects of mate choice and dispersal that occur at fine scales and how they impact the dynamics of two-locus statistics in the presence of continuous geographic structure. In addition, we investigate the effect of spatial sampling of individuals on the estimation of different measures of LD and estimate the impact of sampling strategy on downstream inferences.

902 - Evolutionary patters of Rhinella horribilis in response to anthropic habitat modification in two landscapes in Oaxaca, Mexico

Gerardo Jesús Soria Ortiz

Gerardo J. Soria Ortiz, Ella G. VÃ¡zquez DomÃ-nguez, Leticia M. Ochoa Ochoa, Juan P. Jaramillo Correa
Instituto de EcologÃ-a, UNAM, Instituto de EcologÃ-a, UNAM, Facultad de Ciencias, UNAM, Instituto de EcologÃ-a, UNAM

Anthropized environments increase species vulnerability. Describing evolutionary patterns in species that inhabit such habitats is of great importance to understand population dynamics and ensure their long-term persistence. Rhinella horribilis has generalist habits and resilient toad capable of inhabiting areas with human-induced environmental modifications. Based on genomic data (SNPs by dd-RadSeq), we analyzed the genetic structure and functional connectivity of R. horribilis in two landscapes with different levels of anthropic modification in Oaxaca, Mexico. We also identified candidate loci associated with environmental, climatic and waterbody quality variables. We found that R. horribilis maintains functional connectivity among populations mainly due to the temporary water bodies that form during the rainy season. Vegetation variables and climatic factors affect their connectivity differentially, in close association with the degree of anthropic modification. Furthermore, we found shared genetic variation associated with specific environmental variables in each landscape, encompassing genes and metabolic pathways related to embryonic development and metamorphosis, and immune responses. These evolutionary patterns are key to understand how species that inhabit human-modified

sites manage to cope with selection pressures exerted by rapid environmental changes. Our study is pioneer among the few identifying genetic variation associated with environmental stressors promoted by anthropogenic modification in amphibians.

942 - Determining concordance and drivers of spatial population structure through genogeographic clustering

Shane Lavery

Vanessa Arranz, Rachel Fewster, Shane Lavery

University of Barcelona, University of Auckland, University of Auckland, University of Auckland

In recent years there have been great strides made in the sophistication and detail of both the collection and analysis of spatial population genetic data within single species. However, the methods of comparing patterns of population structure across species have been much slower to develop. This is problematic, as it is becoming clear that an understanding of the patterns of concordance in population structure across multiple species is necessary to better comprehend community patterns, and thus provide insights into the common drivers of those patterns. Here, a novel approach, â€œgenogeographic clusteringâ€•, was used to determine concordance of spatial population structure in a community of 21 species of coastal marine organism in New Zealand waters. The approach involved: 1) use of genetic divergences from previously published studies to quantitatively describe patterns of population structure within each species as a fitted spline curve, 2) quantitatively clustering species by their similarity in geographic pattern using a dendrogram of curve similarities, and 3) then testing whether known life-history and ecological traits are associated with the species sharing similar genetic patterns, using distance-based regression. This approach shows promise in helping us determine both the concordance in spatial patterns among species in a community, and the potential processes driving those patterns.

1015 - Negative frequency-dependence is not a mechanism of coexistence in spatially fragmented populations

Anush Devadhasan

Anush Devadhasan, Oana Carja

Carnegie Mellon University, Carnegie Mellon University

Across scales of organization, understanding the mechanisms that shape and maintainÂ diversity in a population remains a fundamental challenge and there exists a large literature in evolutionary biology describing the processes that can account for persistent coexistence within populations. At their core, the theoretical frameworks are grounded in balancing selection, including negative frequency-dependent (NFD) selection, spatial or temporal habitat heterogeneity, and heterozygote advantage. In particular, NFD selection, by favoring rare types, is often suggested as the main selective force for biodiversity and polymorphism in natural populations. In this talk I will describe an evolutionary model in which, unintuitively, the opposite occurs: NFD selection decreases coexistence relative to neutrality, for certain spatial population structures. We call this the spatial speedup effect and, using a combination of mathematical analysis and simulation, we find that the condition for this effect depends critically on the spatial arrangement of the population. We show that spatial speedup is specifically shaped by the algebraic connectivity of the population: a measure of â€œspatial fragmentationâ€• that derives from spectral graph theory. I will discuss how spatial speedup in fragmented populations can significantly impact biodiversity over evolutionary and ecological timescales by analyzing fixation times and species richness as the respective proxies for coexistence.Â

1021 - Reconstructing the locations of genetic ancestors for a recombining sequence

James Kitchens

James Kitchens, Puneeth Deraje, Matthew Osmond, Graham Coop
University of California, Davis, University of Toronto, University of Toronto, University of California, Davis

Present-day spatial patterns of genetic relatedness among organisms reflect the past movements of their ancestors. The ability to untangle these histories has improved dramatically thanks to advances in ancestral recombination graph (ARG) methods. By extending spatial methods previously applied to trees, we generalize a model of Brownian dispersal to work on ARGs, thereby accounting for correlations along a chromosome when computing the likelihood-based estimates of dispersal rate and locations of genetic ancestors. Further, we develop an efficient algorithm that allows for the application of the method to complex ARGs, scalable to thousands of samples. We evaluate our method's ability to reconstruct spatial histories using simulations. Surprisingly, despite using the fullest information available in the data, we find that our dispersal estimates are biased, uncovering friction between the histories of recombinant lineages and Brownian dispersal models. We identify potential resolutions to this problem based on relaxing the constraints that ARGs place on the movement of lineages and show that ARG-based spatial inference can be used to effectively track the geographic history of admixed individuals. Approaches like this will be key to understanding the interplay of migration, recombination, drift, and adaptation in geographically spread populations.

201 - Elucidating the origin and molecular evolution of the cholesterol biosynthesis pathway.

Janeth Alejandra García

Janeth A. GarcÃ-a, Antonio Lazcano, Ricardo HernÃ¡ndez-Morales
Departamento de BiologÃ-a Evolutiva, Facultad de Ciencias, Universidad Nacional AutÃ³noma de MÃ©xico., Instituto de Investigaciones BiomÃ©dicas, Universidad Nacional AutÃ³noma de MÃ©xico., Departamento de BiologÃ-a Evolutiva, Facultad de Ciencias, Universidad Nacional AutÃ³noma de MÃ©xico., Departamento de BiologÃ-a Evolutiva, Facultad de Ciencias, Universidad Nacional AutÃ³noma de MÃ©xico.

Cholesterol is an essential component of animal cell membranes. The biosynthesis of this sterol has been reported in other eukaryotic organisms, such as fungi and plants. It is believed that the appearance of cholesterol in cell membranes is correlated with the evolution of eukaryotic organisms and the increase in atmospheric oxygen that occurred 2.32 billion years ago, molecular oxygen is already required for the completion of cholesterol biosynthesis, but recent research has found that the halophilic bacterium Enhygromyxa salina may be able to synthesize cholesterol de novo, raising questions about the antiquity of this pathway. Therefore, this project will attempt to address the origin and molecular evolution of the cholesterol biosynthetic pathway by comparing the primary and tertiary structure of the enzymes involved in the biosynthesis. Preliminary results of our study suggest that the ability to synthesize sterols may be more widespread than previously thought, opening new perspectives on the evolution and distribution of this pathway across domains of life.

233 - Evolutionary dynamics of genes involved in pollination syndromes in monocotyledons

Deivid Almeida de Jesus

Deivid Almeida de Jesus, Thiago AndrÃ©, Carlos Guerra Schrago
Universidade Federal do Rio de Janeiro, Universidade de BrasÃ-lia, Universidade Federal do Rio de Janeiro

Flowering evolution boosted plant diversity through the rise of the angiosperm flower, a complex strobilus with varied shapes and colors responding to a rich set of selective agents. Understanding this evolutionary process is vital to address historical processes that drove diversification as well as its future under climate change. To investigate main evolutionary processes that were mediating convergence evolution into pollination syndromes in the monocotyledon lineage, from an exhaustive literature search, 12 genes were listed and tested

for selection in 161 taxons specimen samples deposited in the GenBank. Sequences alignment was performed in Clustal as implemented in MAFT v7.5. Motif analysis was performed on the web server MEME suite 5.5.5 with MEME tools using default parameters. We performed Maximum Likelihood phylogenetic inference in IQTree v1.6.12, using the evolutionary model TVMe+G4, and 1000 replicates of alrt. Selection inferences were obtained using the HyPhy v2.5.7 software with the aBSREL tool. The Swiss model website was used for protein modeling. All genes shared at least two conserved motifs. Phylogenetic inference revealed that almost all genes form a monophyletic group, except the BIG BROTHER gene. Selection analysis showed that five different genes in three species evolved under positive selection: PISTILLATA in Triticum aestivum (Poales); RADIALIS, FLS, and DFR in Musa acuminata (Zingiberales), and a CHS-like gene in Arisaema heterophyllum (Alistamales). Protein structural analysis of these species revealed that the three-dimensional structures are conserved.

241 - Corrin biosynthesis in the last universal common ancestor

Luca David Modjewski

Luca D. Modjewski, Val Karavaeva, Natalia Mrnjavac, Filipa L. Sousa, William F. Martin
Institute of Molecular Evolution, Heinrich Heine University DÃ¼sseldorf, Department of Functional and Evolutionary Ecology, University of Vienna, Vienna Doctoral School of Ecology and Evolution, University of Vienna, Institute of Molecular Evolution, Heinrich Heine University DÃ¼sseldorf, Department of Functional and Evolutionary Ecology, University of Vienna, Institute of Molecular Evolution, Heinrich Heine University DÃ¼sseldorf

Corrins are cobalt-containing tetrapyrroles that serve as cofactors for numerous methyl transfer and radical-dependent reactions. They are essential in the most ancient CO2-fixing biochemical route as cofactors for the corrinoid iron-sulfur protein, CoFeS, which catalyzes a cobalt-to-nickel methyl transfer in the acetyl-CoA pathway. CoFeS occurs in H2-dependent archaeal methanogens, perhaps the most ancient microbial lineage known, dating to 3.5-3.95 billion years ago, based on carbon isotope data, and in the acetyl-CoA pathway of H2-dependent bacterial acetogens. Was the corrin synthesis pathway transferred from the archaeal to the bacterial lineage during evolution, or was it present in the last universal common ancestor, and in either case, how did the primordial acetyl-CoA pathway function prior to the enzymatic synthesis of natureâ€™s most complex cofactor? Here we analyze 26 enzymes for corrin biosynthesis in acetogens and methanogens, showing that phylogenies trace the pathway to their common ancestor. Prebiotically synthesized corrins are unlikely functional forerunners of CoFeS function, yet in laboratory simulations of serpentinizing hydrothermal systems, Fe-, Ni- and Co-containing solid state surfaces catalyze methyl transfer reactions of the acetyl-CoA pathway, presenting natural analogue of corrin function. Corrin biosynthesis was a massive evolutionary innovation, demanding the existence of a steep selective barrier. The data suggest that insolubility of its solid-state predecessor catalyzing essential metal-to-metal methyl transfers in a primordial hybrid acetyl-CoA pathway (part abiotic, part enzymatic) was that hurdle in early biochemical evolution.

250 - Physicochemical Evaluation of Remote Homology in the Twilight Zone

Jamie D Dixson

Jamie D Dixson, Rajeev K Azad
University of North Texas, University of North Texas

A fundamental problem in bioinformatics is determining the degree of evolutionary relatedness among Twilight Zone homologs that by definition have diverged to a point where they share less than 30% identity and yet retain similar structures and/or functions. Many researchers have leveraged experimentally determined structures in the quest to classify Twilight Zone proteins. Such endeavors are difficult and require great investments of time and money. Using both simulated and real-world data we characterize molecular

weight-hydrophobicity physicochemical dynamic time warping (MWHP PCDTW or MWHP DTW for short) as a method requiring only the amino acid sequence to quantify similarity of related proteins in the Twilight Zone. This is a step forward in determination of the degree of relatedness among Twilight Zone proteins and most notably allows for the discrimination of random similarity and true homology in the 0-~20% identity range. This method was previously presented expeditiously as a result of the Covid-19 pandemic because it was able to functionally cluster ACE2-binding betacoronavirus receptor binding domainâ€™s (RBD) a task that has been allusive using standard techniques. Here we show that one reason that MWHP DTW is superior to standard alignment techniques with regard to comparisons within the Twilight Zone is because standard alignment produces incorrect alignment when comparisons are made within the Twilight Zone. Further, we present an extended definition of the Twilight Zone that incorporates the dynamic relationship between structural, physicochemical and sequence-based metrics.

305 - Molecular evolution of the Spike protein of SARS-CoV-2: evidence of natural selection

Georgina I López Cortés

Georgina I LÃ³pez CortÃ©s, Miryam Palacios PÃ©rez, Hannya Velediaz, Margarita Hernandez Aguilar, Marco JosÃ©
Instituto de Investigaciones BiomÃ©dicas, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de Investigaciones BiomÃ©dicas, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de Investigaciones BiomÃ©dicas, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de Investigaciones BiomÃ©dicas, Universidad Nacional AutÃ³noma de MÃ©xico, Instituto de Investigaciones BiomÃ©dicas, Universidad Nacional AutÃ³noma de MÃ©xico

The global scale COVID-19 pandemic made possible to study the evolution of SARS-CoV-2. Here, we describe the molecular evolution of the spike (S) protein of SARS-CoV-2 by applying a neutral substitution test to reported sequences, and relate the probability of occurrence of the amino acids with the real functionality of the Receptor Binding Domain (RBD). The neutrality of mutation of the S protein of Betacoronavirus was computed at 2 different time points. Specifically, Tyr and Asn exhibited higher occurrence rates, whereas His and Arg displayed lower occurrence rates. Moreover, the neutrality of the first variants were more deviated from expected the neutral control than the most recent ones. Some amino acids involved in the receptor binding are under selective pressure, causing them to remain fixed. We identified the amino acids involved in the proteinâ€”protein interaction, being the RBD of Omicron which established the highest contacts. Hence, the mean number of contacts per residue is higher, making the contact thermodynamically more stable.
Our results show that most unique mutations are under selective pressures, which could be related either to the evasion of the immune system, increasing the virusâ€™ fitness or altering protein â€“ protein interactions with host proteins. We explored the consequences of those selected mutations in the structure and protein â€“ protein interaction with the receptor. Altogether all these forces have shaped the spike protein and the continually evolving variants.

530 - Characterizing the origin, evolution, and function of the synaptic protein PSD95 (DLG-4 gene) in single-celled relatives

Riya Nilkant

Riya Nilkant, Lisa Y Mesrop , Samuel Lobo, Anissa Morrison , Onur Sakarya , Soojin V Yi, Kenneth S Kosik
University of California Santa Barbara , University of California Santa Barbara, University of California Santa Barbara, University of California Santa Barbara, University of California Santa Barbara, University of California Santa Barbara, University of California Santa Barbara

The origin and evolution of complex synaptic machinery are facilitated by new protein-protein binding partnerships. Evolutionary changes in protein sequence and binding properties of PDZ domains can evolve new interactions that serve as primers for assembling complex

protein modules with novel synaptic functions. Here we focus on the evolutionary origin and functional divergence of the tri-PDZ domain architecture of PSD-95, a key scaffolding protein in post-synaptic cells. We use phylogenetics, ancestral reconstruction, and protein structure predictions to determine the ancestral PDZ domain and the evolutionary changes that occurred in the binding interactions resulting in new binding partnerships. By comparing PDZ domains in unicellular relatives and Metazoans, we found the n-terminally positioned PDZ is ancestral to the first two PDZ domains and originated along the ancestral lineage leading to the Holozoan clade. Although the interaction of conserved GLGF residues determines the binding specificity of PDZ domains, we found that unicellular PDZ domains still bind in silico to endogenous synaptic-like ligands even in their absence. We observed the hydrophobicity through a molecular dynamics simulation to be similar across these unicellular PDZ domains and Metazoans despite the lack of residue conservation. We further found conservation of hydrogen bond interactions at positions 0 and -2 of PSD-95 ligands across humans, the unicellular relative C.owczarzaki, and certain predicted ancestors. We posit that the hydrogen bond interactions at specific ligand positions, and not conserved residue identities, are likely evolutionary drivers of new ligand partners for PSD-95.

655 - Unicore Enables Ultra-fast and Accurate Phylogenetic Reconstruction with Structural Core Genes

Dongwook Kim

Dongwook Kim, Sukhwan Park, Martin Steinegger
Seoul National University, Seoul National University, Seoul National University

Breakthroughs in AI-driven protein structure prediction are transforming biology. Efforts [1,2] to integrate predicted protein structures into phylogenetic analyses have shown promise, however, have been constrained by the availability of accurate models or computational demands. The ProstT5 [3] protein language model and Foldseekâ€™s 3Di structural alphabet [4] are circumventing the need for slow structure prediction with over three orders of magnitude faster prediction of structural sequences from amino acid input, and are posed to revolutionize phylogenetic reconstruction.

Here, we present Unicore, an ultra-fast and accurate method for structure-based phylogenetic inference that leverages predicted 3Di sequences, focusing on the identification of highly conserved structural core genes. We demonstrate this approach to reliably reconstruct phylogenetic relationships with linear run-time scaling over the size of the target proteomes, while being congruent with conventional approaches using orthologs or core genes. Unicore is universally applicable to any given set of taxa, even spanning superkingdoms and overcoming limitations of previous methods requiring orthologs of fixed taxonomic scope. As an ultra-fast solution for structural phylogenetic reconstruction without taxonomic limitations, we expect Unicore to pave the way to a new era of structural phylogeny.

Unicore is free and open source software and will be available at https://github.com/steineggerlab/unicore.

[1] Moi et al., bioRxiv, 10.1101/2023.12.12.571181, 2023.
[2] Puente-Lelievre et al., bioRxiv, 10.1101/2023.09.19.558401, 2023.
[3] Heinzinger et al., bioRxiv, 10.1101/2023.07.23.550085, 2023.
[4] van Kempen et al., Nature Biotechnology, 2023.

724 - In silico prediction of eukaryotic complexomes with protist genomes

Ore Francis

Ore Francis, Henning Hermjakob, Sandra Orchard, Mark Blaxter, Ore Francis
Wellcome Sanger Institute, European Bioinformatics Institute (EMBL-EBI), European Bioinformatics Institute (EMBL-EBI), Wellcome Sanger Institute, Wellcome Sanger Institute

Protein complexes are responsible for essential biological functions and thus conserved throughout all domains of life. Our understanding of protein complexes is biased towards model species and genomically well-sampled clades. Although initiatives to diversify the taxonomic spread of available genomes are ongoing, discovery, analysis and prediction of protein complexes throughout eukaryotes is hindered by bias in selection of species for sequencing, which are typically of animal, fungi or plant origin. This bias critically misses the huge phylogenetic diversity present in the â€œprotistsâ€•. To address the underrepresentation of protists and facilitate prediction of protein complexes across all eukaryotic phyla, we have predicted protein complex profiles (complexomes) for eukaryotes from genomic and/or predicted proteome records, using the EBI Protein Complex Portal database (https://www.ebi.ac.uk/complexportal/home) as a query dataset for homology searches. Homologs to complex forming extracellular matrix proteins, both metazoan tissue-specific proteases and regulatory proteins were identified sparingly in protist lineages, surprisingly, within the SAR clade. Many of these putative protist protein complexes are referenced, but not confirmed in scientific literature. This long-term project will be key to deciphering the protein complexome of the last eukaryotic common ancestor. As more genomic data becomes available, the predictive capacity of our software will improve. Our database and associated tools will be open and key resources for predictions of protein function in eukaryotic genomes.

731 - The genetic basis for changes in viral receptor-binding specificities

Caroline Craig

Tyler Starr, Ashley Taylor
University of Utah, University of Utah

Animal viruses pose a threat when they evolve traits that enable spillover into humans. The first step necessary for many viruses to cross species boundaries is to bind with high affinity to novel host receptors, allowing entry to initiate infection. However, the molecular mechanisms by which viruses evolve to recognize novel host receptors are not well-characterized. An extreme example of variation in viral receptor usage was recently described in a group of coronaviruses called merbecovirusesâ€”a large subgenus that includes the epidemic human Middle East respiratory syndrome coronavirus (MERS-CoV) and its diverse relatives from mammals. Though MERS-CoV and closest relatives from bats and camels utilize the DPP4 protein as their entry receptor, more distant bat merbecoviruses were recently shown to utilize ACE2 protein receptor; the genetic mechanisms by which these differences in receptor usage evolved have not been characterized. To isolate the evolutionary interval when merbecovirus receptor usage changed, we perform phylogenetic reconstructions of ancestral merbecovirus receptor-binding domains and characterize receptor-binding specificities via a yeast-display-based deep mutational scanning platform. From these data, we identify the evolutionary coincidental sequence substitutions that enable this change in receptor usage. We describe a combinatorial mutational scanning experiment for subsequent genetic dissection of the causal historical substitutions, including the role of epistasis or promiscuous intermediates in this historical evolutionary transition. This work identifies the mechanisms of a dramatic change in protein function while informing efforts to predict or prevent future spillover from this viral family of future pandemic potential.

857 - Secondary structure conservation sheds light on the forces that drive the evolution of condensate proteins

Lisa E Kursel

Lisa E Kursel, Kaan Goktepe, Ofer Rog
University of Utah, University of Utah, University of Utah

Condensates are a newly-appreciated way to organize the cell. Like membrane-bound organelles, condensates enable spatial and temporal regulation. However, condensates are regulated by distinctive molecular interactions. Rather than stable, tight interfaces, condensates assemble through weaker, labile interactions. This mode of assembly could allow for sequence variability, limiting the applicability of current knowledge on protein evolution. Even when carrying out essential functions, condensate members tend to lack conserved amino acid sequence and identification of orthologs relying on per-site identity has been challenging. We addressed this paradox by analyzing the evolution of proteins from two condensates, centrosomes (essential for cell division) and the synaptonemal complex (SC, essential for reproduction), in Caenorhabditis. We found that centrosome and SC proteins have an unusual evolutionary signature; they are significantly more diverged than other proteins, but retain conserved secondary structures including coiled-coil domains. We found similar evolutionary signatures in Drosophila and Eutherian mammals despite no homology across clades. This suggested we could identify condensate proteins by leveraging their diverged primary sequence but conserved secondary structure. We demonstrated this by cloning a novel SC protein, PpaSYP1, in a distantly related nematode, Pristionchus pacificus. PpaSYP1 has no homology to any known SC protein. We made a knockout and tagged version of PpaSYP1 and confirmed that it localizes to the SC and is required for reproduction. Overall, we observe patterns of functional and structural conservation that do not depend on primary sequence conservation, shedding light on the forces that drive the evolution of condensate members.

879 - Truncation as a mechanism of innovation in the evolution of innate immune protein CD14

José A Sánchez-Borbón

JosÃ© A SÃ¡nchez-BorbÃ³n, Kona N Orlandi, Michael J Harms
Institute of Molecular Biology, University of Oregon, Department of Biology, University of Oregon, Institute of Molecular Biology, University of Oregon, Department of Biology, University of Oregon, Institute of Molecular Biology, University of Oregon, Department of Chemistry and Biochemistry, University of Oregon

How does evolution assemble protein complexes made of interdependent parts? We use the Toll-like receptor 4 (TLR4) complex as a model for this question. TLR4 recognizes bacterial lipopolysaccharides (LPS) and activates inflammation. Here, we explored the evolution of the CD14, an important support protein that transports LPS to the TLR4 complex. Prior to our work, it was thought that CD14 evolved in the ancestor of tetrapods, ~70 million years after the TLR4 complex formed. Using phylogenetic and syntenic analyses, we found that CD14 is present in ray-finned fishes. Using cell-based assays, we validated that CD14 from two extant fishes were indeed functional, suggesting CD14 evolved concurrently with the rest of the TLR4 complex. We next sought to understand the evolution of CD14 function. CD14 is a paralog of TLR2 that evolved by serial truncation after duplication. Using ancestral sequence reconstruction and structural modeling, we found that CD14 evolved its LPS delivery function when a truncation exposed a hydrophobic surface, creating a functional site de novo. Ancestral TLR4 and CD14 complexes induce inflammation in experimental assays, demonstrating that the complex was active in the bony vertebrate ancestor. The TLR4 complex was then modified along descendent lineages. In particular, the model vertebrate Danio rerio lost CD14 but gained CD14-independent TLR4 activity. This work improves our understanding of the evolution of the TLR4 complex, reveals that functional sites can evolve by exposure of a hydrophobic surface, and helps map the results of studies in Danio rerio to human biology.

890 - Modeling archaic hominin protein structures reveals evidence of reduced efficacy of selection

Ava Xu

Ava Xu, David Rinker, Bian Li, Christopher Moth, Gregory Sliwoski, Jens Meiler, John Capra
University of California, San Francisco, Vanderbilt University, Vanderbilt University, Vanderbilt University, AB Hancock Jr. Memorial Laboratory for Cancer Research, Vanderbilt University, Vanderbilt University, University of California, San Francisco

Protein-coding genetic variants distinguishing anatomically modern humans and their archaic hominin relatives, the Neanderthals and Denisovans, are of great interest due to their potential contribution to lineage-specific traits. However, the functional effects of most lineage-specific variants are unknown.

We combine high-coverage archaic genome sequences and powerful machine learning methods to evaluate the structural and functional effects of lineage-specific variants. We focus on 1050 fixed or very high-frequency lineage-specific derived missense variants that are likely strongly enriched for positive selection. To account for potential human-specific biases, we consider a diverse range of functional metrics including biophysical (ddG), evolutionary (within and between species sequence constraint), and deep learning (AlphaMissense, ESM1b) models.

Variants derived in the archaic lineage are significantly more disruptive to protein structure and function compared to human-derived variants. This result holds for every functional metric used and is robust to the thresholds used to define lineage-specificity. Moreover, it is robust to filtering sites of potential ancient DNA damage. We also compare the effects caused by lineage-specific variants to a matched set of 409 annotated pathogenic missense variants from the ClinVar database. As expected, both human- and archaic-derived missense variants have smaller effects on proteins than human disease-associated missense variants.

Our results find a substantial role for genetic drift within archaic populations and support previous studies concluding that the effective population size and efficacy of selection were lower in archaic hominins than modern humans.

926 - Quasi-Ordered Recognition Dynamics in Sperm-Egg Interactions in Abalone

William T Higgins

William T Higgins, Damien B Wilburn
The Ohio State University, The Ohio State University

Protein evolution does not occur at a fixed rate over time, but rather as a dynamic process influenced by physiological and genetic factors. Some protein sites are subject to higher degrees of evolutionary conservation than others, such as catalytic residues or residues involved in forming protein-protein interactions. In the marine mollusk abalone, sperm lysin and egg VERL are rapidly coevolving proteins that mediate species-specific interactions during fertilization. The VERL binding interface on lysin consists of many rapidly evolving residues that are loosely packed in a â€œquasi-orderedâ€• state that is primarily stabilized by the coordination an aromatic triad: W3, W62, Y133. This hydrophobic packing is observed in red abalone and closely related taxa, but Y133 is a derived characteristic compared to ancestral N133. This evolutionary transition coincides with a Q63A mutation adjacent to W62. The precise interactions and dynamics between these residues are not well understood. We employed molecular dynamics simulations using hydrogen mass repartitioning to evaluate the effects of A63Q and Y133N mutations in red abalone lysin to gain further insights about the protein dynamics on longer timescales. Preliminary analyses support that epistatic constraints have preserved this spatial coordination via these residues through alternative molecular interactions. Complementary solution NMR experiments to evaluate these trends are ongoing. Few quasi-ordered proteins have been biophysically characterized, and patterns inferred from lysin structural evolution to maintain species-specific fertilization may provide useful insights into other protein systems with complex molecular dynamics.

981 - Evolution and functional role prediction of the CYP6DE and CYP6DJ subfamilies in Dendroctonus (Curculionidae: Scolytinae) bark beetles

Juan Manuel Quijano Barraza

Juan Manuel Quijano Barraza, Gerardo ZÃºÃ±iga, Claudia Cano RamÃ-rez, MarÃ-a Fernanda LÃ³pez, Gema L RamÃ-rez Salinas, Moises Becerril

Instituto PolitÃ©cnico Nacional, Instituto PolitÃ©cnico Nacional, Instituto PolitÃ©cnico Nacional, Instituto PolitÃ©cnico Nacional, Instituto PolitÃ©cnico Nacional, Instituto PolitÃ©cnico Nacional

Dendroctonus beetles are natural components of coniferous forests that spend most of their lives under the bark, where they are exposed to toxic terpenes present in the oleoresin. Cytochrome P450 (CYP) enzymes are involved in the detoxification of these compounds. It has been demonstrated that CYP6DE and CYP6DJ subfamilies hydroxylate monoterpenes, whose derivatives can act as pheromone synergist be pheromones themselves. Given the functional diversity of CYPs, we investigated whether these cytochromes have retained their function throughout the evolution of these insects. We performed a Bayesian phylogenetic analysis to determine subgroups of cytochromes in these subfamilies. Subgroups were mapped and reconciled with the Dendroctonus phylogeny. Molecular docking analyses were performed with the cytochromes of each subgroup and enantiomers of Î±-pinene and Î²-pinene, (+)-3-carene, Î²-myrcene and R-(+)-limonene. In addition, functional divergence analysis was performed to identify changes in catalytic site and/or protein folding. The mapping and reconciliation analysis showed different patterns for each subgroup. Functional predictions indicated that the cytochromes analyzed can hydroxylate all monoterpenes with preferential affinities to different monoterpenes. The CYP6DE subfamily has experimented type I and II divergence, whereas the CYP6DJ subfamily has evolved under strong functional constraints. Results suggest cytochromes of the CYP6DE subfamily evolve to reinforce their detoxifying capacity hydroxylating mainly Î±- and Î²-pinene to (+) and (âˆ')-trans-verbenol, compound used as a pheromone by several Dendroctonus species; whereas the CYP6DJ subfamily appear to retain its original function related to the detoxification of these compounds.

1055 - Biosurveillance of zoonotic coronaviruses through an alignment-free physicochemical clustering approach

Rajeev K Azad

Jamie D Dixson, Rajeev K Azad

University of North Texas, University of North Texas

To infer functional relatedness among betacoronaviruses, we developed an alignment-free remote homology detection technique, which was leveraged to gain new insights into coronavirus evolution, specifically in the context of the COVID-19 pandemic caused by novel coronavirus SARS-CoV-2. Novel methods for predicting the capacity for coronaviruses in general to infect human cells are needed. Our proposed method utilizes physicochemical properties of amino acids to develop a fully dynamic waveform representation of proteins that encodes both the amino acid content and the context of amino acids. These waveforms are then subjected to dynamic time warping and distance evaluation to develop a distance metric that is relatively less sensitive to variation in sequence length or primary amino acid composition. Using this method, we show that in contrast to alignment-based maximum likelihood and neighbor-joining phylogenetic analyses, all bat betacoronavirus spike protein receptor binding domains (RBDs) known to bind to the ACE2 receptor are found within a single physicochemical cluster. Further, other RBDs within that cluster are from pangolin coronaviruses, two of which have already been shown to bind to ACE2 while the others are suspected, yet unverified ACE2 binding domains. This finding is important because both SARS-CoV and SARS-CoV-2 use the host ACE2 receptor for cell entry. Surveillance for coronaviruses belonging to this cluster could potentially guide efforts to stifle or curtail potential and/or early zoonotic outbreaks with their associated deaths and financial

devastation.

89 - Can repetitive elements lead to speciation through chromosomal rearrangements?

Camille Cornet

Camille Cornet, Kay Lucek
University of NeuchÃ¢tel, University of NeuchÃ¢tel

Large-scale chromosomal rearrangements may act as reproductive barriers, contributing to speciation. It has been hypothesised that chromosomal rearrangements are caused by repetitive elements such as transposons. Lepidoptera, i.e. butterflies and moths, which are among the most diverse groups across the Tree of Life, have holocentric chromosomes without defined centromeres. Chromosomal fusions and fissions are thought to be less deleterious in holocentric organisms and therefore more likely to be retained. Accordingly, some butterfly clades show a tremendous diversity of chromosome numbers. These clades often show bursts of species diversity, suggesting a role of rearrangements in speciation. To advance our understanding on the importance of repetitive elements and chromosomal rearrangements in speciation, we have generated and compared chromosome-scale genome assemblies for Erebia butterflies with differing chromosome numbers. We found support for a role of some repetitive elements in chromosomal rearrangements. Therefore, differences in repeat expansions might explain why some clades show a greater variation in chromosome numbers and a greater species diversity than others. Overall, we are disentangling the association between repetitive elements, chromosomal rearrangements, and species diversification in holocentric organisms.

102 - Directed evolution of a nascent inhibitor of HIV-1 budding

Steve Cho

Steve Cho, Nels Elde
University of Utah, University of Utah, Howard Hughes Medical Institute

Evolutionary processes have left fingerprints throughout the development of eukaryotic and prokaryotic immune systems. As such, understanding the impact of evolutionary forces on host immunity is crucial to not only the discovery of uncharacterized infection dynamics, but leverage of these findings into orthogonal therapeutic approaches. One such immune process we described is an intriguing adaptation in South American primates and mice; the retrogene, retroCHMP3 manipulates the deeply conserved ESCRT pathway to provide a novel immune function in the setting of ESCRT-dependent virus budding such as that of HIV-1. Primates and mice independently acquired retroCHMP3 about 45 and 7 million years ago respectively, and each ortholog exhibits unique adaptations that converge on the same antiviral property. Here, we developed an experimental evolution model selecting for adaptations to retroCHMP3 and associated ESCRT factors to understand the mutational landscape of retroCHMP3 and the tradeoffs thereof. Our results show directed evolution of retroCHMP3 can drastically improve survival under multiple independent, ESCRT-mediated phenotypes, leading to hints about the context and mechanism for the acquisition of novel immune processes.

141 - Transposable Element expression across mammalian organ development

Rocio Enriquez-Gasca

Rocio Enriquez-Gasca, Margarida Cardoso-Moreira
The Francis Crick Institute, The Francis Crick Institute

Mammalian genomes are rich in transposable elements (TEs), which have been shown to contribute to phenomena as varied as the

concerted regulation of genes to specific phenotypic changes, all of which can impact an organismâ€™s capacity to adapt to its environment. Nonetheless, TEs are often omitted from genomic studies. Here, we explore the transcriptional landscape of TEs throughout the development of seven organs (cerebrum, cerebellum, heart, kidney, liver, ovary and testis) from early organogenesis to adulthood for human, rhesus macaque, mouse, rat, rabbit, opossum, and chicken as an outgroup. We identified considerable differences in TE expression among organs and developmental stages within species and between species. Taken together with each speciesâ€™ TE repertoire, with this work, we seek to uncover niches of transcriptional permissibility, some of which could functionally impact mammalian organ development.

177 - Transposon-mediated evolution of species-specific interleukin receptor variants

Andrea D Ordoñez

Andrea D OrdoÃ±ez, Conor J Kelly, Edward Chuong
University of Colorado, Boulder, University of Washington , University of Colorado, Boulder

Interleukins (ILs) are immune signaling molecules that mediate recruitment, proliferation, and differentiation of cells. IL-signaling must be highly regulated to maintain a balance between immune stimulation and regulation. Variants of IL-receptors have evolved to modulate signaling, as agonists or antagonists. I am exploring the role of transposable elements (TEs) in driving the evolution of species-specific IL-receptor variants. TEs can mediate variable expression of host coding-regions by providing alternative promoters, splice sites, or poly-adenylation signals. Interleukin pathways are highly conserved across mammals, yet transposition events can be unique to species. I'm leveraging RNA-sequencing and TE annotations of reference genomes to computationally uncover a variety of TE-modified transcripts, followed by functionally testing the protein products to characterize their roles in immune regulation. We have identified a gene duplication of IL2 receptor beta (IL2RÎ²), specific to bovine, resulting in a novel coding sequence that retains only the N-terminus domain of the canonical gene. IL-2 is a master signaling molecule of lymphocyte proliferation and differentiation. Expression of this transcript, referred to as bovine IL2RÎ²-soluble (IL2RÎ²-s), is driven by a promoter from a MER41 TE. The protein product is predicted to be a soluble modulatory receptor. To characterize the role of this protein, I will be purifying it and testing it in various cell models and primary bovine immune cells to evaluate differential phenotypic responses, including cell growth, phosphorylation of downstream effector proteins, and differential gene expression of IL2 target genes.

194 - Novel Endogenous Retroviruses Identified in Anuran

Hayley Beth Free

Hayley B Free, Ravinder Kanda, Richard A Nichols
Oxford Brookes University, Zoological Society of London, Oxford Brookes, Queen Mary University of London

Endogenous retroviruses (ERVs) are a class of transposable elements which comprise 10-15% of vertebrate genomes and are known to be a major source of genomic structural variation, contributing to vertebrate genome evolution. While most ERV research has focused on mammals, we are interested in ERV presence and distribution in Anuran species (frogs and toads). We have designed a bioinformatic pipeline to identify and characterise ERV insertions, using iterative in-silico screening to refine detection and assess insertion loci and structure. This pipeline was used to screen for alpha- and gamma- ERVs. While gamma- ERVs are ubiquitous in vertebrate genomes, alpha-like ERVs appear avian-specific, with detection in Amphibians only in the sister order to Anura, Gymnophiona. Here we processed publicly available Anuran genomes (n=48) and have identified multiple loci of alpha-like and gamma-like ERV elements, some of which are shared across Anuran species. Additionally, through phylogenetic and motif analysis, the genome organisation of some full-length insertions appeared variable, suggesting recombination amongst ERVs in Anura. We propose that the rarity of previous reports of

amphibian alpha-like ERVs is likely explained by the limited number of studies using methods that can detect them, as well as their possible recombined states in the genome. These new discoveries suggest that alpha-like ERVs are more widely distributed in vertebrates than originally proposed. These initial steps in characterising ERVs in Anura can provide insight into whether ERV impact on genome evolution is consistent across vertebrates.

252 - Across-species exploration of gene-TE chimeras using an integrated bioinformatics workflow

Albert Garcia Andrés

Albert Garcia AndrÃ©s, Marta Coronado-Zamora, Josefa GonzÃ¡lez
Institute of Evolutionary Biology, CSIC, UPF, Institute of Evolutionary Biology, CSIC, UPF, Institute of Evolutionary Biology, CSIC, UPF

Transposable elements (TEs) constitute a significant portion of eukaryotic genomes, playing a crucial role in genomic evolution. While most studies focus on the role of TEs as providers of regulatory sequences, the presence of gene-TE chimeras, the fusion of a transcript and a TE sequence, introduces a new layer of complexity in understanding genome dynamics. A recent study in Drosophila melanogaster, quantified the contribution of TEs to transcriptome diversification and showed that chimeric transcripts contribute, on average, 43% to the total gene expression, and they provide protein domains for DNA binding, catalytic and DNA polymerase activity. Given the substantial contribution of TEs to transcriptome and protein novelty, we present a workflow designed to detect and analyze chimeric gene-TE transcripts, aiming to unravel their prevalence and dynamics across species, body parts, and conditions.
We have designed an intuitive workflow which integrates cutting-edge bioinformatic tools for TE identification, sequence alignment, and chimeric gene-TE transcripts detection, facilitating a systematic exploration of transcriptomic landscapes. The workflow can take both short and long RNA-seq reads, and provides a simple parametrization that allows it to run with all types of eukaryotic species sequences. By applying this workflow to diverse transcriptomic datasets exposed to various stressors, we are investigating how different stress types influence the generation and abundance of chimeric TEs in Drosophila species. Furthermore, applying the workflow to diverse species including mammals, plants, and fungi, will allow us to explore the contribution of TEs to transcriptome diversification across distantly related species.

339 - Two centuries of transposable element invasions in Drosophila melanogaster

Almorò Scarpa

AlmorÃ² Scarpa, Riccardo Pianezza, Sarah Signor, Robert Kofler
University of Veterinary Medicine, Vienna, Vienna Graduate School of Population Genetics, University of Veterinary Medicine, Vienna, Vienna Graduate School of Population Genetics, North Dakota State University, University of Veterinary Medicine, Vienna

Transposable elements (TEs) significantly influence genome evolution and phenotype. However, the rate of TE horizontal transfer remains a key unresolved question. Previous studies have demonstrated four TE families invading D. melanogaster in the 20th century. Expanding on this, we examined historical specimens spanning an additional 100 years, revealing three LTR retrotransposons (Blood, Opus, and 412) spreading in the 19th century. Furthermore, we identified a novel TE, named Spoink, invading D. melanogaster post-1990s. Employing a novel methodology, we detected three more TEs, the latest appearing globally less than ten years ago. In total, eleven TE families have invaded D. melanogaster over 200 years, expanding the genome by approximately 1 Mbp. This rapid horizontal transfer rate, coupled with the origin of some TEs from American species, suggests human activity may have facilitated inter-species contact, contributing to TE spread.

360 - "Un-trapping”: Does the size of piRNA clusters predict the abundance of transposable element

insertions?

Sarah Saadain

Sarah Saadain, Riccardo Pianezza, Robert Kofler

Institute of Population Genetics, University of Veterinary Medicine Vienna, Vienna Graduate School of Population Genetics, Institute of Population Genetics, University of Veterinary Medicine Vienna, Vienna Graduate School of Population Genetics, Institute of Population Genetics, University of Veterinary Medicine Vienna, Vienna Graduate School of Population Genetics

Identifying the factors that determine the abundance of individual transposable element (TE) families remains a major open question in the field. Under the trap model, it is assumed that an invasion of a TE is stopped when a member of the invading family jumps into a piRNA cluster and thereby triggers the production of piRNAs, which in turn silences the TE. Simulation studies suggest that the total size of all piRNA clusters (as a percentage of the genome) should be the most crucial aspect, where species with a smaller proportion of clusters ought to accumulate many TE insertions and the inverse being true for those with larger proportions of piRNA clusters.
By estimating the size of piRNA clusters and the abundance of individual TE families in 15 Drosophila species, we tested whether the cluster size can predict the observed abundance of the individual TE families. To exclude negative selection against TEs as a potential bias, we focused on young TE families, i.e. families with a low pairwise divergence among the insertions. Contrary to expectations under the trap model, we find no correlation between the size of piRNA clusters and the abundance of TEs. Our work suggests that other mechanisms, such as siRNAs, may largely be responsible for triggering host defense mechanisms to combat invading TEs.

412 - Transposable elements rapidly induce chromosomal rearrangements in locally adapted island Drosophila

Brandon Turner

Brandon Turner, Theresa Erlenbach, Nicholas Stewart, Cathy Moore, Robert Reid, Rebekah Rogers, Rebekah Rogeres

University of North Carolina - Charlotte, University of Georgia, Southern Oregon University, University of North Carolina - Charlotte, University of North Carolina - Charlotte, University of North Carolina - Charlotte, University of North Carolina - Charlotte

By evaluating rearrangements in island Drosophila and comparing to mainland Africa Drosophila we can utilize population differentiation, population genetic statistics and differential expression data to quantify the magnitude of these various genetic mechanisms and their role on evolution. By evaluating how chromosomal rearrangements result in phenotypic changes, we can gain insight on how rapid evolution can reshape populations. We hypothesize that chromosomal rearrangements can be a rare source of innovation that potentially produces cellular changes and, while generally detrimental or neutral, can potentially result in adaptive variation during habitat shifts. We wish to discern whether these "hopeful monsters" among genes can lead to adaptive success in new environments while firmly separating out their effects from that of random genetic drift.

489 - The dynamic genomes of Hydra reveal the anciently active repeat complement of animal chromosomes

Tetsuo Kon

Tetsuo Kon, Koto Kon-Nanjo, Tracy Chih-Ting Koubkova Yu, Diego Rodriguez-Terrones, Francisco Falcon, Daniel E. MartÃ-nez, Robert E. Steele, Elly Margaret Tanaka, Thomas W. Holstein, Oleg Simakov

University of Vienna, University of VIenna, Heidelberg University, Institute for Molecular Pathology, Institute for Molecular Pathology, Vienna BioCenter PhD Program, Doctoral School of the University of Vienna and Medical University of Vienna, Pomona College, University

of California, Irvine, Institute for Molecular Pathology, Heidelberg University, University of Vienna

Many animals exhibit highly conserved chromosomal homologies that pre-date their ancient origin. Despite such conservation, evolutionary forces behind the retention, expansion, and contraction of chromosomal elements, and the resulting macro-evolutionary implications, remain unclear. We present a comprehensive stem-cell resolved genomic and transcriptomic investigation of the fresh-water cnidarian Hydra, which show remarkable regenerative capacity, an ability to propagate clonally, and a lack of aging. Using single-haplotype telomere-to-telomere genome assemblies of two recently diverged hydra strains, we elucidated how the macro-evolutionary history of chromosomal elements is shaped by both ancient and recent transposable element (TE) expansions. Using fluorescence-activated cell sorting (FACS) followed by genome sequencing we revealed distinct TE insertion profiles in hydraâ€™s three stem cell lineages. Transcriptomic analyses revealed that over 14,000 transcripts contained almost entire sequences of TEs, and further classification into families, subfamilies, and individual loci reveals cell type-specific TE expression. The active TEs include elements that differentially contribute to genome size changes and persistent structural variants around loci associated with cell proliferation. We identified 14 TE families primarily involved in this role which are largely composed of DNA elements such as TcMar-Tc1, hAT-Ac, and CMC-EnSpm. Evolutionary analysis showed that these families constitute a highly conserved TE core in eukaryotic and metazoan genomes. Our results suggest an ancient role for these core TEs as genomic components that persist beyond ancient chromosomal homologies.

623 - Evolution of retrocopies of protein-coding genes in the context of HUSH complex repression.

Izabela Makalowska

Joanna Kozlowska-Maslon, Joanna Ciomborowska-Baseer, Izabela Makalowska
Adam Mickiewicz University, Greater Poland Cancer Centre, Adam Mickiewicz University, Adam Mickiewicz University

Retroposition is a process in which mRNA is reverse transcribed into a cDNA and reintegrated into the genome. Like other retroelements, retrocopies provide new genetic material, but also pose a threat to genome integrity. Therefore, it is fundamental to control this threat of invasion of RNA-derived elements. One of developed defence strategies is a repression by the HUSH complex. It was determined that sequences repressed by the HUSH are reverse-transcribed, long, intronless, and transcriptionally active elements with a high adenine content.

We performed an analysis of 4463 retrocopies of human genes focused on three factors: exon length, intron presence, and adenine content. Based on RNA-seq analysis and annotations, retrocopies were divided into three groups: protein-coding, noncoding, and nonexpressed. All retrocopies had significantly lower expression than their progenitors. The expression of noncoding retrocopies was negatively correlated with exon length. Protein-coding retrogenes had the highest expression level and accumulated the highest number of substitutions. However, their adenine content, unlike other categories, did not differ from their progenitors.

The results of our study show that the presence of long exons and a high level of adenine have a negative effect on the level of retrocopiesâ€™ expression. However, exon length appears to be more important than adenine content. We also showed that intron gain provides some protection from possible HUSH silencing and makes the expression level less depending on the transcript length. Nevertheless, this is true only when the level of adenine is low.

784 - Functional study of DPLG3, a PIF transposable element domesticated protein, in Drosophila melanogaster

María del Pilar Castellanos

MarÃ-a del Pilar Castellanos, Chathuri Devmika Wickramasinghe, Fatema Ruma, Esther BetrÃ¡n
The University of Texas at Arlington, The University of Texas at Arlington, Coriell Institute for Medical Research, The University of Texas at

Arlington

Molecular domestication of transposable element (TE) proteins can be viewed as the process of recycling the TE selfish protein/s for cellular functions. This process is gaining attention in evolutionary biology due to increased genome sequencing and functional analyses. Here, we analyze a PIF/Harbinger family of TE proteins domesticated gene in Drosophila melanogaster, Drosophila PIF-like gene 3 (DPLG3). We tagged DPLG3 with HA and observed its localization and co-localization with DNA in ovaries. Knock-out flies generated via CRISPR-Cas9 displayed abnormal gonadal phenotypes that are partial penetrant and maternally inherited, with 10-15% of females and fewer males affected. RNA-Seq analysis of non-rudimentary DPLG3-KO ovaries revealed differentially expressed genes, TEs and long non-coding RNAs suggesting a potential regulatory role for DPLG3. Our hypothesis posits that DPLG3, because of its nuclear localization, has retained its DNA binding domain and likely functions as a transcription regulator. Further studies aim to elucidate the role of this domesticated transposase in Drosophila germline development.

791 - Chromosome-level genome assembly of Candidia barbatus (Cypriniform) reveals an underlying mechanism in chromosome evolution

Jui-Hung Tai

Jui-Hung Tai, Tzi-Yuan Wang, Yu-Wei Wu, Feng-Yu Wang, Shih-Pin Huang, Shu-Miaw Chaw, Hurng-Yi Wang
National Taiwan University, Academia Sinica, Taipei Medical University, National Applied Research Laboratories, Academia Sinica, Academia Sinica, National Taiwan University

The tribe Opsariichthyini is a group of small East Asia minnow, belonging to Xenocyprididae, Cypriniformes. While all Xenocyprididae species share a common chromosome count of n=24, the tribe Opsariichthyini exhibits chromosome number ranging from 24 to 39. The formation of the chromosome number difference in this group is important for studying the evolution of chromosomes in vertebrates. We assembly the first chromosome-level genome of Candidia barbatus, which is one of the tribe Opsariichthyini members and endemic to Taiwan. Through the synteny analysis between C. barbatus (n=24) and Opsariichthys bidens (Opsariichthyini, n=39), we found the dramatic increase of chromosome number is caused by 15 chromosome fission events. Most importantly, we observed a high percentage of LINE/L2 repeat elements surrounding the fission regions, which are rare throughout the entire genome. Unlike in other cyprinids, LINE/L2 has recently undergone an expansion in the tribe Opsariichthyini. In addition, a similar pattern of chromosome fission can be found on Human chromosome 14 and 15. We hypothesized that the expansion of repeat elements may play an important role in chromosome fission and evolution.

808 - Transposable element copy variation in cultivated apples, and their wild apple relatives

Anthony Venon

Anthony Venon, Xilong Chen, AurÃ©lie Mesnil, Karine Alix, Amandine Cornille
INRAE, CNRS, CNRS, AgroParisTech, CNRS

Transposable elements (TEs) are repetitive sequences of the genome that can play critical roles in major agronomic traits. In addition, TEs can be involved in the adaptation of populations to new environments. The role of TEs in the domestication of perennial crop, including fruit trees, still needs to be studied. Exploiting short-read sequencing data, we annotated TEs and estimated their copy numbers among populations of the cultivated apple tree (Malus domestica) and its wild relatives, representing a total of 230 accessions. To do so, first, we built an original bioinformatic pipeline combining short-reads and TE predictors (RepeatExplorer and REPdenovo) to obtain an exhaustive TEs database for the genus Malus. Subsequently, we annotated this database with different bioinformatic tools (PASTEC and

TEsorter). This database contains a total of 158k consensus sequences with a predominance of LTR-retrotranposons. We then mapped the raw short read sequences of the 21 accessions onto this TE database to get the copy number of each TE consensus. Statistical analyses revealed significant TE copy number variations between cultivated and wild apple trees and among the cultivated apple trees depending on their uses. Further analyses are being done to identify which specific TEs mainly explain the observed variation. Our study represents a significant advance in understanding the role of TEs in the domestication of fruit trees and their role in genome evolution and adaptation.

810 - The extensive diversity and speciation in the Impatiens genus (Balsaminaceae) may be linked to TE-associated differential gene disruption

Sudhindra R Gadagkar

Sudhindra R Gadagkar, Kristina Buss
Midwestern University, Arizona State University, Arizona State University

The Impatiens genus (Balsaminaceae) is one of the most speciose among the angiosperm genera, comprising over 1000 species spread over much of the tropical and temperate parts of the old world. Of particular interest is the contrast with its only sister genus, Hydrocera, which is monotypic and confined to specific regions of South and Southeast Asia. We are interested in unraveling genomic reasons, if any, for this contrast and queried the only two Impatiens species with fully sequenced genomes (I. capensis and I. glandulifera), with particular emphasis on the role of transposable elements (TEs) in generating diverse phenotypes in the Impatiens genus. We found twice as many TEs in I. glandulifera as in I. capensis, although I. capensis contained five times as many genes associated with TEs. In particular, the Gypsy family and LTRs, and to some extent, Copia, are highly enriched in their association with genes in I. capensis. While the number of GO annotated genes available for functional enrichment analysis by TE density is similar between the two genomes, a larger than expected enrichment of cytochrome assembly, carboxylic acid metabolism, protein binding functions, and mitochondrial membrane localization is seen in I. glandulifera, while terms associated with transcription regulation processes, hydrolase functions, dsRNA binding, extracellular space, and ribosome association are more enriched in I. capensis. We speculate that enrichment of energy-associated functional terms, such as cytochrome assembly and mitochondrial localization, are at least partially responsible for the phenotypic differences between these two species.

819 - Unraveling the Repetitive Element Landscape: Insights into Trypanosomatid Genomes.

Percy Omar Tullume-Vergara

Percy Omar Tullume, Kelly Olivos, Joao M. Alves
Department of Parasitology, Institute for Biomedical Sciences, University of Sao Paulo, Department of Parasitology, Institute for Biomedical Sciences, University of Sao Paulo, Department of Parasitology, Institute for Biomedical Sciences, University of Sao Paulo

Trypanosomatid flagellate parasites constitute a well-recognized group within Kinetoplastida. Currently, the family Trypanosomatidae can be categorized into monoxenous and dixenous based on their lifecycle. Presently, four superfamilies have been identified in trypanosomatids, which are classified under Class I retrotransposons. Specifically, Ingi and SLACS are located within the LINE order, while VIPER and TATE are associated with DIRS elements. In this study, we conducted annotation and classification of repeat content across 57 trypanosomatid whole-genomes, shedding light on the proportions and diversity of repetitive elements (RE) and transposable elements (TE) within a phylogenetic framework. Our analysis yielded 583 consensus sequence models across the dataset. Notably, we observed instances of TE family contraction, particularly within the genus Leishmania.
Moreover, our findings revealed that the abundance of transposable elements does not exhibit a direct correlation with genome size,

although repetitive elements displayed a weak positive correlation. We constructed a phylogenetic tree for the 57 trypanosomatid species using a supermatrix approach, employing both Maximum Likelihood (ML) and Bayesian Inference (BI) methods. In addition to these insights, we have compiled a species-specific library intended to aid the trypanosomatid community in improving the annotation process. Our analysis demonstrates that all four superfamilies identified in trypanosomatids are present from monoxenous early stages to dixenous parasites, with some species encompassing all superfamilies observed thus far. This comprehensive examination of TE landscapes contributes significantly to our understanding of the evolutionary dynamics shaping the genomes of trypanosomatids.

875 - Insights into PIF transposase-derived gene functionality in germline development in Drosophila

Chathuri Devmika Wickramasinghe

Chathuri Devmika Wickramasinghe, MarÃ-a del Pilar Castellanos, Diwash Jangam, Fatema B. Ruma, Dragomira N. Markova, CÃ©dric Feschotte , Esther BetrÃ¡n
University of Texas at Arlington, University of Texas at Arlington, University of Texas at Arlington, University of Texas at Arlington, University of Texas at Arlington, Cornell University, University of Texas at Arlington

Many studies have demonstrated the contribution of transposable elements (TE) derived proteins to the emergence of new host proteins. Across Drosophila genus, there are seven PIF TE-derived genes known as Drosophila PIF Like Genes (DPLG1-7), domesticated from PIF/Harbinger transposase superfamily. We are studying potential functions of the four DPLGs (DPLG1-4) present in D. melanogaster, which are highly conserved and were likely independently domesticated. DPLG1-4 are highly expressed in gonads and nervous system. Here, we focus on DPLG1 and DPLG4 female germline function. We have studied localization of DPLG1-HA and DPLG4-HA tagged proteins in ovaries. RNA-Seq analysis of ovaries from CRISPR-Cas9 generated knock-out flies revealed that the absence of these genes causes differential expression of a set of overlapping genes in ovaries, with positively correlated changes in expression. Several genes in piRNA pathway were found among DE genes in DPLG1-KO ovaries. Further, several TE families were differentially expressed in both DPLG1-KO or DPLG4-KO ovaries. Most upregulated TEs belonged to Gypsy LTR retrotransposon superfamily, whereas downregulated TEs were telomeric non-LTR retrotransposons. Interestingly, preliminary small RNA-Seq analysis revealed an increase in piRNAs targeting these telomeric elements, which may explain the reduction in telomeric elements. These findings strongly suggest that DPLG1 and DPLG4 may have potential regulatory functions in Drosophila telomere elongation through piRNA pathway. Collectively, our studies imply complex interactions among transposase-derived genes repeatedly co-opted from the same TE superfamily, to function within the same cellular pathways.

920 - Epistasis analysis reveals adaptive separation-of-function mutants in experimentally evolved Saccharomyces cerevisiae

Michelle Hays

Michelle Hays, Katja Schwartz, Gavin Sherlock
Stanford University, Stanford University, Stanford University

Previously, we showed that several mutational mechanisms give rise to adaptation when Saccharomyces cerevisiae are experimentally evolved under fluctuating nitrogen starvation. Adaptive mutations arise from SNPs, but also from other mutational mechanisms such as increased retrotransposon activation and microhomology mediated recombination leading to structural variation and gene conversion. Different adaptive alleles differ in their fitness benefit. The notable differences in adaptive alleles at the same loci, both in fitness and putative gene products, led us to pursue epistatic analysis of several beneficial mutations that arose in the experimental evolution at recurrently mutated genomic loci (such as MEP1, GAT1, and PAR32). Here, through double mutant analysis, we see that negative

epistasis is most common, with double-mutant fitnesses most often being similar to the fitness increase of the higher of the two single mutants. These data largely best fit a model of diminishing returns, but with a few notable exceptions. In specific examples we observe unexpected combinatorial fitness effects, suggesting that some adaptive alleles may in fact be separation of function or gain of function mutants. For example, beneficial missense mutations at the nitrogen transporter MEP1 exhibit different epistatic interactions and overall fitness changes than novel 3â€™ retrotransposon insertions at the same genomic locus. Nitrogen catabolism regulation in S. cerevisiae is complex, with several feedback loops: these evolved mutations provide means for disentangling adaptation in a complex regulatory network and exploring the mechanism of specific subfunctions in gene products most likely to impact fitness and organism adaptation.

923 - Transposable element polymorphism across the tree of life varies with speciation rate

Pío Sierra

PÃ-o Sierra, Richard Durbin

Department of Genetics - University of Cambdrige, Department of Genetics - University of Cambdrige

Transposable element (TE) activity makes a major contribution to genome diversity within species. We compared high quality assemblies of hundreds of species across the tree of life using a new polymorphism-first TE detection tool, pantera, that has high power to find recently active, low copy number TE families, and found great differences in the levels of polymorphic TE content. Interestingly, species within lineages that have undergone recent radiation are enriched for TE polymorphisms, reflecting high levels of recent TE activity. Amongst the most dynamic TE families are Mavericks, also known as Polintons, large DNA transposons closely related to adenoviruses that can be found in many eukaryotes. In 404 new genome assemblies of Lepidoptera from the Darwin Tree of Life project, pantera found 312 new families of Maverick elements, more than the 234 previously in RepBase, most of which appear to still be active. Their protein sequences are under strong purifying selection and a phylogenetic analysis shows there are at least five different clades of Mavericks in Lepidoptera which can be identified not only by homology of the proteins, but also by the order and orientation of the proteins in the sequence. Most of the branches in their phylogeny are consistent with vertical transmission, suggesting that, if present, horizontal transmission has been very rare in Lepidoptera. Furthermore, we suggest that the structure of Mavericks is better explained as the combination of two individual elements, in a similar way to the segments of Bombyx mori densovirus 3 (BmDNV-3).

990 - SINE elements contribute to regulatory regions during cardiomyocyte differentiation in humans and chimpanzees

John Dennis Hurley

John D Hurley, Raodatullah O Abodunrin, Michelle C Ward

University of Texas Medical Branch, University of Texas Medical Branch, University of Texas Medical Branch

Humans and chimpanzees differ in anatomy and physiology despite their genomes being approximately 98% identical. Differential gene regulation is the major driver of the phenotypic differences between these closely related species. Transposable Elements (TEs) comprise fifty percent of primate genomes and can contribute to gene regulation within and between species. However, their relative contribution appears to differ between pluripotent and somatic cell types. To study the evolutionary dynamics of TE regulation during development, we utilize an in vitro system to differentiate iPSCs of humans and chimpanzees into cardiomyocytes. We collected cells at four stages of differentiation, as determined by 50% of the cell population expressing stage-specific markers by flow cytometry. Stages include pluripotent cells (OCT3/4+), mesoderm (Brachyury+), cardiac mesoderm (ISL1+), and terminal cardiomyocytes (TNNT2+). To study TE regulatory dynamics, we profiled the enrichment of H3K27ac, a marker of active regulatory regions, by CUT&Tag followed by sequencing. We detected tens of thousands of H3K27ac regions in orthologous regions in both species, with the majority being variable across

differentiation stages. Stage-variant regions in humans are less likely to be conserved across species than stage-invariant regions. Overall, conserved H3K27ac regions are enriched for orthologous SINE elements amongst all annotated orthologous TEs, suggesting involvement in gene regulation. Conserved stage-variant regions are more likely to overlap an annotated orthologous TE than conserved stage-invariant regions, suggesting TEs contribution to stage-specificity. We are currently collecting data from additional individuals to investigate the stage- and species-specificity of TE family regulation.

1130 - Evolution of transposable elements as the source of extracellular RNA

Cei Abreu-Goodger

Isaac Martinez-Ugalde, Kyriaki Neophytou2, Yenetzi Villagrana-Pacheco, Xiaochen Du, Mike Evans, Lewis Stevens, Mark Blaxter, Amy Pedersen, Amy H Buck, Cei Abreu-Goodger
Institute of Ecology and Evolution, University of Edinburgh, Institute of Immunology and Infection Research, University of Edinburgh, Institute of Ecology and Evolution, University of Edinburgh, Institute of Immunology and Infection Research, University of Edinburgh, Institute of Immunology and Infection Research, University of Edinburgh, Tree of Life, Wellcome Sanger Institute, Tree of Life, Wellcome Sanger Institute, Institute of Ecology and Evolution, University of Edinburgh, Institute of Immunology and Infection Research, University of Edinburgh, Institute of Ecology and Evolution, University of Edinburgh

Beyond self-replicating, transposable elements (TEs) can be a source for new functions. Small RNAs (sRNAs) and Argonaute proteins are important for silencing TEs, however, sRNAs can also be effector molecules for some parasites. The parasitic nematode Heligmosomoides bakeri can modulate the immune response of mice by secreting a variety of molecules including proteins, lipids, sRNAs and the extracellular worm Argonaute (exWAGO). Importantly, secreted sRNAs are mostly derived from TEs, suggesting a relationship between TE silencing and parasitism. H. bakeri has one of the largest, TE-rich, nematode genomes. Nevertheless, little is known about the evolution of these TEs, or their functions in parasitism.
Using immunoprecipitation (IP) of exWAGO in H. bakeri, and sequencing bound sRNAs, we found that they map to all major classes of TEs, with a preference for particular families. LTR/Pao and LTR/Gypsy retrotransposons seem particularly relevant, producing ~8-fold more sRNAs than expected. IP experiments in H. polygyrus, Nippostrongylus brasiliensis, Teladorsagia circumcincta, and Ancylostoma ceylanicum confirm that, although their genomes vary in size and TE composition, exWAGO preferentially loads sRNAs derived from LTRs. Furthermore, we found that specific TE families are relevant for secreted exWAGO in H. bakeri, and are conserved across parasites and free-living nematodes. We propose exWAGO as an ancestral regulator of retrotransposon activity. Parasitic nematodes have since gained the capacity to secrete exWAGO and sRNAs during infection, alongside an expanded and diverged repertoire of transposable elements in their genomes.

47 - Gene expansion and multiple combinations confer retromer functional diversity in traffic in Entamoeba histolytica

Tomoyoshi Nozaki

Natsuki Watanabe, Kumiko Nakada-Tsukui, Tomoyoshi Nozaki
The University of Tokyo, National Institute of Infectious Diseases, The University of Tokyo

The protease transport is essential for various cellular functions such as protein degradation, maturation, activation, and extracellular matrix remodeling, migration, and intercellular communication. It is elaborately regulated in both antero- and retrograde directions, the latter of which is mainly mediated by the retromer complex. The retromer core complex, cargo-recognition complex (CRC), is composed of Vps26, Vps29, and Vps35, and accessory proteins including sorting nexin and Rab7, which interact with CRC to control its localization and

function. It has been demonstrated that the retromer regulates recycling of plasma membrane- and endosome-localized receptors. The human intestinal protozoan parasite Entamoeba histolytica causes amoebic dysentery and its parasitism and pathogenesis largely depend upon proteases and their transport mechanisms. In this study, we investigated gene repertoire expansion and functions of highly diversified heterogeneous retromer CRCs in E. histolytica. Immunoprecipitation of individual Vps26 and Vps35 isotypes led to identification of multiple heterogeneous CRCs, which form multimeric super complex. Vps35-1 and Vps35-2, two predominant Vps35 isotypes, are in the multiple forms of CRCs, selectively interact with specific effectors, and are engaged with distinct functions. Vps35-1 controls the formation of super complex and recruitment of CRC to the trogosome, while Vps35-2 regulates CP transport and secretion. This study provides that unprecedented diversification of the repertoire of CRC components and heterogeneous combination and oligomer formation of CRCs allow functional diversity in membrane traffic.

56 - Sexual size dimorphism is associated with gene family contractions in brain development in mammals

Benjamin Padilla-Morales

Benjamin Padilla-Morales, Alin P. AcuÃ±a-Alonzo, Huseyin Kilili, Atahualpa Castillo-Morales, Karina Diaz, Kathryn H. Maher, Evangelos Mourkas, Tamas Szekely, Martin-Alejandro Serrano-Meneses, Diego Cortez, Sergio Anccona, Araxi O. Urrutia
University of Bath, Milner Centre for Evolution, Instituto de Ecologia, University of Bath, Milner Centre for Evolution, Cardiff University, University of Bath, Milner Centre for Evolution, Instituto de Ecologia, NERC environmental Omics Facility, Ineos Oxford Institute, University of Bath, Milner Centre for Evolution, Universidad de las Americas Puebla, Centro de Ciencias GenÃ³micas, Instituto de Ecologia, University of Bath, Milner Centre for Evolution, Instituto de Ecologia

In mammals, sexual size dimorphism can serve as a proxy for sexual selection strength. However, its association with genomic evolution remains unexplored. Gene family size evolution can reflect shifts in the relative importance of different molecular functions. Our analysis of 124 mammalian species reveals significant changes in gene family size related to sexual size dimorphism. Expanding gene families enriched in olfactory sensory perception functions display pronounced sex-biased gene expression. However, contracting gene families were associated with brain development functions, exhibiting higher gene expression in the brain and prenatal stages. These results suggest that sexual selection is associated with selective pressures influencing highly sexually dimorphic species, where olfactory-associated gene families are under expansion. In contrast, brain development-associated gene families are under contraction. These insights illustrate the complex interplay between sexual selection, gene family size evolution, and their roles in mammalian brain development, providing a valuable understanding of mammalian genome evolution.

57 - The role of gene duplication in facilitating divergent patterns of gene expression across the monarch butterfly metamorphosis

James G. DuBose

James G. DuBose, Jacobus C. de Roode
Emory Univeristy, Emory University

The diversification of many lineages throughout natural history has frequently been associated with evolutionary changes in life cycle complexity. However, our understanding of the processes that facilitate divergence in the morphologies and functions expressed by organisms throughout their life cycles is limited. Theory suggests that the expression of traits is decoupled across life stages, thus allowing their evolutionary independence. Although trait decoupling between stages has been described in many studies, explanations of how said decoupling evolves have seldom been considered. Because the different phenotypes expressed by organisms throughout their

life cycles are coded for by the same genome, trait decoupling must be mediated through divergence in gene expression between stages. Gene duplication has been identified as an important mechanism that enables divergence in gene function and expression between cells and tissues. Therefore, we suggest that because stage transitions across life cycles require changes in tissue types and functions, gene duplication has an important role in facilitating the evolution of life cycle complexity. To test this idea, we examined the temporal changes in gene expression across the monarch butterfly (Danaus plexippus) metamorphosis. We found that within homologous groups, more phylogenetically diverged genes exhibited more distinct temporal expression patterns, and that this relationship scaled such that more phylogenetically diverse homologous groups showed more diverse patterns of gene expression. Furthermore, we found that duplicate genes showed increased stage-specificity relative to singleton genes. Overall, our findings suggest an important role of gene duplication in the evolution of trait decoupling across complex life cycles.

68 - Evolution and Development of the Autopodium in Microteiid Lizards (Squamata: Gymnophthalmidae): An Evo-Devo Approach

Joana F Morais

Joana F Morais, Tiana Kohlsdorf

University Of SÃ£o Paulo, University Of SÃ£o Paulo

The Gymnophthalmidae family of lizards has a Neotropical distribution and is characterized by great ecological and morphological diversity. This group brings together two lineages composed of species that exhibit different degrees of reduction in locomotor limbs, with the absence of several elements in the autopodium region (manus and pes) and even serpentine morphology, and evidence of reversal of the loss of digits in the genus Bachia. Limb reduction is a very frequent phenomenon in Squamata, and microteiid lizards represent an ideal group for studying genetic processes related to the reduction and loss of osteological elements in limbs. The evaluation of genes expressed during the formation of locomotor limbs, including regulatory regions related to their control and regulation, makes it possible to evaluate the evolution of morphological diversity in the Gymnophthalminae family from a perspective that integrates evolution and development. Knowledge of genetic pathways related to the emergence of evolutionary novelties in certain clades, such as fossoriality in the group studied, enriches studies of evolutionary developmental biology (EvoDevo) by allowing the assessment of how evolutionary changes originate and are disseminated. The present research project aims to investigate genetic differences in developmental genes between species of Gymnophthalmini (reduction mainly of the forelimbs) and species of Bachiini (reduction mainly of the hind limbs). The results will contribute significantly to the knowledge of the genetic bases related to the evolution of phenotypic diversity in Squamata, reinforcing the central role of EvoDevo in the process of expanding the Modern Synthesis of Evolution.

81 - Polygenic adaptation to heavy metals in a legume host plant and structural adaptation in a symbiotic rhizobia partner

Tim Paape

Timothy Paape, Aditi Bhat, Reena Sharma, Sanhita Chakraborty, Ankita Mishra, Robert Bowers , Tanja Woyke, Peter Tiffin, Jose Pueyo

USDA-ARS Children's Nutrition Research Center, Institute for Advancing Health through Agriculture, Brookhaven National Laboratory, Quorum Bio, Brookhaven National Laboratory, Institute for Advancing Health through Agriculture, Institute for Advancing Health through Agriculture, DOE Joint Genome Institute, DOE Joint Genome Institute, Dept. of Plant and Microbial Biology, CSIS Madrid Institute of

Agricultural Sciences

The heavy metals cadmium (Cd) and mercury (Hg) are highly toxic to most organisms including plants and their associated root microbes. In the Medicago-Sinorhizobium system, both partners possess ion detoxification mechanisms that are essential for free-living and in symbiotic life stages. We used a combination of genome assembly of heavy metal adapted rhizobia, dual-transcriptomics of host-plants and rhizobia in nodules, and X-ray fluorescence microscopy to identify ion distributions in root nodules. Genome assemblies of Sinorhizobium medicae with high variation in Cd and Hg tolerance showed structural differences between strains, including a Mer operon in the most Hg-tolerant strains. While independent mercury reductase A (merA) genes are prevalent in Sinorhizobium, complete Mer operons are rare and often vary in their gene organization. Differentially expressed genes (DEGs) in S. medicae inside of nodules showed that rhizobia and host-plant tolerance affected the number of DEGs. Aside from Mer operon genes, nif genes which are involved in nitrogenase activity in S. medicae showed significant upregulation in the most Hg-tolerant strain in nodules of the most Hg-accumulating host-plant, indicating a genotype-by-genotype interaction that influences nitrogen-fixation under stress conditions. Transfer of the Mer operon to low-tolerant strains resulted in an immediate increase in Hg tolerance, indicating that the operon is solely necessary to confer hypertolerance to Hg, despite paralogous merA genes present elsewhere in the genome. This study demonstrated the role of the Mer operon in Hg hypertolerance in nitrogen-fixing rhizobia.

84 - Fine tuning AAV: machine learning the capsid

Ravi Kalia

Ravi Kalia, Joseph Rabinowitz
Genial Labs LLC, Kriya Therapeutics, Kriya Therapeutics

Viruses have evolved mechanisms to enter cells, deliver their nucleic acid to the correct
compartment, express genes, make copies of the delivery system and genome, and exit the
cell to continue the cycle. Adeno-associated virus (AAV), a dependoparvovirus, lost the ability to complete its life cycle without functions carried by a helper virus, a desirable quality for
a gene therapy vector. AAV serotypes have achieved commercial success as a gene medicine
treating RPE65 mutation-associated retinal dystrophy (LuxturnaÂ®, AAV2 subretinal injection), spinal muscular atrophy (ZolgensmaÂ®, AAV9 intravenous injection [IV]), AADC deficiency (AAV2, putamen injection 2.8e11 vg/putamen), hemophilia B (RoctavianÂ®/HemgenixÂ®,
AAV5 IV), Duchenne muscular dystrophy (Elevidys, AAVrh74 IV). However, the natural serotypes may have limitations depending on the therapeutic application in transduction efficiency, tropism, neutralizing antibody resistance, and blood brain barrier crossing among
other shortcomings. To address these limitations artificial capsids are being developed using
synthetic biology to deliver the gene therapy. This review will focus on machine learning
(ML) as the guiding design principle to improve AAV capsids.

88 - Hidden in Chaos: An Ancient Pathogen Resistance Complex in a Hyper-Divergent Region Challenges Evolutionary Paradigms

Suha Naser-Khdour

Suha Naser-Khdour, Dieter Ebert
University of Basel, University of Basel, University of Helsinki

Unveiling the hidden layers of genomic diversity within species is critical for understanding the complex relationships between genes, environment, and evolution. This study investigates a fascinating case study: the Pasteuria Resistance Complex (PRC) in the crustacean Daphnia magna, offering crucial defense against the virulent bacterium Pasteuria ramosa.

While traditional studies often focus on single nucleotide polymorphisms (SNPs) and copy-number variations (CNVs), we employed advanced techniques like pan-genome analysis and tree reconciliation. These revealed an unexpected location for the PRC â€“ a highly diverse region within the Daphnia genome. This region boasts remarkable ancient origins, tracing its roots back over 200 million years, making it one of the oldest identified resistance loci to date. The level of polymorphism within this region is so extreme that haplotypes â€“ usually similar stretches of DNA â€“ have lost most of the sequence homology, highlighting its remarkable evolutionary vitality.

This unique combination of an ancient resistance complex and a dynamic, hyper-divergent genomic environment presents an interesting scientific mystery. By delving into this case study, we aim to shed light on the potential role of such regions in shaping the evolution and long-term maintenance of resistance loci. Our findings offer valuable insights into the evolutionary forces shaping disease resistance and adaptation, not only in Daphnia but potentially across the entire Cladocera class. This research opens exciting avenues for further exploration into the complex relationships between diverse genomic landscapes and the evolution of crucial defense mechanisms in various organisms.

157 - Comparative genome microsynteny illuminates the fast evolution of nuclear mitochondrial segments (NUMTs) in mammals

Zixia Huang

Zixia Huang, Marek Uvizl, Sebastien Puechmaille, Sarahjane Power, Martin Pippel, Samuel Carthy, Wilfried Haerty, Eugene Myers, Emma Teeling

University College Dublin, Charles University, University of Montpellier, University College Dublin, Max Planck Institute of Molecular Cell Biology and Genetics, University College Dublin, Earlham Institute, Max Planck Institute of Molecular Cell Biology and Genetics, University College Dublin

The escape of DNA from mitochondria into the nuclear genome (nuclear mitochondrial DNA, NUMT) is an ongoing process. Although pervasively observed in eukaryotic genomes, their evolutionary trajectories in a mammal-wide context are poorly understood. The main challenge lies in the orthology assignment of NUMTs across species due to their fast evolution and chromosomal rearrangements over the past 200 million years. To address this issue, we systematically investigated the characteristics of NUMT insertions in 45 mammalian genomes and established a novel, synteny-based method to accurately predict orthologous NUMTs and ascertain their evolution across mammals. With a series of comparative analyses across taxa, we revealed that NUMTs may originate from nonrandom regions in mtDNA, are likely found in transposon-rich and intergenic regions, and unlikely code for functional proteins. Using our synteny-based approach, we leveraged 630 pairwise comparisons of genome-wide microsynteny and predicted the NUMT orthology relationships across 36 mammals. With the phylogenetic patterns of NUMT presence-and-absence across taxa, we constructed the ancestral state of NUMTs given the mammal tree using a coalescent method. We found support on the ancestral node of Fereuungulata within Laurasiatheria, whose subordinal relationships are still controversial. This study broadens our knowledge on NUMT insertion and evolution in mammalian genomes and highlights the merit of NUMTs as alternative genetic markers in phylogenetic inference.

173 - Evolution of Evolvability In Rapidly Evolving Populations (if accepted for talk); The Dynamics of Horizontal Gene Transfer In Rapidly Adapting Populations (if accepted for poster)

James Ferrare

James Ferrare, Benjamin Good
Stanford University, none, Stanford University

(if accepted for talk): Mutations can alter the short-term fitness of an organism, as well as the rates and benefits of future mutations. While numerous examples of these evolvability modifiers have been observed in rapidly adapting microbial populations, existing theory struggles to predict when they will be favored by natural selection. Here, we develop a mathematical framework for predicting the fates of genetic variants that modify the rates and benefits of future mutations in linked genomic regions. We derive analytical expressions showing how the fixation probabilities of these variants depend on the size of the population and the diversity of competing mutations. We find that competition between linked
mutations can dramatically enhance selection for modifiers that increase the benefits of future mutations, even when they impose a strong direct cost on fitness. However, we also find that modest direct benefits can be sufficient to drive evolutionary dead-ends
to fixation. Our results suggest that subtle differences in evolvability could play an important role in shaping the long-term success of genetic variants in rapidly evolving microbial populations.
(if accepted for poster): In rapidly adapting populations, horizontal gene transfer can facilitate adaptation by breaking down linkage disequilibrium and recombining beneficial mutations onto a single genetic background. When this process is sufficiently rare, basic aspects of evolutionary dynamics are poorly understood. Here, we derive analytical expressions showing when recombination events are favored by natural selection and how these events affect adaptation.

176 - A population genetics-based approach to uncover 3D genomics variation among closely related individuals

Juan Antonio Rodríguez

Juan Antonio Rodriguez, Sarah S. T. Mak, Maria Marti-Marimon, Sofia Marcos, Claudia Fontsere, Antton Alberdi, Marc A. Marti-Renom, M Thomas P Gilbert
Center for Evolutionary Hologenomics, Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Center for Evolutionary Hologenomics, Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Centre Nacional dâ€™AnÃ lisi GenÃ²mica, CNAG, Center for Evolutionary Hologenomics, Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Applied Genomics and Bioinformatics, University of the Basque Country (UPV/EHU), Center for Evolutionary Hologenomics, Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Center for Evolutionary Hologenomics, Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Centre Nacional dâ€™AnÃ lisi GenÃ²mica, CNAG, Center for Evolutionary Hologenomics, Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen

Characterizing genomic diversity is crucial to exploiting genetics in the basic and applied sciences. A key aspect that is gaining increased attention, is how the genome is not just a one-dimensional linear thread, but rather a three-dimensional (3D) folded polymer. Nuclear DNA fiber folds into non-random 3D structure, whose preservation is vital for correct cell homeostasis and gene-regulatory mechanisms. These mechanisms have recently gained considerable interest, thanks to chromosomal conformation capture techniques, that allow reconstruction of the genomic structure. However, 3D genomics analyses are yet to explore how 3D genome structure varies across individuals within a population. Neglecting to consider such variation at the population scale (should it exist), could result in sampling biases and/or yield exaggerated differences in 3D structure when comparing pairwise conditions. Using a cohort of 16 chickens from a single population raised in highly controlled conditions as a model, we evaluated the populational variability of 3D genome structure at several levels. We found that up to a 10% of the genome compartmentalization may show variability between individuals of a same population. While a significant fraction of 3D genomic variability for autosomes can be explained by sex and/or age, a large fraction of the

remaining individual, populational variance stays unaccounted. As our chicken individuals derive from a controlled captive breeding population, we hypothesize that 3D genome variation in wild populations can be even larger. Hence, acknowledging the understudied standing 3D genomic variation of samples can be helpful to correctly frame and interpret discoveries.

184 - The Genome of the Blueberry Stem Gall Wasp

Stephanie O Castro-Marquez

Stephanie O Castro-Marquez, Glen R Hood, Mike Sergeant, Carlos A Farkas-Pool, Walter Wolfsberger, Ellen O Martinson, Taras K Oleksyk

Oakland University, Wayne State University, Wayne State University, Universidad CatÃ³lica de la SantÃ-sima ConcepciÃ³n, Oakland University, University of New Mexico, Oakland University

The blueberry stem-gall wasp (Hemadas nubilipennis) is a small insect pest native to eastern North America that infests highbush and lowbush blueberry cultivation. This infestation significantly impacts the yield and quality of both wild and predominantly commercial crops. Although multiple studies have been published regarding managing and characterizing of this species and the gall, the exact mechanism for the gall formation has yet to be discovered. The first genome size estimate for this species (~547.91 Mb) was established from the short-read data using GenomeScope. The quality of our assembly was underscored by a BUSCO completeness score of 95% and average coverage of 96.49x, with sequence distributed across 20,813 contigs with an N50 of 40.8 Kbps. This study is the first look at the efforts towards the complete genome assembly and annotation of this impactful agricultural species. The blueberry stem-gall wasp genome's high-quality assembly will provide the foundation for gaining new insight into gall formation mechanisms, co-evolution, host-specific behaviors, and genetic markers for targeted pest control management strategies.

222 - tRNA gene repertoires are dynamically constructed

Dylan Sosa

Dylan Sosa, Jianhai Chen, Shengqian Xia, Manyuan Long

University of Chicago, University of Chicago, University of Chicago, University of Chicago

The duplicative nature of the tRNA gene family has been well described, yet how its multi-copy members arise and are naturally selected in eukaryotic genomes and populations remains largely unexplored. We inferred the tempo of Drosophila melanogaster tRNA gene evolution and their mode of origination through comparative genomic analyses of 24 Drosophilid species. Throughout 60 million years of evolution, Drosophilid tRNA isodecoding and isoaccepting genes have been continuously lost and gained at a rate of 1.92 tRNA gene gains and 3.98 losses per million years. The differential temporal distribution of tRNA gene origination in Drosophilids has led to taxonomically restricted copies and we report that 1.03% of D. melanogaster tRNA genes have originated within the last 10 million years resulting in one copy being specific to D. melanogaster. We used site-frequency spectra tests of neutrality to probe how nuclear tRNA gene repertoires are constructed using 727 natural D. melanogaster population samples. We observed departures from neutral expectations consistent with purifying selection in patterns of polymorphisms at loci of 23% of D. melanogaster tRNA genes. To assess the effect of gene age on tRNA function we utilized a novel small RNA sequencing methodology (MSR-seq) to measure expression, charging, and base modifications of the D. melanogaster and D. pseudoobscura tRNA gene sets in gonad and head tissues. Our data suggest that the evolutionary process of eukaryotic tRNA gene complements is characterized by continuous changes in membership rather than simply retention of ancient copies.

266 - Comparative study of population TSS maps in yeast unravels genetic basis underlying

divergence of transcription initiation

Zhenguo Lin

Zhenguo Lin, Zhenbin Hu, Junhao Chen, Yu Zhan, William Park

Saint Louis University, Saint Louis University, Saint Louis University, Saint Louis University, Saint Louis University

Genetically mediated changes in gene expression play a significant role in phenotypic diversification and adaptation. As the first step of gene expression, transcription initiation integrates regulatory signals, determining transcription start sites (TSS) and transcript abundance. TSS turnovers during evolution often change gene expression and gene structure, and dysregulation of TSS usage is associated with many human diseases. Yet, its genetic basis is poorly understood. The great genetic diversity and rich population genomic data of the budding yeast Saccharomyces cerevisiae offer an ideal system for such studies. Here, we generated and characterized quantitative TSS maps for 24 S. cerevisiae strains by CAGE sequencing. Comparative studies of these TSS maps reveal a highly dynamic transcription initiation landscape among yeast populations, outpacing their sequence substitution rates. Integrative interrogations of these TSS maps and SNP data showed that transversion substitutions near a TSS (i.e., -8, -1, and +1 sites) profoundly changed TSS usage within a core promoter, but had limited impact on its overall expression level. Major changes in TSS locations and activities are linked to mutations in upstream binding sites of gene-specific transcription factors and trans-acting factors. Comparative studies of allele-specific TSS usage in an artificial hybrid strain with its parental strains support the major role of trans-acting factors in TSS turnovers. Our study provides a detailed picture depicting the evolutionary dynamic of transcription initiation landscape and its genetic basis, contributing to a better understanding of the genetic mechanisms underlying the diversification of gene expression.

281 - Significance of aneuploidy in understanding population structure of hexaploid coast redwood

Alexandra Sasha Nikolaeva

Alexandra Sasha Nikolaeva, Rasmus Nielsen, James Santangelo

University of California, Berkeley, University of California, Berkeley, University of California, Berkeley

Simulation studies have indicated that variations in ploidy levels can lead to erroneous clustering on a PCA within the same species (Meirmans, 2018), suggesting population structure that does not exist. However, such clustering might also obscure the existing population structure within polyploid species with the same ploidy level. In coast redwood (Sequoia sempervirens), an autohexaploid species, only minimal population structure has been detected in previous studies, in stark contrast to its diploid counterpart, the giant sequoia (Sequoiadendron giganteum). Our research suggests that the population structure of the coast redwood, previously overlooked, might be concealed by chromosomal number variations (aneuploidy). Revealing this hidden structure is crucial for understanding the coast redwood's adaptive responses and informing conservation strategies for the species.

295 - Alternative splicing and environmental adaptation in wild house mice

David N. Manahan

David N. Manahan, Michael W. Nachman

University of California, Berkeley, Museum of Vertebrate Zoology, University of California, Berkeley, Museum of Vertebrate Zoology

A major goal of evolutionary genetics is to understand the genetic and molecular mechanisms underlying adaptation. Previous work has established that changes in gene regulation may contribute to adaptive evolution, but most studies have focused on mRNA abundance and only a few studies have investigated the role of post-transcriptional processing. Here, we use a combination of exome sequences and short-read RNA-Seq data from wild house mice (Mus musculus domesticus) collected along a latitudinal transect in eastern North

America to identify candidate genes for local adaptation through alternative splicing. First, we identified alternatively spliced transcripts that differ in frequency between mice from the northern-most and southern-most populations in this transect. We then identified the subset of these transcripts that exhibit clinal patterns of variation among all populations in the transect. Finally, we conducted association studies to identify cis-acting splicing quantitative trait loci (cis-sQTL), and we identified cis-sQTL that overlapped with previously ascertained targets of selection from genome scans. Together, these analyses identified a small set of alternatively spliced transcripts that may underlie environmental adaptation in house mice. Many of these genes have known phenotypes associated with body size, a trait that varies clinally in these populations. We observed no overlap between these genes and genes previously identified by changes in mRNA abundance, indicating that alternative splicing and changes in mRNA abundance may provide separate molecular mechanisms of adaptation.

298 - Using nematodes to study natural variation in mutagen sensitivity

Sophia Tintori

Sophia Tintori, Matthew Rockman
New York University, New York University

All living cells are continuously challenged by exogenous, endogenous, and spontaneous DNA damage. While conserved DNA damage response keeps organisms safe broadly, natural variants in these mechanisms can carry profound consequences for cell biology, health, and evolutionary genomics. In this study we use wild nematode strains to study the DNA repair and mutagen tolerance variants that exist in natural populations, as well as the cell biological and mutagenic consequences of those variants.

We have developed a highly sensitive assay to identify which wild strains are naturally more or less tolerant to different DNA damaging agents during chronic exposure over multiple generations. We have applied this assay to 35 strains across two species: C. elegans strains that were previously collected from around the world, and O. tipulae that we collected from the Chornobyl Exclusion Zone, a radioactive landscape that we hypothesized may have already selected for mutagen tolerance. We have tested these strains for their diverging sensitivities to a range of DNA damaging chemicals, and have identified phenotypic outlier strains. We have performed bulk segregant analyses with hybrid populations of these outlier strains, identifying alleles of interest for mutagen tolerance phenotypes.

393 - Relating FST to Information Theory gives multiple insights.

Daniel R Tabin

Daniel R Tabin, Nick J Patterson, David E Reich
Harvard, Harvard, Broad Institute, Harvard

Wright's FST can be thought of in Information Theory terms. Specifically FST can be seen to be related to the ratio of Shannon entropy without subpopulation knowledge and the Shannon entropy with subpopulation knowledge, where the "subpopulation" can be any subdivision in the data meaningful or otherwise. Thinking of FST in this way offers a number of benefits. First and foremost, this makes the meaning of FST, as Wright himself describes it, much clearer than the traditional description of differentiation, which is easily confused with the related but non identical concept of difference. Likewise, we view thinking of FST as the ratio of information with and without subpopulation knowledge as more intuitive to students and the general public when compared with thinking of FST as variance. Many seemingly paradoxical facts about FST, such as the fact that Han and French have FSTs similar to those between Amazonian groups, and that African-Eurasian FST went down after Eurasians mixed with Neanderthals, become intuitive when thinking of FST in terms of information. Additionally, many insights can be gleaned from this view. For example, we show how Information Theory can be

used to compare FST with other types of relatedness, such as genealogical relatedness, and we make an argument that the formula for FST should depend on the ploidy of the species, with haploid and tetraploid organisms requiring a different formula than the traditional diploid one.

396 - The role of promiscuous molecular recognition in the evolution of RNase-based self-incompatibility

Tamar Friedlander

Keren Erez, Amit Jangid, Ohad Noy Feldheim, Tamar Friedlander
The Hebrew University of Jerusalem, The Hebrew University of Jerusalem, The Hebrew University of Jerusalem, The Hebrew University of Jerusalem

How do biological networks evolve and expand? We study these questions in the context of the plant collaborative-non-self recognition self-incompatibility system. Self-incompatibility evolved to avoid self-fertilization among hermaphroditic plants. It relies on specific molecular recognition between highly diverse proteins of two families: female and male determinants, such that the combination of alleles an individual possesses determines its mating partners. Though highly polymorphic, previous models struggled to pinpoint the evolutionary trajectories by which new specificities evolved. Here, we construct a novel theoretical framework, that crucially affords interaction promiscuity and multiple distinct partners per protein, empirical findings disregarded by previous models. We demonstrate spontaneous self-organization of the population into distinct 'classes' with full between-class compatibility and a dynamic long-term balance between class emergence and decay.
Our work highlights the importance of molecular recognition promiscuity to network evolvability. Promiscuity was found in additional systems suggesting that our framework could be more broadly applicable.

413 - Micro-C reveals 3D genome conservation patterns in rice species

Amina Kurbidaeva

Amina Kurbidaeva, Sonal Gupta, Michael Purugganan
NYU, NYU, NYU

Chromatin structure plays a central role in evolution of genomic architecture. It plays an important role in speciation and domestication. Thus, it is of fundamental importance to determine the dynamics of 3D chromatin structure during evolution. The structure of chromosomes at different scales has been an object of intense study in many organisms. Yet, there hasnâ€™t been a study which compares and quantifies the overall chromosome structure of entire genomes, and correlates that with the evolutionary times since speciation events. Here, we performed a comprehensive analysis of 3D chromatin structure in five rice genomes, including three different species and three domesticated rice varieties. We focus on short evolutionary distances, spanning not more than 2.41 MYA. We show that even at such small evolutionary scale, there are clear differences in overall chromatin structure, and we can quantify those differences using CHESS. We show that these differences positively correlate with evolutionary times and with sequence similarity between the genomes. We also identify the genetic and epigenetic determinants of 3D genome similarities and differences. At high resolution, TADs are not evolutionary stable, and we identified signatures of conserved TADs and boundaries. We conclude that the overall 3D chromatin structure can be estimated and quantified with CHESS, and it positively correlates with evolutionary distances between genomes, and the individual TAD conservation between genomes depends on their genetic and epigenetic composition.

421 - Genetic and Epigenetic Data Reveal Complementary Climate Adaptation Pathways in an

Endangered Sea Turtle

James D Gilbert

James D Gilbert, Eugenie C Yen, Samuel J Shrimpton, Stuart RB Negus, James O Bazely, Alice Balard, Kirsten Fairweather, Albert Taxonera, Christophe Eizaguirre

Queen Mary University of London, Queen Mary University of London, Queen Mary University of London, Queen Mary University of London, Queen Mary University of London, Queen Mary University of London, Project Biodiversity, Project Biodiversity, Queen Mary University of London

Genomic diversity underpins a populationâ€™s adaptive potential and resilience to environmental change. While variation in genetic sequence is often used to predict population persistence, genomes have other levels of diversity. DNA methylation is an epigenetic modification that regulates gene activity and can be both genetically controlled and environmentally induced. How genetic and epigenetic diversities determine a populationâ€™s ability to respond to climate change, particularly for wild, non-model organisms remains elusive. Here, we combined nanopore and bisulfite whole genome sequencing applied to endangered loggerhead sea turtles (Caretta caretta) nesting on different islands of the Cabo Verde archipelago, along a temperature cline. Sea turtles must acquire substantial energy reserves between breeding seasons, making them sensitive to temperature changes that affect energy storage and their ability to return to their natal place to reproduce. While we detected a large number of highly differentiated genetic regions among turtles from different islands, none overlapped with regions of high variation in DNA methylation. Many regions were also detected with low genetic differentiation but differential methylation. We also find island-specific methylation of genes related to fatty acid metabolism, suggesting variation in energy storage linked to migration along the temperature cline. These insights highlight how epigenetic mechanisms, operating independently of genetic diversity, may underpin resilience to climate change.

425 - The dynamics of accumulation of gametic segregation distorters in the male germline of Arabidopsis

Christopher Condon

Christopher Condon, Camille Roux, Russ Corbett-Detig, Vincent Castric

UC Santa Cruz, Universite de Lille, UC Santa Cruz, Universite de Lille

Dobzhansky-Muller incompatibilities are one proposed contributor to speciation, where incompatible alleles result in segregation distortion and thereby impact population-wide allele frequencies. In this study we detect and quantify segregation distortion loci in crosses produced from a diverse sampling of Arabidopsis lyrata and A. halleri populations. We show that both the frequency of occurrence and the effect size of distortion loci increase as the parentsâ€™ genetic distance from one another increases. We also show that distorter loci occur not only within interspecific hybrids, but also in intraspecific hybrids produced from isolated population crosses. Finally, we identify correlated distortion effects between loci that repeatedly appear on two different chromosomes in multiple F1 individuals. Our study demonstrates that pollen-acting segregation distortion may be ubiquitous, and shows how segregation distortion contributes not only to the ongoing reproductive isolation between A. halleri and A. lyrata, but also between very recently diverged populations of the same species. More generally, this work demonstrates that the rapid evolution of haploid-affecting incompatibilities are likely to disproportionately contribute to suppressing gene flow between diverging populations.

481 - Maize and wild relatives show distinct patterns of genome downsizing following polyploidy

Samantha J Snodgrass

Samantha J Snodgrass, Margaret R. Woodhouse, Arun S. Seetharam, Michelle C. Stitzer, Matthew B. Hufford
UC Davis, USDA-ARS Corn Insects and Crop Genetics Research Unit, Department of Ecology Evolution and Organismal Biology, Iowa State University, Department of Genetics, Development, and Cellular Biology, Iowa State University, Institute for Genomic Diversity, Cornell University, Department of Ecology, Evolution, and Organismal Biology, Iowa State University

Polyploidy and fractionationâ€”the processes by which genomes grow and shrink in sizeâ€”characterize plant genome evolution. Work on single ancient polyploid species and species descended from separate whole genome duplication (WGD) events has shown fractionation happens quickly and likely before subsequent speciation events. New, chromosome level assemblies for most members of the Tripsacinae subtribe represent multiple descendant lineages from the same WGD event ~5-12 MYA. The Tripsacinae includes the wild genera Tripsacum, Zea, and the culturally and economically important domesticate maize, and spans diverse geographical and ecological ranges centered in Central and North America. We describe shared and segregating fractionation at the exon level across 35 Tripsacinae genomes. Tripsacum (haploid chromosome number n=18) shows less fractionation than any Zea species (haploid chromosome number n=10), with more exons retained in pairs. Using parsimony to estimate the relative timing of fractionation, most exons from either ancestral subgenome (~80%) are either completely fractionated or completely retained across all genomes, suggesting the majority of fractionation occurred soon after the inciting WGD event in the ancestral lineage. However, ~5-6% of exons fractionated after divergence of the genera, but before later speciation events. The higher amount of fractionation observed in Zea may be connected to extensive chromosomal rearrangements observed in this genus while Tripsacum has remained relatively stable. While most fractionation patterns of exon pairs are shared across all genomes (~65.8%), thousands of exon pairs segregate for fractionation pattern. These results suggest that fractionation is a much more dynamic and ongoing process than previously hypothesized.

491 - Hybrid adaptation is hampered by Haldane’s sieve

Carla Bautista

Carla Bautista, Isabelle Gagnon-Arsenault, Mariia Utrobina, Anna Fijarczyk, Devin P Bendixsen, Rike Stelkens, Christian R Landry
Institut de Biologie IntÃ©grative et des SystÃ¨mes (IBIS), UniversitÃ© Laval, DÃ©partement de Biologie, FacultÃ© des Sciences et de GÃ©nie, UniversitÃ© Laval, Regroupement quÃ©bÃ©cois de recherche sur la fonction, la structure et l'ingÃ©nierie des 11 protÃ©ines (PROTEO), UniversitÃ© Laval, Centre de Recherche en DonnÃ©es Massives (CRDM), UniversitÃ© Laval, Institut de Biologie IntÃ©grative et des SystÃ¨mes (IBIS), UniversitÃ© Laval, DÃ©partement de Biologie, FacultÃ© des Sciences et de GÃ©nie, UniversitÃ© Laval, Regroupement quÃ©bÃ©cois de recherche sur la fonction, la structure et l'ingÃ©nierie des 11 protÃ©ines (PROTEO), UniversitÃ© Laval, Centre de Recherche en DonnÃ©es Massives (CRDM), UniversitÃ© Laval, DÃ©partement de Biochimie, de Microbiologie et de Bio-informatique, FacultÃ© des 15 Sciences et de GÃ©nie, UniversitÃ© Laval, Institut de Biologie IntÃ©grative et des SystÃ¨mes (IBIS), UniversitÃ© Laval, DÃ©partement de Biologie, FacultÃ© des Sciences et de GÃ©nie, UniversitÃ© Laval, National University of Kyiv-Mohyla Academy, Institut de Biologie IntÃ©grative et des SystÃ¨mes (IBIS), UniversitÃ© Laval, DÃ©partement de Biologie, FacultÃ© des Sciences et de GÃ©nie, UniversitÃ© Laval, Regroupement quÃ©bÃ©cois de recherche sur la fonction, la structure et l'ingÃ©nierie des 11 protÃ©ines (PROTEO), UniversitÃ© Laval, Centre de Recherche en DonnÃ©es Massives (CRDM), UniversitÃ© Laval, Department of Zoology, Stockholm University, Department of Zoology, Stockholm University, Institut de Biologie IntÃ©grative et des SystÃ¨mes (IBIS), UniversitÃ© Laval, DÃ©partement de Biologie, FacultÃ© des Sciences et de GÃ©nie, UniversitÃ© Laval, Regroupement quÃ©bÃ©cois de recherche sur la fonction, la structure et l'ingÃ©nierie des 11 protÃ©ines (PROTEO), UniversitÃ© Laval, Centre de Recherche en DonnÃ©es Massives (CRDM), UniversitÃ© Laval, DÃ©partement de Biochimie, de Microbiologie et de Bio-informatique, FacultÃ© des 15 Sciences et de GÃ©nie, UniversitÃ© Laval

Hybrids between species exhibit plastic genomic architectures that foster phenotypic diversity. Their genomic instability also incurs costs,

potentially limiting adaptation. When challenged to evolve in an environment containing a UV mimetic drug, yeast hybrids have reduced adaptation rates compared to parents. We hypothesized that this reduction could result from a faster accumulation of genomic changes, but we found no such association. Alternatively, we proposed that hybrids might lack access to adaptive mutations occurring in the parents, yet, we identified mutations in the same genes (PDR1 and YRR1), suggesting similar molecular adaptation mechanisms. However, mutations in these genes tended to be homozygous in the parents but heterozygous in the hybrids. We hypothesized that a lower rate of loss of heterozygosity (LOH) in hybrids could limit fitness gain. Using genome editing, we demonstrated that mutations display incomplete dominance, requiring homozygosity to show full impact and to circumvent Haldaneâ€™s sieve, which favors the fixation of dominant mutations. We used frozen â€˜fossilsâ€™ to track genotype frequency dynamics and confirmed that LOH occurs at a slower pace in hybrids than in parents. Together, these findings show that Haldaneâ€™s sieve slows down adaptation in hybrids, revealing an intrinsic constraint of hybrid genomic architecture that can limit the role of hybridization in adaptive evolution.

492 - Unveiling the Diversity and Evolutionary Significance of Tandem Repeats

Luis Gerardo Fernandez Luna

Luis G Fernandez Luna, Rori Rohlfs, Carolina De Lima Adam, SebastiÃ¡n Iturbe, Diego Ortega-Del Vecchyo
University of Oregon, University of Oregon, University of Oregon, International Laboratory for Human Genome Research, National Autonomous University of Mexico, International Laboratory for Human Genome Research

Repetitive sequences, known as tandem repeats (TRs), are highly polymorphic. Over 1 million identified loci constitute more than 3% of genomic DNA in humans. TRs exhibit significant diversity due to their complex underlying mutational processes, characterized by high mutation rates that depend on changes in repeat length, motif structure, and sequence. These fast-evolving regions serve as a significant source of genetic variability that may underlie adaptation processes within and between species.
These regions have largely remained unexplored and ungenotyped due to their complexity. However, recent advances in long-read sequencing technologies and computational genotyping tools facilitate a more systematic study of TR variation. Now is feasible to address the roles of TRs in regulatory regions, their involvement in health and disease, and their contributions to complex traits such as height. Despite these advances, TRs remain an under-studied source of adaptive genetic variation due to the lack of a statistical framework supporting the comparison of TRs to test evolutionary hypotheses.
To comprehend the evolutionary processes underlying the selection of specific TRs that play pivotal roles in the dynamic nature of the genome, we propose the development of a tool named TR evolutionary variance analysis (TREVA). TREVA aims to assess the variance of TRs within and between species, providing a framework to test evolutionary hypotheses regarding TR variation. To validate and draw meaningful inferences from TREVA results, we are simulating comparative TR data under several mutational and evolutionary models. These data will provide a test for TREVA and other similar methods.

501 - The distribution of fitness effects of mutations in enhancers, promoters, and conserved non-coding regions

Chenlu Di

Chenlu Di, Kirk Lohmueller
University of California, Los Angeles, University of California, Los Angeles

Functional non-coding regions are increasingly identified in the human genome by various methods. The challenge now is to understand the fitness effects of mutations in these non-coding regions. We developed a pipeline to infer the distribution of fitness effects (DFE) of non-coding mutations. Using polymorphism data from the 1KGP, we inferred DFEs of new mutations in enhancers, promoters, and

conserved non-coding regions. We found that mutations in enhancers are frequently neutral, with a small proportion of weakly and strongly deleterious mutations. Promoters, in contrast to enhancers, show a lower proportion of neutral mutations and more weakly deleterious mutations. To investigate the temporal nature of the evolution of noncoding sequences and whether selection pressure changed over time, we estimated the fitness effects of mutations in regions of the genome showing different degrees of sequence constraint in mammals. We found the highest proportion of deleterious mutations in regions of the genome that are most conserved in both mammals and primates alone. However, regions of the genome conserved in primates alone showed a higher proportion of deleterious mutations than mutations in regions of the genome that were only conserved when considering all mammals. This pattern appears in putative enhancers, promoters, and when classifying conserved regions at two thresholds. Our results highlight the dynamic nature of the evolution of gene regulation, with changing selection coefficients over deep evolutionary time. More broadly, our work has implications for using comparative genomic approaches to detect non-neutrally evolving sequences in the human genome.

632 - The genetic basis of long-distance migration in Common swifts (Apus apus)

Heiman Ho

Heiman Ho, Leyi Su, Zhou Wu, Chenqing Zheng, Yang Liu

School of Ecology, Sun Yat-sen University, IMBIM, Uppsala University, The Roslin Institute and Royal (Dick) School of Veterinary Studies R(D)SVS, University of Edinburgh, School of Ecology, Sun Yat-sen University, School of Ecology, Sun Yat-sen University

The genetic basis of long-distance migration has long been one of the interests in evolutionary biology. While integration of genetic, behavioural and physiological factors has complicated the decoding of general signs in migration adaptation, research on different systems is also essential to achieve this goal. Here, we applied WGS on 47 swifts, including 13 Apus pallidus, which migrate from southern Europe to Central Africa (~4500km), along with 17 of its sister species â€“ A. apus apus â€“ which have relatively longer migratory distance from Europe to South East Africa (~7000km), as well as 17 A. a. pekinensis moving between Asia and South Africa (~13000km). 11,405,577 SNPs were yield with at an average sequencing depth of ~20.5X. We adopted sliding windows approach with 50kb windows and estimated FST, dXY and XP-nSL to find the top 5% diverged regions which likely to be selected when comparing A. pallidus to A. a. apus and A. a. pekinensis separately. Correspondingly, we found a total of 128 and 75 non-overlapping genes, suggesting the 2 subspecies having different evolutionary routes. Candidate genes relating to mitochondrial functions such as IMMP2L and RYR2 proved the importance of mitochondria during long-distance migration. In addition, we found an overlapping gene â€“ Interleukin 1 Receptor Accessory Protein Like 1(IL1RAPL1) â€“ in both comparisons with signs of selection, which is likely to play a role in cerebellar development, cognitive functions and memory abilities. We suggest that these genes might be important in maintaining long-distance migratory behaviour.

636 - Investigating BUSCO databases and performance in the light of a new generation of chromosome-level molluscan genomes

Marcela Uliano da Silva

Marcela Uliano-Silva, Mark Blaxter

Wellcome Sanger Institute, Wellcome Sanger Institute

The Benchmarking of universal single copy orthologs (BUSCO) is a widely used tool for evaluation of genome assembly quality and completeness. Mollusca is a phylum known to recover low BUSCO scores, and a comprehensive analysis of the possible causes is lacking. We have run BUSCO v5.6.1 with metauek and mollusca_db10 for 60 highly contiguous chromosome-level genomes produced by the Darwin Tree of Life Project and others, plus the more fragmented genomes of the 7 species included in mollusca_db10. Our analyses

included genomes from main molluscan groups (Scaphopoda, Gastropoda, Cephalophoda, Bivalvia). Out of the 5295 mollusca_db10 BUSCOs, only 41 were found as Complete in all species, and no common BUSCO genes were found to be Missing, Fragmented or Duplicated in all species analysed. A presence/absence analysis of Complete BUSCOs showed clearly that completeness BUSCO recovery has a strong bias towards the species included in mollusca_db10 and its close-relatives, and recovery decreases as further away species are in the phylogeny. We are further investigating if the mostly common missed genes are (i) difficult to predict by homology, (ii) if they are affected by the predictor used (metaeuk/augustus) and the mode run (genome/proteome), or (ii) if these missed genes may not represent universally conserved elements across all molluscan taxa. This investigation aims to better inform our use and understanding of genome quality assessment tools in the light of new chromosome-level genomes for multiple never-before-sequenced species across the main phylogenies.

682 - If they don’t have proteins, let them have RNA: evolutionary rescue of defective ribosome mutants by changes to transcription

Enea Franceschini

Enea Franceschini, Paul B Rainey

Department of Microbial Population Biology, Max Planck Institute for Evolutionary Biology, Department of Microbial Population Biology, Max Planck Institute for Evolutionary Biology, Laboratory of Biophysics and Evolution, CBI, ESPCI Paris, UniversitÃ© PSL, CNRS

The ribosome is one of the more conserved molecular machines present in all kingdoms of life. In the last couple of decades much progress has been made to describe it and determine its function. Despite this, there is still a lack of understanding of both the specific role performed by some of its protein components and of factors involved in the loss and acquisition of said components over evolutionary time.

In our experiments we used an experimental evolution framework to test if and how bacteria can adapt to the loss of ribosomal proteins. Five different ribosomal proteins of varying ancestry were deleted from the chromosome of the model bacterium Pseudomonas fluorescens. As expected, these mutations caused a decrease in growth rate and translation speed.

After more than 500 generations each evolving population showed partial or full recovery of fitness. Genome sequencing of derived populations allowed compensatory mutations to be identified, many of which targeted genes that couple transcription and translation.

Further experiments showed stronger loss of fitness caused by the deletions we produced when there is lack of phosphate in the growth medium. Since phosphate is needed for nucleotide synthesis, this could point to a higher need of RNA transcription to maintain fitness.

These results together suggest that ribosomal protein gene deletions not only affect translation speed, but also alter the required transcription rate needed to maintain growth. The need to optimize the balance between transcription and translation in a given environment could affect the evolution of the translation apparatus.

730 - DNA replication initiators have rapidly evolving chromatin-binding elements for adaptive specification of replication start sites

Tasia Bos

Tasia Bos, Matthew W Parker

The University of Texas Southwestern Medical Center, The University of Texas Southwestern Medical Center

The first step of eukaryotic DNA replication is the so-called â€œlicensingâ€• step where the Mcm2-7 helicase is loaded onto chromosomal loci known as origins. Three proteins are required for licensing, including the origin recognition complex (ORC), Cdt1, and Cdc6. These licensing factors are conserved across eukaryotes. ORC is composed of six subunits, Orc1-Orc6, and is responsible for

recognizing and binding the origin sites where Mcm2-7 will be loaded. The Orc1 subunit is primarily responsible for origin selection in metazoans and, consistently, possesses two chromatin-tethering appendages in addition to its core AAA+ domain: an intrinsically disordered region (IDR) that binds DNA non-specifically and a bromo-adjacent homology (BAH) domain that binds a specific epigenetic mark. To determine whether origins of replication are conserved, we have assessed the molecular evolution of ORC and its chromatin-binding elements across the metazoan lineage. Despite its central role in origin binding, we find that Orc1 is the fastest evolving ORC subunit. In fact, the chromatin binding elements themselves (i.e., IDR and BAH) are the most rapidly evolving domains with alterations that suggest dramatic changes in origin specificity. Ongoing work is investigating how these changes alter the specificity of DNA and chromatin binding in vitro and in a cellular setting. Altogether, these studies demonstrate that ORCâ€™s essential chromatin binding elements are rapidly evolving. Our future work will investigate whether the observed plasticity is essential to adapt this fundamental cellular process to the varied cellular and environmental conditions of metazoan organisms.

774 - Unraveling the evolutionary history of polyploid Arabidopsis kamchatica

Maria Vasilarouo

Maria Vasilarou, Uliana Kolesnikova, Alison D Scott, Polina Novikova

Max Plank for Plant Breeding research, Max Plank for Plant Breeding research, Max Plank for Plant Breeding research, Max Plank for Plant Breeding research

Polyploidy is an enduring force in plant evolution and speciation. The study of allopolyploid species presents a fascinating avenue for exploring the dynamics of polyploid evolution. In allotetraploids, the fusion of two divergent genomes in a single nucleus can lead to dynamic interactions and genomic changes. For example, homoeologous exchange between the two subgenomes may occur, though these are more often detected in synthetic or very young polyploids. Established allopolyploids may instead demonstrate genomic stability and absence of large homeologous exchanges. Arabidopsis kamchatica is an established allotetraploid with a complex evolutionary history, involving multiple hybridization events between A. lyrata and A. halleri. Using de novo genome assemblies and population level resequencing data, we explore the evolutionary trajectory of the two subgenomes within A. kamchatica. In particular, we investigate how the subgenomes interact since their fusion into the allopolyploid state. To that end, we generated a reference genome that allowed us to detect evidence for large genomic rearrangements between the two subgenomes of our A. kamchatica accession. Ultimately, the study of A. kamchatica will enhance our comprehension of polyploid genome evolution while also underscoring the necessity of a population-level analysis for resolving the complex evolutionary origins of allopolyploidization.

788 - Population history and chromosomal speciation in gelada monkeys

Brooklynn R. Scott

Brooklynn R. Scott, Ferehiwot Ayele, Alemayehu Lemma, Jacinta C. Beehner, India A. Schneider-Crease, Amy Lu, Thore J. Bergman, Kenneth L. Chiou, Noah Snyder-Mackler

Arizona State University , Addis Ababa University, Addis Ababa University, University of Michigan, Arizona State University , Stony Brook University , University of Michigan, Arizona State University , Arizona State University

Genomic studies of the demographic history of extant populations have revealed a myriad of processes underlying genetic diversity, barriers to hybridization, and speciation. Geladas (Theropithecus gelada) are an especially interesting example of a long-lived primate that may be in the midst of an ongoing speciation event. We recently identified a novel, derived karyotype (2n=44) in one population (â€œnorthern geladasâ€•) that resulted from a fission on chromosome 7. This karyotype evolved recently in geladas, differing from the ancestral 2n=42 karyotype that other populations (â€œcentral geladasâ€•) share with geladasâ€™ close relatives, including baboons

(Papio spp.) and macaques (Macaca spp.). Here, we used population resequencing data from northern and central geladas to reconstruct their population history and identify when and how this fission event occurred. Using two complementary approaches, we estimate that the two populations diverged ~175 kya, with substantial changes in population size after the split. This suggests that the novel karyotype emerged in the northern population around the time of divergence. The limited migration estimated between the populations suggest that the fission represents a strong barrier to hybridization. The genetic divergence and karyotype difference between populations suggest emerging reproductive barriers, which carry substantial implications for gelada speciation and conservation efforts.

796 - Antarctic microbial genomics: insights into cold adaptation and secondary metabolite biosynthesis

Michele Giovannini

Michele Giovannini, Walter Vieri, Angelina Lo Giudice, Renato Fani, Emanuele Bosi, Marco Fondi, Elena Perrin
University of Florence, University of Florence, Institute of Polar Sciences, National Research Council (CNR-ISP), Messina, University of Florence, University of Genova, University of Florence, University of Florence

Antarctica, one of the most extreme environments on Earth, hosts diverse microbial communities. These microbes have evolved and adapted to survive in these hostile conditions, but knowledge on the molecular mechanisms underlying this process remains limited. The Italian Collection of Antarctic Bacteria [Collezione Italiana Batteri Antartici (CIBAN)], managed by the University of Messina, represents a valuable repository of cold-adapted bacterial strains isolated from various Antarctic environments. In this study, we analysed the genomes of 62 Î³-Proteobacteria strains from the CIBAN collection, isolated during Italian expeditions from 1990 to 2005. Genomic DNA was extracted, sequenced using Illumina technology, and de novo assembled. By employing average nucleotide identity (ANI) analysis, we confirmed the genomic relatedness between strains. This enabled us to discern four distinct clusters belonging to different genera: Pseudomonas, Pseudoalteromonas, Shewanella and Psychrobacter. Genome annotation and comparative analyses revealed insights into the functional potential of these strains, including the presence of secondary metabolite biosynthetic gene clusters and antibiotic resistance genes. In addition, phylogenomic analyses provided a framework for understanding the evolutionary context. Whereas assessment of cold-shock protein presence shed light on adaptation mechanisms. Therefore, our findings underscore the importance of the culture collection as a resource for understanding Antarctic microbial life and its biotechnological potential. The genomic data open new horizons for understanding bacterial life in Antarctica and offer practical solutions for preserving this delicate ecosystem.

This research was funded by the Italian National Program of research in Antarctica â€œProgramma Nazionale di Ricerca in Antartide (PNRA)â€•, grant number PNRA18_00075.

859 - Characterisation and quantification of deleterious genetic variants in non-model organisms: from present to extinct species.

Julia Höglund

Julia HÃ¶glund, Seyan Hu, Martijn F.L. Derks, Love DalÃ©n, Mirte Bosse
Centre for Palaeogenetics, Dept. of Zoology, Stockholm University, Dept. of Animal Breeding and Genomics, Wageningen University & Research, Dept. of Animal Breeding and Genomics, Wageningen University & Research, Dept. of Animal Breeding and Genomics, Wageningen University & Research, Centre for Palaeogenetics, Dept. of Zoology, Stockholm University, Dept. of Animal Breeding and Genomics, Wageningen University & Research, Section Ecology & Evolution, Amsterdam Institute for Life and Environment (A-LIFE), Vrije Universiteit Amsterdam

Many animal populations are currently undergoing dramatic decline, to which human activities have significantly contributed. When a population becomes smaller, threats affecting the speciesâ€™ genome become larger. A reduced population size increases the risk of inbreeding and loss of beneficial genetic variation, results in a worse ability to adapt to changes, and can potentially lead to extinction. Hence, there is a critical need to quantify damaging genetic variation to aid conservation efforts and understand past extinctions. One approach to characterise damaging genetic variants is to score them based on predicted deleteriousness. However, such models are currently species-specific and often unreliable beyond model organisms. To address this, we utilised the domesticated pig reference genome and estimated substitution and mutation rates from several ancestral nodes to compare the impact of evolutionary time on scoring and annotation. By using dense sequencing data, we will train a scoring model within domesticated pig, test it in wild boar and lastly extending it to endangered pig species. Preliminary results suggest similar distributions of variant effects across ancestral nodes. In contrast to previous similar scoring models, this model extends across similar species already in the initial score estimation. These scores will then be utilised in quantifying genetic load, to estimate how much genetic variation is contributing to animal extinction. By extending the model beyond model species, we enhance its applicability to several species, hopefully providing insights into the possible genetic factors underlying species decline.

940 - Genome-wide analysis of transcription initiation in Paramecium

Berenice Jimenez-Marin

Berenice Jimenez-Marin, Jean-Francois Gout, Weibo Zheng, Taylor Raborn, Tia Swenty, Timothy Licknack, Michael Lynch
Arizona State University, Arizona State University, Arizona State University, Arizona State University, Arizona State University, Arizona State University, Arizona State University, Arizona State University

The timing and intensity of transcription of different genes is fundamental to survival of the organism, and thus is orchestrated by sequences located upstream of the protein-coding portion of the gene itself. In the single-celled eukaryotes of genus Paramecium, the macronuclear genome is remarkably compact and streamlined for gene expression. Interestingly, member species of the Paramecium aurelia species complex have undergone two whole genome duplications, resulting in differential patterns of gene retention and loss which are influenced by their expression levels. As means to explore eukaryotic transcription in a streamlined genome and shed light on the role of sequence architecture on gene expression and loss, we analyzed the distribution and diversity of candidate transcription initiation sites in Paramecium sexaurelia and Paramecium tetraurelia. A total of 10,652 genes (30.5% of the P. sexaurelia genome) and 22,098 genes (55.9% pf the P. tetraurelia genome) were queried for their transcription initiation site (TIS) diversity and conservation within and between species. Our analysis suggests that both Paramecium aurelia genes preferentially have narrow transcriptional initiation regions, and >90% of genes in both species have considerably short 5â€™ UTRs (<55 bp). The short region in which transcription initiation occurs most frequently and the short length of the Paramecium 5â€™ UTRs suggest differences in promoter regions between Paramecium species, in addition to divergence in 5â€™UTR architecture under relatively strict sequence-length constraints, influence gene retention and loss.

941 - Emergence of saliva protein genes in the secretory calcium-binding phosphoprotein (SCPP) locus and accelerated evolution in primates

Petar Pajic

Petar Pajic, Luane Landau, Omer Gokcumen, Stefan Ruhl
University at Buffalo, University at Buffalo, University at Buffalo, University at Buffalo, University at Buffalo

Genes within the secretory calcium-binding phosphoprotein (SCPP) family evolved in conjunction with major evolutionary milestones: the

formation of a calcified skeleton in vertebrates, the emergence of tooth enamel in fish, and the introduction of lactation in mammals. The SCPP gene family also contains genes expressed primarily and abundantly in human saliva. Here, we explored the evolution of the saliva-related SCPP genes by harnessing currently available genomic and transcriptomic resources. Our findings provide insights into the expansion and diversification of SCPP genes, notably identifying previously undocumented convergent gene duplications. In primate genomes, we found additional duplication and diversification events that affected genes coding for proteins secreted in saliva. These saliva-related SCPP genes exhibit signatures of positive selection in the primate lineage while the other genes in the same locus remain conserved. We found that regulatory shifts and gene turnover events facilitated the accelerated gain of salivary expression. Collectively, our results position the SCPP gene family as a hotbed of evolutionary innovation, suggesting the potential role of dietary and pathogenic pressures in the adaptive diversification of the saliva composition in primates, including humans.

950 - A Comparative Multi-Omic Analysis of Plethodontid Salamander Mucus

Paul Anthony Nicolosi

Paul A Nicolosi
The Ohio State University

Amphibian cutaneous secretions are critical for diverse and rapidly evolving physiological functions, such as territorial signaling, courtship, and immune responses. As their mucus mediates both interspecific and intraspecific interactions, its utility as a general mechanism of communication mechanism suggests potential coevolution of its components to those of other species. Despite the large body of work demonstrating the diverse functionality of amphibian cutaneous secretions, little is known about their biochemical composition. In this study, we combine proteomic and transcriptomic methods to build high quality â€œmucosomesâ€• for two lungless salamander species (family Plethodontidae), Plethodon shermani and Desmognathus ocoee. Plethodon salamanders, known as slimy salamanders, produce substantial amounts of mucus relative to their size. The quantity of mucus along with past works demonstrating its involvement in communication makes them ideal models for studying the composition of cutaneous secretions. The distantly related D. ocoee is found in the same habitats and produce similarly complex but biochemically distinct mucus. Using high performance liquid chromatography (HPLC) in tandem with mass spectronomy (MS), we characterized the protein contents of these excretions, with transcriptomic analyses of skin samples collected from matched animals used to construct search databases for proteomics. Our results demonstrate the diverse functionality of Plethodontid mucus, illuminating aspects of their biology. The use of cutaneous secretions as a system of information exchange between salamanders and their environment provides a starting point for understanding how these interactions have shaped their life histories.

951 - Exploring the molecular coevolution of species-specific fertilization in marine abalone through high-throughput mutagenesis using Pichia

Trenton John Winters

Trenton J Winters, Damien B Wilburn
The Ohio State University, The Ohio State University

Biochemical interactions in sperm-egg fusion can impose strong selection on functionally important residues between binding proteins. Sperm lysin and egg VERL in the marine mollusk abalone are a classic model of sexual coevolution whose species-specific interactions restrict hybridization of natural populations. VERL is both glycosylated and disulfide-bonded such that recombination expression requires a eukaryotic host. The methylotrophic yeast Komagataella pastoris â€“ better known as Pichia â€“ is an extremely cost-effective and efficient expression system for such proteins, where available genetically engineered strains allow for precise control over

post-translational modifications. However, evolutionary studies using high throughput mutagenesis screens have been limited in this system due to a lack of efficient cloning methods. By combining a suite of recently characterized genetic tools, we have developed a new yeast display vector to explore the genetics and biophysics of lysin-VERL coevolution. Only recently were universal yeast origins of replication discovered to enable stable plasmid maintenance in Pichia for efficient transformation of many mutants. The expression cassette includes a GPI-anchor for cell wall attachment with engineered fluorescence and protease susceptibility for multiple layers of surface control. Pichia expression allows us to display mutagenized libraries of proteins with controlled PTMs and screen binding with simple fluorescence measurements. Lysin and VERL coevolve rapidly such that many sites vary between the orthologs of any two abalone species. Our Pichia display system is a robust and accessible tool for understanding the structural and functional consequences of mutations in this complex evolutionary system.

957 - Untargeted nanopore-based genomics enable phylogenetic analysis of highly repetitive genes required for species-specific fertilization in marine abalone.

Trevor Gregory Stepanyan

Trevor G Stepanyan, Damien B Wilburn

The Ohio State University, The Ohio State University

Repetitive sequences are common throughout animal genomes but are notoriously difficult to uniquely assign to loci through short-read sequencing. This creates challenges for phylogenomic analyses between different species. Marine abalone are gastropods found along the Pacific coast of the United States where 7 species exist with overlapping habitat ranges. Natural hybrids are rare despite external fertilization with common breeding seasons and lab-bred hybrids being viable. Abalone eggs contain a glycoprotein known as VERL that comprises ~30% of their egg extracellular matrices. Interactions between egg protein VERL and sperm protein lysin are critical for species-specific fertilization in nature. VERL is a highly repetitive molecule with many tandem repeats of a lysin-binding domain which experience concerted evolution. This repetitive domain structure has created enormous experimental challenges for enriching and sequencing the locus via traditional molecular biological approaches. Here we demonstrate that shotgun whole-genome nanopore sequencing with ultra-high molecular weight genomic DNA provides an untargeted, cost-effective, and bioinformatically simple strategy for performing phylogenetic analyses of complex, difficult-to-sequence loci. Optimizations were performed for DNA extraction methods, tissue type, and library conditions to produce >10X coverage of multiple abalone genomes, including previously un-sequenced genomes, with median read lengths >60 kb. Phylogenetic comparisons after extracting full-length VERL using simple BLAST searches has enabled identifying signatures of molecular evolution in a critical fertilization protein. This method provides a roadmap for affordable characterization of formerly difficult-to-sequence genomic regions within evolutionary critical non-model systems.

1016 - Genome size variations and putative length-variable sex chromosomes within surface, subterranean and F1 hybrid populations of Asellus aquaticus isopods.

Mohammed M.Tawfeeq

Mohammed M.Tawfeeq, Meredith Protas, Serban M. Sarbu, Jean-FranÃ§ois Flot

UniversitÃ© Libre de Bruxelles, Dominican University of California, â€œEmil RacoviÅ£Äƒâ€• Institute of Speleology, California State University, UniversitÃ© Libre de Bruxelles

Genome size is a fundamental biological trait that varies a lot among eukaryotic species. However, research on its intraspecific diversity remains limited. Previous studies comparing surface and subterranean isopod species found that cave species had larger genomes than surface ones (LefÃ©bure et al. 2017), hence we wondered whether the same pattern would be observed within species. To fill this gap,

we examined intraspecific variations in genome sizes in Asellus aquaticus, a freshwater-isopod crustacean with surface and cave-dwelling populations. A total of 39 specimens were collected from surface and subsurface populations in Slovenia and Romania, as well as seven individuals from a hybrid F1 population obtained by crossing hypogean males with epigean females (both from Romania). Using a new, refined Feulgen Image Analysis Densitometry protocol, we generated precise genome size estimates with ranges 0.979-1.98pg (surface), 1.98-2.86pg (cave) and 1.75-2.91pg (F1) for Romania vs. 1.95-2.43 (surface) and 2.12-4.11 (cave) for Slovenia. Intraspecific 1.25 to 2-fold genome size variations were detected in every population. For surface populations, variation was similar among males as among females, but for cave populations the variation was 1.7-2 times larger in males than in females, suggesting strong variations in Y chromosome length in cave populations. Cave populations also had larger average genome sizes than surface ones, with the F1 population having a genome size intermediary between those of its parent populations. The observed variations raise the question of the underlying genomic features, which future whole-genome sequencing will uncover.

1053 - Evolutionary Constraints on Gene Expression in Humans: Implications for Drug Efficacy and Personalized Medicine

Alber Aqil

Alber Aqil, Omer Gokcumen
University at Buffalo, University at Buffalo

Previous studies have shown that highly expressed genes in humans tend to be evolutionarily constrained at the sequence level. However, the constraints on the regulation of genes with different magnitudes of expression remain understudied. Here, we use gene expression variability among individuals, measured by the Gini coefficient, as an inversely related proxy for constraint on gene regulation. Our study systematically quantifies the relationship between the expression magnitude (mean expression across individuals) and expression variability for 30,000 genes across 27 tissues using the Genotype-Tissue Expression (GTEx) database. We observe a consistent negative correlation (Spearmanâ€™s rho < -0.9; p < 2.2 x 10^-16 across tissues) between expression magnitude and variability. This indicates a stronger constraint on the regulatory architecture of genes with a higher magnitude of expression. Genes that are exceptions to this trend may be interesting in the context of both balancing selection and disease. In parallel, the analysis of these expression trends allowed us to test the hypothesis that the efficacy of drugs is negatively affected by the high expression variability of target genes, which hinders the determination of optimal drug dosage for individual patients. Indeed, genes targeted by FDA-approved drugs that were subsequently withdrawn exhibit greater expression variability compared to drugs still in use (p < 2.2 x 10^-16; Cohenâ€™s d=0.5). These insights will inform drug discovery studies and advance the development of personalized medicine.

1074 - Unveiling a cryptic player: understanding the role of accessory elements in fungal pathogens

Dinah M Parker

Dinah M Parker, Carolina Nogueira, Michael Habig, Tue K Nielsen, Knud N Nielsen, Andi Wilson, Lars H Hansen, Henrik H De Fine Licht
University of Copenhagen, University of Copenhagen, Kiel University, University of Copenhagen, University of Copenhagen, University of Copenhagen, University of Copenhagen, University of Copenhagen

Within Eukarya, fungi exhibit remarkable genomic plasticity, featuring abundant accessory regions integrated either within core chromosomes or as dispensable accessory chromosomes. This variability influences phenotypic characteristics and is particularly evident in pathogenic species. Utilizing seven chromosomal-level assembled and reference genomes from four continents, we establish a reference-quality pangenome to identify accessory regions present in Metarhizium acridum, a specialist fungal pathogen that infects orthopteran insects. Here, we identify a novel accessory chromosome unique to a single isolate, and show for the first time in fungal

biology, the presence of MAT loci on an accessory chromosomeâ€”a necessary element for fungal sexual reproduction. Additionally, we also find evidence of functional compartmentalization of the fungal genome, as the core genome of M. acridum is enriched in carbohydrate-active enzymes, while the accessory component is enriched in effectors that play a major role in host-pathogen interactions. Using this foundational genomic dataset, three isolates displaying divergent genotypic and virulence profiles were then selected for a serial passage experiment to investigate how standing genetic variation in a pathogen contributes or constrains the ability to colonize novel hosts. Specifically, M. acridum was passaged for several generations through Locusta migratoria grasshoppers, representing the natural host, and Tenebrio molitor beetles, representing a novel host. Clear changes in virulence and gene expression emerge, shedding light on how standing genetic variation and accessory genomic compartments influence early adaptation to a novel host. This investigation in genome plasticity in fungal pathogens has implications for understanding and managing emerging infectious diseases.

1081 - Exploring the evolutionary basis of gene expression trends in murine salivary glands

Luane Jandira Bueno Landau

Luane Jandira Bueno Landau, Achisha Saikia, Sarah Knox, Omer Gokcumen

University at Buffalo, University at Buffalo, UT Southwestern Medical Center, University of California San Francisco, University at Buffalo

The major salivary glands (parotid, submandibular and sublingual glands) show a high diversity in gene expression in mammals. More specifically, sex-bias in gene expression is present for the three glands, but variation across mouse strains was not investigated yet. Here, we use mouse as a model to understand the evolution of gene expression in the salivary glands. We generated sex- and tissue-specific transcriptome from the three major salivary glands of two mouse strains (CD1 and C57). We compared expression in each gland for strain and sex. Submandibular (SM) shows the higher number of differentially expressed genes (DEGs) across sexes. When performing a GO enrichment analysis for the sex-biased genes, we found genes highly expressed in females show an enrichment for transmembrane transport ($p<10^{-7}$). This is consistent with sex differences in cell composition in the murine SM gland. As females present an increased number of acinar cells, these patterns could be related to differences in salivary fluid secretion. Alternatively, we found that the sublingual (SL) gland shows the higher number of DEGs between strains. Highly expressed genes in CD1 strain show an enrichment for developmental processes ($p<10^{-2}$). This is consistent with reported differences in development across mouse strains for multiple tissues. Two highly expressed genes in CD1 strain (Galtn10 and St6GalNa) are related to glycosylation, which are known to have major phenotypic impacts on saliva biology. This dataset provides the necessary means for the analysis of the evolutionary bases of the salivary glands gene expression.

139 - Non-invasive sampling of Inca mummies yields undocumented pathogens and human ancestry

Michelle Hämmerle

Michelle HÃ¤mmerle, Meriam Guellil, Dagmara Socha, Susanna Sawyer, Olivia Cheronet, Mario Bernaski, Gabriela Recagno, Martin Kuhlwilm, Pere Gelabert, Ron Pinhasi

University of Vienna, Human Evolution and Archaeological Sciences (HEAS), University of Vienna, Human Evolution and Archaeological Sciences (HEAS), Center for Andean Studies, University of Warsaw, University of Vienna, Human Evolution and Archaeological Sciences (HEAS), University of Vienna, Human Evolution and Archaeological Sciences (HEAS), Museo de ArqueologÃ-a de Alta MontaÃ±a (MAAM), Museo de ArqueologÃ-a de Alta MontaÃ±a (MAAM), University of Vienna, Human Evolution and Archaeological Sciences (HEAS), University of Vienna, Human Evolution and Archaeological Sciences (HEAS), University of Vienna, Human Evolution and Archaeological Sciences (HEAS)

In 1999, the mummies of the "Children of Llullaillaco" were discovered at the top of the Llullaillaco volcano in Argentina at 6739 m. These

mummies are among the best-preserved mummies ever found, which can be attributed to a combination of freezing temperatures, mild humidity, and an anaerobic environment. The Inca sacrificed them as part of the ritual of â€œCapacochaâ€•, likely for reasons such as enforcing social control and gaining socioeconomic benefits. In collaboration with the Museo de ArqueologÃ-a de Alta MontaÃ±a in Salta, Argentina, we performed non-invasive and non-destructive sampling on the mummies under strict supervision from the curators.
Through extensive sequencing of 27 samples of skin, anal and buccal tissue obtained from three individuals, we have recovered 5.4-fold, 14.3-fold, and 25.3-fold nuclear genomes from each of the children, respectively. We have used these to A) infer their affinities to native American populations, which gives insights into their geographical origins, and B) refine the demographic models of pre-contact South American populations. Our non-invasive non-destructive sampling approach has also enabled us to detect both commensal and pathogenic microbial and fungal species. Through our metagenomic analysis, we have recovered a high coverage Bartonella quintana genome from one of the children. To the best of our knowledge, this is the first ancient genome of this bacterium. We could also identify the presence of Human gammaherpesvirus 4 (HHV4) sequences. The study of the commensal microbiota promises to shed light on the living conditions of the individuals.

167 - Pre-Columbian Treponema pallidum in the Americas and the origin of treponematoses

Fernando González-Candelas

Fernando GonzÃ¡lez-Candelas, Marta Pla-Diaz, Kerttu Majander, Louis du Plessis, Natasha Arora, Jose Filippini, Luis Pezo Lanfranco, Sabine Eggers, Verena J. Schuenemann
University of Valencia, FISABIO, University of Valencia, University of Basel, University of Basel, ETH ZÃ¼rich, University of ZÃ¼rich, University of Sao Paulo, University of Sao Paulo, Natural History Museum , University of Basel

The origins of treponemal diseases have long remained unknown, especially considering the sudden onset of the first syphilis epidemic in the late 15th century in Europe and its hypothesized arrival from the Americas with Columbusâ€™ expeditions. Recently, ancient DNA evidence has revealed various treponemal infections circulating in early modern Europe and colonial-era Mexico. However, there is no genomic evidence of treponematosis recovered from either the Americas or the Old World that can be reliably dated to the time before the first trans-Atlantic contacts. We have obtained a high-quality (33.6X coverage) treponemal genome from nearly 2,000-year-old human remains from Brazil, most closely related to the bejel-causing agent Treponema pallidum endemicum (TEN). Three additional, low-coverage genomes from the same site were also obtained, all clustering in the TEN clade. For their analysis, along with representative ancient and modern genomes of the main T.pallidum lineages, we have developed a new mapping strategy that substantially improves variant calling, thus providing better resolution for phylogenetic inference, including the identification of new recombination events. Contradicting the modern-day distribution of bejel, the results call into question the previous palaeopathological characterization of treponeme subspecies and showcase their adaptive potential. Also, we have obtained improved molecular clock estimations, placing the divergence of modern T.pallidum subspecies firmly in pre-Columbian times. Overall, our study demonstrates the potential of paleogenomics to uncover key events in pathogen evolution and emergence, paving the way to new hypotheses on the origin and spread of treponematoses.

257 - A comprehensive investigation of woolly mammoth remains associated microbes

Benjamin Guinet

Benjamin Guinet, Nikolay Oskolkov, Love DalÃ©n, Tom van der Valk
Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Centre for Palaeogenetics, Department of Biology, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Lund University, Centre for Palaeogenetics, Department of

Zoology, Stockholm University, Centre for Palaeogenetics, Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Centre for Palaeogenetics

Throughout history, the exploration of ancient DNA (aDNA) has primarily centered on the evolution and demographics of humans and animals, emphasizing the analysis of eukaryotic aDNA. However, the discovery of microbial sequence data has revolutionized this field by uncovering the significance of host-associated microbial aDNA found within eukaryotic remains. Once considered a mere sequencing by-product, this microbial aDNA has emerged as a crucial source of insights into ancient pandemics, lifestyle patterns, and population movements. Despite the broad potential of obtaining evolutionary microbial insights from ancient DNA remains, the vast majority of studies to date have remained focused on humans. In this study we identified the microbial DNA found within 495 woolly mammoth remains, comprising tooth, tusk and skins samples, collected from a broad geographic range and timescale spanning the holocene (~4300 years old) up to samples over a million years in age. This extensive dataset enables us to assess and formulate hypotheses regarding numerous potential interactions between these ancient microorganisms and their hosts. Among the identified microorganisms, several related bacteria are known to interact with contemporary animals in either commensal or pathogenic ways. Taken together, these results will pave the way for a better understanding of the microbial interactions with mammoth populations over time.

280 - Genomic traces of ancient pathogens in Central Patagonia (6000-100yBP)

Florencia -- Alvarez Gallego

Florencia Alvarez Gallego, Laura Carrillo Olivas, Miriam Jetzabel Bravo LÃ³pez , Kelly Blevins , Viridiana Villa Islas , Ernesto Garfias Morales , Alejandra Castillo , Jorge Alejandro Suby, MarÃ-a del Carmen Ã•vila-Arcos , MarÃ-a Laura Parolin
Instituto de Diversidad y EvoluciÃ³n Austral, CCT- Conicet-Cenpat., Universidad Nacional de la Patagonia San Juan Bosco, International Laboratory for Human Genome Research. UNAM, International Laboratory for Human Genome Research. UNAM, Center for Bioarchaeological Research, School of Human Evolution and Social Change, Department of Archaeology, Durham University, International Laboratory for Human Genome Research. UNAM, International Laboratory for Human Genome Research. UNAM, International Laboratory for Human Genome Research. UNAM, Instituto de Investigaciones ArqueolÃ³gicas y PaleontolÃ³gicas del Cuaternario Pampeano (INCUAPA), UNCPBA-CONICET, International Laboratory for Human Genome Research. UNAM, Instituto de Diversidad y EvoluciÃ³n Austral, CCT- Conicet-Cenpat

Understanding the prevalence and evolution of ancient pathogens in Patagonia, the last peopled part of the Americas, is crucial for revealing health dynamics and evolutionary history of the hunter-gatherer populations that inhabited this region. This study provides molecular insights into endemic infections among hunter-gatherers in Central Patagonia during a 6000 yBP transect. We examined a total of 36 human remains, including dental and bones samples with and without pathological changes. Unmapped reads to the human genome were taxonomically classified using Kraken2 and MALT. As an addition parameter was assessed through coincidences in the virulence factor database VFDB2022. We identified Tannerella forsythia, Treponema denticola, and other oral bacteria linked with various stages of periodontal disease in an individual from the Atlantic coastal region dating back to 950 (yBP). In addition, systemic pathogens, including Proteus mirabilis and Corynebacterium urealyticum, associated with ulcers and urinary infections, were detected from the valley and coastal regions, dating back to 6000-2400 yBP, respectively. A noteworthy finding was the identification of Mycobacterium lepromatosis from Patagonia dating back to 2,000 yBP (n=1) and 4,000 yBP (n=2). This discovery has the potential to unveil previously unknown aspects of the health of ancient Patagonians and contribute to the poorly characterized evolutionary history of this pathogen, which was recently identified as a causative agent of leprosy. The next stage involves a capture-enrichment strategy of the pathogens of interest to increase the genomic coverage and facilitate phylogenetic analyses to reveal potential connections with other ancient and modern strains.

426 - Metagenomic Analysis of Oral Pathogens in Mammuthus columbi Remains from Mexico

Victoria Pastor

Victoria Pastor Martinez Ulloa, Miriam J Bravo Lopez, Alejandra Castillo Carbajal, Ernesto Garfias Morales, Angeles Tavares Guzman, Viridiana Villa Islas, Alejandro Lopez Jimenez, Mashaal Sohail, Joaquin Arroyo-Cabrales, Federico Sanchez Quinto, Maria C. Avila Arcos

International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), QuerÃ©taro, MÃ©xico., International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), QuerÃ©taro, MÃ©xico., International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), QuerÃ©taro, MÃ©xico. , International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), QuerÃ©taro, MÃ©xico., International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), QuerÃ©taro, MÃ©xico, International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), QuerÃ©taro, MÃ©xico., Laboratorio de ArqueozoologÃ-a, SubdirecciÃ³n de Laboratorios y Apoyo AcadÃ©mico, Instituto Nacional de AntropologÃ-a e Historia, Mexico City, Mexico. 3 Centro de Ciencias GenÃ³micas, Universidad Nacional AutÃ³noma de MÃ©xico, Avenida Universidad s/n, Cuernavaca, Morelos, 62220, MÃ©xico., International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), QuerÃ©taro, MÃ©xico, . Laboratorio de ArqueozoologÃ-a, SubdirecciÃ³n de Laboratorios y Apoyo AcadÃ©mico, Instituto Nacional de AntropologÃ-a e Historia, Mexico City, Mexico. 3 Centro de Ciencias GenÃ³micas, Universidad Nacional AutÃ³noma de MÃ©xico, Avenida Universidad s/n, Cuernavaca, Morelos, 62220, MÃ©xico., . International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), QuerÃ©taro, MÃ©xico, . International Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), QuerÃ©taro, MÃ©xico.

Mammuthus is an extinct genus of elephants that appeared in Africa about five million years ago. Mammuthus columbi, a species within this genus, was the only mammoth endemic to the Americas. Excavations near the Santa Lucia Military Air Base in Mexico have unearthed at least 80 Mammuthus columbi remains. These findings present an ideal opportunity to gain insight into the evolutionary and ecological dynamics of this extinct species. The study of the oral microbiome through paleogenomics offers a window into the past oral health of individuals, and it has not yet been fully explored in Mammoths. In this study, we characterized the oral metagenomic profiles of 72 mammoths found in Santa Lucia Military Air Base and Tultepec Mexico. Sequencing libraries were built from dentin samples; the resulting shotgun sequence data were used to profile ancient microorganisms' DNA. The taxonomic classification revealed a significant presence of reads associated with Propionibacterium acidifaciens and Streptococcus mutans, oral pathogens implicated in dental caries lesions in humans and other animals, in three samples. These findings provide insights into mammoth health, contributing to our understanding of ancient mammalian biology. In our future research, we plan to perform a phylogenetic placement analysis to understand the relationship between these pathogen sequences and modern strains.

429 - A Neolithic genome reveals the ancient origins and evolutionary pathways of Brucella melitensis

Louis L'Hôte

Louis L'HÃ´te, Lionel Gourichon, Kevin Daly

Smurfit Institute of Genetics, Trinity College Dublin, Ireland, CNRS,CEPAM (UMR 7264) - UniversitÃ© Nice CÃ´te d'Azur, School of Agriculture and Food Science - University College Dublin - Ireland

Brucellosis, a zoonotic bacterial infection prevalent in livestock, is a significant threat to animal and human welfare due to chronic infection and spontaneous abortion. This disease detrimentally impacts both livestock herds and communities reliant on animal

husbandry, particularly pastoralist societies. The Neolithic era in Southwest Asia witnessed the emergence of livestock domestication, fostering increased human-animal and animal-animal interactions and a greater risk of pathogen host jumping We report the retrieval and analysis of a 3.4X ancient genome of Brucella melitensis from a Neolithic sheep from MenteÅŸe
, Turkey (6024-5842 cal BCE). This ancient genome represents the oldest Brucella melitensis genome recovered to date and serves as a foundation for elucidating the global modern phylogeny and evolution of this pathogen. We estimate the speciation event between Brucella melitensis,primarily infecting sheep and goats, and Brucella abortus (primarily infecting cattle) to have occurred around 11,000 Before Present, immediately predating evidence of incipient management of goats, cattle, and sheep. Our findings suggest a plausible association between the onset of livestock domestication and the diversification and evolution of Brucella. Furthermore, we timestamp pseudogenization events within the Brucella melitensis lineage, , including several associated with the niche and pathogenesis of Brucella melitensis in its animal and human hosts.

504 - A Glimpse into Antiquity: Paleoproteomics Approach to Investigate Ancient Pathogens

Miguel Alejandro Navarro

Miguel A Navarro, Miriam B. Bravo-Lopez, Glendon J. Parker, Marcela Sandoval-Velasco, Julie K. Wesp, Maria L. Moreno-Cabrera, Juan G. CCervantes-Rosado, Alejandro Meraz-Moreno, Lourdes Marquez-Morfin, Jorge A. Gomez-Valdes, Alejandra Castillo, Ernesto Garfias, Daniel Blanco-Melo, Maria C Avila-Arcos

Undergraduate program on Genomic Sciences, UNAM, International Laboratory for Human Genome Research, UNAM, Fred Hutchinson Cancer Research Center, International Laboratory for Human Genome Research, UNAM, University of California, Center for GeoGenetics, Natural History Museum of Denmark, North Carolina State University, National Institute of Anthropology and History, National Institute of Anthropology and History, National Institute of Anthropology and History, National Institute of Anthropology and History, National Institute of Anthropology and History, International Laboratory for Human Genome Research, UNAM, International Laboratory for Human Genome Research, UNAM, Fred Hutchinson Cancer Research Center, International Laboratory for Human Genome Research, UNAM

Proteins are a rich source of evolutionary information, capable of withstanding degradation better than nucleic acids. However, the use of paleoproteomics to identify and study ancient pathogens remains limited.

Previous research in our lab identified the presence of ancient pathogen DNA in several samples from individuals buried in Mexico City dating to the 16th century, a period characterized by epidemics spread following the aftermath of the Spanish colonization. For this project, tryptic peptides were extracted from six of these samples using protocols optimized for ancient proteins. This included the implementation of a novel enzymatic digestion method for collagen. Subsequently, the tryptic peptides underwent LC-MS/MS analysis, and the resulting raw data files were processed using protein identification software, including MaxQuant, Protein Discoverer, and PEAKS.

Besides the various peptides from environmental bacteria, we detected numerous peptides of pathogenic bacteria from the genus Salmonella, as well as several peptides from the oral pathogen Tanerella forsythia. Importantly, Salmonella and T. forsythia were identified in the same individuals who had previously tested positive for aDNA from these pathogens, further supporting their presence and revealing the possibility of recovering ancient pathogen peptides from archaeological remains. This exploratory study highlights the potential of paleoproteomics in ancient pathogen research and the synergistic benefits of a multidisciplinary approach to unravel the mysteries of ancient epidemics.

Keywords: Paleoproteomics, Paleogenomics, pathogens, epidemics.

566 - Paleogenomic Analysis of Viruses from Pre-Hispanic and Colonial Mexico

Anahi Tania Sanchez

Anahi T Sanchez, Miriam J Bravo, Viridiana Villa, Ernesto Garfias, Alejandra Castillo, Julie Wesp, MArcela Sandoval, Lourdes MÃ¡rquez, Jorge Valdes, Alejandro Meraz, Daniel Blanco, MarÃ-a Ã•vila
LABORATORIO INTERNACIONAL DE INVESTIGACIÃ“N DEL GENOMA HUMANO, LABORATORIO INTERNACIONAL DE INVESTIGACIÃ“N DEL GENOMA HUMANO, LABORATORIO INTERNACIONAL DE INVESTIGACIÃ“N DEL GENOMA HUMANO, villa.islas.vi@gmail.com , LABORATORIO INTERNACIONAL DE INVESTIGACIÃ“N DEL GENOMA HUMANO, LABORATORIO INTERNACIONAL DE INVESTIGACIÃ“N DEL GENOMA HUMANO, DEPARTMENT OF SOCIOLOGY & ANTHROPOLOGY, NORTH CAROLINA STATE UNIVERSITY, CENTRO DE CIENCIAS GENÃ“MICAS, UNAM, ESCUELA NACIONAL DE ANTROPOLOGÃ•A E HISTORIA, ESCUELA NACIONAL DE ANTROPOLOGÃ•A E HISTORIA, DIRECCIÃ“N DE SALVAMENTO ARQUEOLÃ“GICO, INSTITUTO NACIONAL DE ANTROPOLOGÃ•A E HISTORIA, FRED HUTCHINSON CANCER CENTER, LABORATORIO INTERNACIONAL DE INVESTIGACIÃ“N DEL GENOMA HUMANO

The colonization of Mexico facilitated the transmission of pathogenic viruses, leading to a crisis in the indigenous population which experienced epidemics and mass deaths caused by previously unseen pathogens. The presence of infectious diseases during the pre-Hispanic and colonial periods in Mexico is evidenced in the historical and archaeological record, suggesting the existence of various pathogens. However, these type of record is often imprecise, and the exact causative agents remain unknown in many cases. Paleogenomics has enabled the reconstruction of ancient viral genomes of DNA and RNA present in archaeological remains. In this study, we leveraged paleogenomic data available for seven pre-Hispanic and 106 colonial individuals from Mexico and focused on identifying and analyzing human viral genomes. We taxonomically classified non-human reads using MALT/0.6.2 and MEGAN 6.2 using a reference database containing 5,339 viral genomes. Libraries with a higher abundance of ancient viral sequences will be selected for subsequent genome capture and enrichment. This project builds upon previous studies that have identified viral genomes in colonial Individuals from Mexico on smaller sample sizes and aims to expand our understanding of the origin and evolutionary history of viruses that circulated before and after the Spanish colonization of Mexico.

651 - Ancient Salmonella enterica genomes and the evolutionary path of the Para C lineage

Gunnar U. Neumann

Gunnar U. Neumann, Lyazzat Musralina, Maria A. Spyrou, Ainash Childebayeva, Marcel Keller, Xiaowen Jia, Angela MÃ¶tsch, Luka Papac, Sandra Penske, Vanessa Villalbla-Mouco, Arman Beisenov, Lorenc Bejko, Natalia Berezina, Alexandra Buzhilova, Kevin Daly, Michal ErnÃ©e, Miroslav Furmanek, Michaela Harbeck, Agata HaÅ‚uszko, Daniela Heilmann, Egor Kitov, RÃ¼diger Krause, Petr KriÅ¡tuf, Stephanie Larson, Maria Liston, Erdene Myagmar, Ilya Merz, Victor Merz, Mathieu Ott, Aleksandra Papazovska, Jaroslav PeÅ¡ka, Erika PrÅ¯chovÃ¡, Claude Raynaud, Roberto Risch, Zainolla Samashev, Svetlana Sharapova, Elena Sintes Olives, Adela Sobotkova, Astrid Stobbe, Tatjana Stojoska Vidovska, OndÅ™ej Å vejcar, Emma Usmanova, Michael McCormick, Elmira Khussainova, Leyla Djansugurova, Christina Warinner, Philipp W. Stockhammer, Wolfgang Haak, Alexander Herbig, Johannes Krause
Max Planck Institute for Evolutionary Anthropology, Al-Farabi Kazakh National University, Institute of Genetics and Physiology, Max Planck Institute for Evolutionary Anthropology, Eberhard-Karls-University TÃ¼bingen, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, University of Texas at Austin, University of Basel, University of Tartu, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Institute of Evolutionary Biology, CSICâ€“Universitat Pompeu Fabra, Begazy-Tasmola Research Center of History and Archeology, University of Tirana, Research Institute and Museum of Anthropology of Lomonosov Moscow State University, Research Institute and Museum of Anthropology of Lomonosov Moscow State University, Bucknell University, Institute of Archaeology, Czech Academy of Sciences, University of Wroclaw, Ludwig Maximilian University, Maria Curie-SkÅ‚odowska University, Ludwig Maximilian University , Center of Human Ecology, Institute of Ethnology and Anthropology, Russian Academy of Science, Goethe-UniversitÃ¤t Frankfurt am Main, University of West Bohemia, Bucknell

University, University of Waterloo, Department of Anthropology and Archaeology, National University of Mongolia, Institute of Archaeology, Department of History of S. Toraighyrov PSU, INRAP, Centre ArchÃ©ologique de NÃ®mes, Archaeological Museum of North Macedonia, Archaeological Centre in Olomouc, PrÃ¡cheÅˆ Museum in PÃ-sek, UMR 5140, ArchÃ©ologie des SociÃ©tÃ©s MediterranÃ©ennes, Universitat AutÃ²noma de Barcelona, State Historical and Cultural Museum-Reserve â€œBerelâ€• village Zhambyl, Branch of Institute of Archaeology by A.Kh. Margulan, Institute of History and Archaeology, Ural branch of Russian Academy of Science, Universitat AutÃ²noma de Barcelona, Aarhus University, Goethe-UniversitÃ¤t Frankfurt am Main, Institute for Protection of Cultural Monuments and National Museum Ohrid, Czech Academy of Sciences, Saryarka Archaeological Institute, Buketov Karaganda State University, Harvard University, Institute of Genetics and Physiology, Institute of Genetics and Physiology, Max Planck Institute for Evolutionary Anthropology, Harvard University, Ludwig Maximilian University , Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology, Max Planck Institute for Evolutionary Anthropology

Salmonellosis, caused by Salmonella enterica, ranks among the most prevalent food-borne illnesses worldwide. Most of the approximately 2600 S. enterica distinct serovars have the capacity to infect a broad spectrum of host species, typically resulting in self-limiting gastrointestinal illness in humans. However, a few serovars are specifically adapted to humans, inducing enteric fever. Within this group, Paratyphi C belongs to the so-called Para C lineage, alongside Typhisuis (primarily infecting swine), and Choleraesuis (affecting both humans and swine). Previous aDNA studies suggested a link between host adaptation in this lineage and the Neolithization process involving pseudogenization and gene gain/loss.
Here, we present a temporal and geographical transect of 35 new ancient S. enterica genomes spanning from 5000-500 years BP, collected from across Eurasia. The majority of genomes older than 3000 years form a distinct phylogenetic subclade lacking any modern representatives, suggesting probable extinction. The remaining genomes, dating to the last 2700 years, position along the ParatyphiC and Choleraesuis branches. Their sequential distribution enables a comprehensive examination of host adaptation, involving both the acquisition and loss of genomic loci and gene pseudogenization, and allows the reconstruction of the temporal sequence of these events. For instance, our findings indicate that the acquisition of the Salmonella pathogenicity island SPI-7 preceded the loss of the tcfABCD operon of SPI-6 in ParatyphiC. These findings provide valuable insights into the evolutionary trajectory and host adaptation process within the S. enterica Para C lineage, revealing how changes in genetic composition influenced its virulence.

701 - Metagenomic Profiling of Ancient Pathogens in Britain

Anastasia Brativnyk

Anastasia Brativnyk, Kyriaki Anastasiadou, Christopher Barrington, Thomas Booth, Alexandre Gilardet, Isabelle Glocke, Sarah Johnston, Monica Kelly, Jesse McCabe, Marina Silva, Pooja Swali, Frankie Tait, Mia Williams, Pontus Skoglund
The Francis Crick Institute, The Francis Crick Institute, The Francis Crick Institute, The Francis Crick Institute, The Francis Crick Institute, The Francis Crick Institute, The Francis Crick Institute, The Francis Crick Institute, The Francis Crick Institute, The Francis Crick Institute, The Francis Crick Institute, The Francis Crick Institute

In the last decade, the characterisation of ancient pathogens has significantly expanded our understanding of paleoepidemiology and the evolution of infectious diseases, such as plague, smallpox, and tuberculosis. Our laboratory has sequenced ancient DNA samples from across Britain, spanning an 11,000-year period. This time transect provides a basis for a comprehensive study of pathogens within a single region. Investigating the prevalence and dynamics of infectious diseases, clustered by time and space, may help us to uncover previously unknown outbreaks. Furthermore, constructing a spatiotemporal map of pathogens in Britain, allows us to study phylogenetic patterns to date divergences between lineages and taxa and to understand the timing of evolutionary transitions. Additionally, we plan to explore the evolutionary trajectory of specific pathogens, focusing on alterations in functional genomic elements, such as point

mutations, insertions and deletions, that signify gains or losses of function over time. Mapping the landscape of ancient pathogens in Britain and examining key hypotheses into the evolutionary forces driving infectious diseases, has the potential not only to illuminate the history of diseases but also to contribute to modern public health strategies and improve our preparedness for future outbreaks.

727 - Exploring the landscape of infectious diseases in the Americas from pre-colonial to modern times with ancient DNA approaches

Maria Lopopolo

Maria Lopopolo, NicolÃ¡s Rascovan

Institut Pasteur, Universitè de Paris, CNRS UMR 2000, Microbial Paleogenomics Unit, Universitè de Paris, INSERM, System Engineering and Evolution Dynamics, Institut Pasteur, Universitè de Paris, CNRS UMR 2000, Microbial Paleogenomics Unit

Pathogen paleogenomics is an emerging field that offers unique opportunities to investigate how pathogens and human microbes emerged, spread and changed at the genomic level over the course of human history. This information can be used to better understand the complexity of modern infectious diseases. To do so, this field brings together and/or borrows tools, concepts, and approaches from other areas, such as ancient DNA, phylogenetics, microbial genomics, evolution, population genetics, bioinformatics, history and archaeology. While most of the current knowledge about infectious diseases derives from studies on modern microbial strains, ancient pathogens genomes can provide key additional information about where and when these species were in the past and determine the relationships and genomic changes that took place over time. We used these approaches to investigate a long-standing question regarding which infectious diseases prevailed in the Americas before the European invasion and how this determined epidemiological scenarios today in the continent. We identified pathogenic bacterial species in archaeological samples that continue to infect human populations in the Americas today. Comparison of modern and ancient strains allowed us to reconstruct epidemiological scenarios to infer how human mobility could have determined the expansion of infectious diseases in the continent. Our work represents a glimpse into a much more complex scenario for the history of infectious diseases in the continent than it is commonly considered.

782 - Streptococcus evolutionary diversity in ancient Great Britain and its associations with oral health outcomes

Ava Gabrys

Ava Gabrys, Abigail Gancz, Laura Weyrich

Department of Biology, The Pennsylvania State University, Department of Anthropology, The Pennsylvania State University, Department of Anthropology, The Pennsylvania State University, Huck Institutes of Life Science, The Pennsylvania State University, Australian Centre for Ancient DNA, University of Adelaide

Streptococcus species play key roles in tooth decay, opportunistic infections, and structuring the oral microbiome through adhesion and interspecies interactions. Retracing the evolutionary history of these bacteria is critical to understanding the origins of modern diseases, including the costly development of dental caries. We used ancient DNA recovered from the dental calculus of 162 British individuals dating 2100BCE-1853CE to explore changes in streptococci communities over time. Employing a competitive mapping approach to finely characterize species abundances, we found major shifts in dominant Streptococcus species and their genomes occurred after the post-medieval period. Streptococcus sinensisâ€”a species poorly described in populations today except in its involvement in infective endocarditisâ€”was the dominant Streptococcus species in 54% of ancient samples, and phylogenomics suggest it had a stable population closely related to Streptococcus cristatus before significantly decreasing in abundance in the modern day. Streptococcus sanguinis is consistently the most abundant species in healthy modern individuals but was dominant in only 22% of ancient samples,

where it was missing loci for membrane transport, stress response, and sugar usage that are variably present in modern individuals with periodontitis. Surprisingly, although the species thrives in modern carbohydrate-rich oral environments, Streptococcus mutans was nearly absent. We did not find significant associations between the presence of any Streptococcus species and oral pathologies, including caries and periodontitis. These compositional, genomic, and functional shifts suggest changing roles and competition of Streptococcus species post industrialization, increasing our understanding of their relationships to human health and disease.

834 - Reconstruction of one Salmonella enterica Paratyphi C genome from 19th-century Mexico City

Miriam Bravo-Lopez

Miriam Bravo-Lopez, Eduardo Arrieta, Ã…shild Joanne Vagene, Viridiana Villa-Islas, Ana Beatriz VillaseÃ±or, Sandra Elena Guevara-Flores, Ernesto Garfias-Morales, Jean Cury, Maria MorenoCabrera, Alejandro Meraz-Moreno, Jorge Gomez-Valdes, Flora Jay, Emilia Huerta-Sanchez, Maria Avila-Arcos

nternational Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), nternational Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), Section for Evolutionary Genomics, GLOBE Institute, University of Copenhagen, nternational Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), Brigham and Womenâ€™s Hospital in the Pulmonary and Critical Care Medicine Division, Center for Anthropological Studies, Faculty of Political and Social Sciences, National Autonomous University of Mexico (CEA-FCPyS-UNAM), nternational Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM), UniversitÃ© Paris-Saclay, National School of Anthropology and History, UNAM, National School of Anthropology and History, UNAM, National School of Anthropology and History, UNAM, UniversitÃ© Paris-Saclay, CNRS UMR 9015, INRIA, Laboratoire Interdisciplinaire des Sciences du NumÃ©rique, Ecology, Evolution and Organismal Biology, Brown University, nternational Laboratory for Human Genome Research, Universidad Nacional AutÃ³noma de MÃ©xico (UNAM)

The European Imperial expansion during the 19th century in Mexico, resulted in new infectious diseases due to ecological disruptions. To gain insights into the infectious agents introduced during this period, we generated paleogenomic data from seven individuals dated to the 19th century from Mexico City. Taxonomic classification of the obtained sequences enabled us to identify ancient DNA recovered from one individual as Salmonella enterica Paratyphi C. This pathogen causes paratyphoid fever in humans. Remarkably, S. Paratyphi C has been proposed as one of the causative agents of the 1545-1550 cocoliztli epidemic, resulting in an estimated 15 million deaths. To better understand the evolutionary history of S. Paratyphi C in ancient Mexico, we reconstructed its genome at a 10x depth using an in-house capture-enrichment strategy. The reconstructed genome was phylogenetically analyzed alongside ancient and modern genomes of S. Paratyphi C. Our genome falls within a branch closely related to the ancient S. Paratyphi C genomes from south Mexico. The divergence time between them was estimated to 500 years BP, which coincides with the Spanish conquest. Altogether, this study demonstrates that a replacement occurred within S. enterica Paratyphi C after the arrival of the Spanish conquistadors in New Spain.

845 - Exploring the History of Malaria in the Americas Using Ancient DNA

Megan Michel

Megan Michel, Eirini Skourtanioti, Federica Pierini, Evelyn K. Guevara, Angela MÃ¶tsch, Arthur Kocher, Rodrigo Barquera, Raffaela A. Bianco, Selina Carlhoff, Lorenza Coppola Bove, Suzanne Freilich, Karen Giffin, Taylor Hermes, Alina HiÃŸ, Florian Knolle, Elizabeth A. Nelson, Gunnar U. Neumann, Luka Papac, Sandra Penske, Adam B. Rohrlach, Nada Salem, Lena Semerau, Vanessa Villalba-Mouco, Isabelle Abadie, Mark Aldenderfer, Jessica F. Beckett, Matthew Brown, Franco G. R. Campus, Tsang Chenghwa, MarÃ-a Cruz Berrocal, Ladislav DamaÅ¡ek, Kellie Sara Duffett Carlson, RaphaÃ«l Durand, Michal ErnÃ©e, Cristinel FÃ¢ntÃ¢neanu, Hannah Frenzel, Gabriel GarcÃ-a AtiÃ©nzar, Sonia GuillÃ©n, Ellen Hsieh, Maciej Karwowski, David Kelvin, Nikki Kelvin, Alexander Khokhlov, Rebecca L.

Kinaston, Arkadii Korolev, Kim-Louise Krettek, Mario KÃ¼ÃŸner, Luca Lai, Cory Look, Kerttu Majander, Kirsten Mandl, Vittorio Mazzarello, Michael McCormick, Patxuka de Miguel IbÃ¡Ã±ez, Reg Murphy, Rita E. NÃ©meth, Kerkko Nordqvist, Friederike Novotny, Martin Obenaus, Lauro Olmo-Enciso, PÃ¤ivi Onkamo, JÃ¶rg Orschiedt, Valerii Patrushev, Sanni Peltola, Alejandro Romero, Salvatore Rubino, Antti Sajantila, Domingo C. Salazar GarcÃ-a, Elena Serrano, Shapulat Shaydullaev, Emanuela Sias, Mario Ã laus, Ladislav StanÄ•o, Treena Swanston, Maria Teschler-Nicola, Frederique Valentin, Katrien Van de Vijver, Tamara L. Varney, Alfonso Vigil-Escalera Guirado, Christopher K. Waters, Estella Weiss-Krejci, Eduard Winter, Thiseas C. Lamnidis, Kay PrÃ¼fer, Kathrin NÃ¤gele, Maria Spyrou, Stephan Schiffels, Philipp W. Stockhammer, Wolfgang Haak, Cosimo Posth, Christina Warinner, Kirsten I. Bos, Alexander Herbig, Johannes Krause

Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Human Evolutionary Biology, Harvard University, Max Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Max Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Forensic Medicine, University of Helsinki, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Max Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Transmission, Infection, Diversification and Evolution Group, Max Planck Institute of Geoanthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Max Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Max Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Department of Legal Medicine, Toxicology and Physical Anthropology, University of Granada, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Evolutionary Anthropology, University of Vienna, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Anthropology, University of Arkansas, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Medical Engineering and Biotechnology, University of Applied Sciences Jena, Microbial Paleogenomics Unit, Department of Genomes and Genetics, Institut Pasteur, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Max Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, School of Computer and Mathematical Sciences , University of Adelaide, Adelaide Data Science Centre, University of Adelaide, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Max Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Instituto Universitario de InvestigaciÃ³n en Ciencias Ambientales de AragÃ³n, IUCA-Aragosaurus, Universitity of Zaragoza, Max Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Inrap - Institut national de recherches archÃ©ologiques prÃ©ventives, CRAHAM - Centre Michel de BoÃ¼ard - Centre de recherches archÃ©ologiques et historiques anciennes et mÃ©diÃ©vales, UniversitÃ© de Caen Normandie, Department of Anthropology and Heritage Studies, University of California, Merced, Private contractor, Sociology and Anthropology Department, Farmingdale State College, Medieval Archeologist, Freelance, Institute of Anthropology, National Tsing Hua University, Institute of Heritage Sciences (INCIPIT), Spanish Research Council (CSIC), Institute of Classical Archaeology, Faculty of Arts, Charles University, Human Evolution and Archaeological Sciences, University of Vienna, Service dâ€™archÃ©ologie prÃ©ventive Bourges plus, UMR 5199 PACEA, UniversitÃ© de Bordeaux, Department of Prehistoric Archaeology, Institute of Archaeology of the Czech Academy of Sciences,

National Museum of Unification Alba Iulia, Anatomy Institute, University of Leipzig, Instituto Universitario de InvestigaciÃ³n en ArqueologÃ-a y Patrimonio HistÃ³rico, Universidad de Alicante, Centro Mallqui, Institute of Anthropology, National Tsing Hua University, Institut fÃ¼r Urgeschichte und Historische ArchÃ¤ologie, University of Vienna, Department of Microbiology and Immunology, Dalhousie University, Division of Ancient Pathogens, BioForge Canada Limited, Journal of Infection in Developing Countries, Samara State University of Social Sciences and Education, BioArch South, Griffith Centre for Social and Cultural Studies, Griffith University, Nathan Campus, Samara State University of Social Sciences and Education, Senckenberg Centre for Human Evolution and Palaeoenvironment at the University of TÃ¼bingen, Thuringian State Office for Heritage Management and Archeology, Department of Anthropology, University of North Carolina at Charlotte, Department of Anthropology, Brooklyn College, Department of Environmental Science, Integrative Prehistory and Archaeological Science, University of Basel, Department of Evolutionary Anthropology, University of Vienna, Department of Biomedical Sciences, University of Sassari, Initiative for the Science of the Human Past at Harvard and Max Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Department of History, Harvard University, Instituto Universitario de InvestigaciÃ³n en ArqueologÃ-a y Patrimonio HistÃ³rico, Universidad de Alicante, Servicio de Obstetricia, Hospital Virgen de los Lirios, SecciÃ³n de AntropologÃ-a, Sociedad de Ciencias Aranzadi, University of Nebraska-Lincoln, MureÈ™ County Museum, Helsinki Collegium of Advanced Studies, University of Helsinki, Department of Anthropology, Natural History Museum Vienna, Silva Nortica ArchÃ¤ologische Dienstleistungen, Department of History, University of AlcalÃ¡, Department of Biology, University of Turku, Landesamt fÃ¼r Denkmalpflege und ArchÃ¤ologie Sachsen-Anhalt, Institut fÃ¼r PrÃ¤historische ArchÃ¤ologie, Freie UniversitÃ¤t Berlin, Mari State University, Centre of Archaeological and Ethnographical Investigations, Faculty of Biological and Environmental Sciences, University of Helsinki, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Instituto Universitario de InvestigaciÃ³n en ArqueologÃ-a y Patrimonio HistÃ³rico, Universidad de Alicante, Departamento de BiotecnologÃ-a, Universidad de Alicante, Department of Biomedical Sciences, University of Sassari, Department of Forensic Medicine, University of Helsinki, Forensic Medicine Unit, Finnish Institute for Health and Welfare, Departament de PrehistÃ²ria, Arqueologia i HistÃ²ria Antiga, Universitat de ValÃ¨ncia, Department of Geological Sciences, University of Cape Town, Universidad de Cantabria, Instituto Internacional de Investigaciones PrehistÃ³ricas, TAR ArqueologÃ-a, Faculty of History, Termez State University, Centro Studi sulla CiviltÃ del Mare, Anthropological Center, Croatian Academy of Sciences and Arts, Institute of Classical Archaeology, Faculty of Arts, Charles University, Department of Anthropology, Economics and Political Science, MacEwan University, Department of Anthropology, Natural History Museum Vienna, Department of Evolutionary Anthropology, University of Vienna, CNRS, UMR 8068, Royal Belgian Institute of Natural Sciences, Center for Archaeological Sciences, University of Leuven, Dienst Archeologie - Stad Mechelen, Department of Anthropology, Lakehead University, Departamento de Humanidades: Historia, GeografÃ-a y Arte, Universidad Carlos III de Madrid, Heritage Department, National Parks of Antigua and Barbuda, Nelsons Dockyard, English Harbour, , Austrian Archaeological Institute, Austrian Academy of Sciences, Institut fÃ¼r Ur- und FrÃ¼hgeschichte, Heidelberg University, Department of Social and Cultural Anthropology, University of Vienna, Department of Anthropology, Natural History Museum Vienna, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Archaeo- and Palaeogenetics, Institute for Archaeological Sciences, Department of Geosciences, University of TÃ¼bingen, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Max Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Ludwig Maximilian University, Institute for Pre- and Protohistoric Archaeology and Archaeology of the Roman Provinces, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Archaeo- and Palaeogenetics, Institute for Archaeological Sciences, Department of Geosciences, University of TÃ¼bingen, Senckenberg Centre for Human Evolution and Palaeoenvironment at the University of TÃ¼bingen, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Anthropology, Harvard University, Max

Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, HistoGenes, Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Max Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM)

Malaria is a vector-borne disease caused by parasitic protozoa of the genus Plasmodium. It exerts one of the strongest selective pressures acting on the human genome and causes over 600,000 deaths each year. Nevertheless, significant questions remain regarding when and how particular malaria-causing parasites emerged as human pathogens and spread around the globe. In particular, debate persists over whether the two most widespread and clinically relevant malaria agents, Plasmodium falciparum and Plasmodium vivax, were present in the Americas prior to European contact, or whether they spread from Europe and/or Africa during the colonial period. To investigate this question, we present genome-wide nuclear and mitochondrial data from 36 ancient P. vivax, P. falciparum, and P. malariae strains, spanning 16 countries and 5,500 years of human history, including a high-coverage P. vivax genome recovered from a Chachapoya individual from the peri-contact Peruvian site of Laguna de los CÃ³ndores. Genomic analysis of this dataset reveals close links between now-eliminated European P. vivax and ancient and modern Latin American lineages, suggesting a European source for strains circulating in the Americas today. Moreover, the Laguna de los CÃ³ndores strain shows greater affinity to present-day Peruvian P. vivax compared to other Latin American populations, providing evidence for persistence of an endemic malaria focus from the early colonial period until today. Overall, our results provide a first glimpse into the processes which allowed P. vivax to gain a foothold in the Americas during the period of the Columbian Exchange.

966 - Exploring the Health of Colonial Enslaved Communities at Hacienda La Quebrada: A Preliminary Study using Metagenomic Methods

Jaime Zolik

Jaime Zolik, Laura Pott, JosÃ© Luis Santa Cruz AlcalÃ¡, Claire Maass, Maria A Nieves-ColÃ³n
University of Minnesota-Twin Cities, University of Minnesota-Twin Cities, Universidad Nacional Mayor de San Marcos, APACHETA Consultores, MUAFRO (Afro-Peruvian Museum of San Luis), SNA International, APACHETA Consultores, MUAFRO (Afro-Peruvian Museum of San Luis), University of Minnesota-Twin Cities

The La Quebrada Archaeology Project, an ongoing community-led initiative, investigating an 18th century unmarked historical cemetery located in coastal Peru. The cemetery holds the remains of enslaved African and African-descendant laborers from the former plantation Hacienda La Quebrada. To contribute to the overarching goal of unraveling the history of these individuals and the broader context of African enslavement in Latin America, our laboratory conducted the extraction and analysis of ancient DNA (aDNA) from the dental pulp of 30 cemetery individuals. The recovered endogenous human aDNA had degradation patterns consistent with authentic aDNA. In conjunction with human aDNA analysis, our aim was to detect the presence of ancient microbes within our shotgun sequences to provide insights into the health, life histories, and potential disease dynamics of enslaved communities during the colonial period of Peru. Using established metagenomic tools, such as KrakenUniq and MALT, preliminary results suggest the presence of authentic Streptococcus gordonii, an opportunistic oral pathogen, in one individual displaying osteological indicators of dental disease (e.g., gum resorption, dental plaque). Aligned reads covered 54.85% (average 2.5x depth) of the S.gordonii reference genome. Although S. gordonii can live in environments outside of the human body, the aligned reads exhibited the same aDNA degradation pattern observed in the recovered human aDNA, indicating its authenticity. Future work includes examining phylogenetic context and conducting comparative analysis to modern local S.gordonii genomes. This study shows how combining paleogenomics and bioarch can lead to more holistic reconstructions of health in the past.

1095 - Characterizing the cariogenic bacterium Streptococcus mutans in ancient and modern Chile

Shreya L Ramachandran

Shreya L Ramachandran, Ricardo A Verdugo, Paola Gonzalez, Mauricio Moraga, Sergio Gonzalez Providell, Josefa NuÃ±ez Belmar , Jaime AndrÃ©s Rivas Pardo, Constanza de la Fuente, Maanasa Raghavan

University of Chicago, NA, Universidad de Talca, Universidad de Talca, Universidad de Chile, Universidad Mayor, Universidad Mayor , Universidad Mayor, Universidad de Chile, University of Chicago

Here, I present preliminary data from tooth samples from eight individuals from an archaeological site in the Near North region in Chile, dating between 1200 and 1400 BCE. The individuals, representing different times and a variety of archaeological contexts, all show genomic evidence of ancient Streptococcus mutans, a bacterium responsible for causing cavities. I designed a custom capture panel to gain higher-coverage data of a set of evolutionarily and medically relevant S. mutans genes. Using the combined data from shotgun sequencing and the capture panel, I conducted phylogenetic analyses that use these ancient time-stamped samples to refine estimates of the evolutionary trajectory of S. mutans and its relationship to human agricultural transitions. I will also discuss the adaptation of S. mutans to its local environment, by analyzing the presence and absence of virulence factors in an archaeological context compared with contemporary S. mutans strains. In collaboration with archaeologists and anthropologists, I will study the samples in the context of the diet and lifestyle of these individuals.

From these analyses, we can gain a greater understanding of determinants of oral health in ancient South America, a region that has been historically understudied from an ancient metagenomic perspective. At the same time, by studying the evolutionary history of pathogens that continue to affect us today, we can gain insight into potential targets for improving human health.

1107 - Depletion of the plasminogen activator (pla) virulence gene across pandemics of plague

Ravneet Kaur Sidhu

Ravneet K Sidhu, Guillem Mas Fiol, Christian E. Demeure , Pierre Le Bury, Katherine Eaton, Jesper L. Boldsen, Jennifer Klunk, Dan Salked , Ann Carmichael, Nukhet Varlik, David Earn, Brian Golding, Olivier Dussurget, Edward C. Holmes, Javier Pizarro-Cerda, Hendrik N. Poinar

McMaster Ancient DNA Centre, Department of BIology, McMaster University, Yersinia Research Unit, Institut Pasteur , Yersinia Research Unit, Institut Pasteur, Yersinia Research Unit, Institut Pasteur, McMaster Ancient DNA Centre, McMaster University, Department of Forensic Medicine, Unit of Anthropology (ADBOU), University of Southern Denmark, McMaster Ancient DNA Centre, McMaster University, Daciel Arbor Biosciences, Department of Biology, Colorado State University, History Department, Indiana University, Department of History, Rutgers University, Department of Mathematics and Statistics, McMaster University , Michael G. DeGroote Institute of Infectious Disease Research, McMaster University, Department of Biology, McMaster University, Yersinia Research Unit, Institut Pasteur, Faculty of Medicine and Health, The university of Sydney, Yersinia Research Unit, Institut Pastuer, McMaster Ancient DNA Centre, Departments of Anthropology, Biochemistry, and Biology, McMaster University, Michael G. DeGroote Institute of Infectious Disease Research, McMaster University, Humans and the Microbiome Program, Canadian Institute for Advanced Research

Yersinia pestis killed ~30-50% of Europeâ€™s population during its Black Death epidemic (1345-1353). This was followed by approximately 400 years of post-Black Death outbreaks throughout Afro-Eurasia. These persistent post-Black Death waves are believed to have had lower mortality rates, potentially due to virulence attenuation, host immunity, cultural adaptations, or synergistic interactions between them. Genomic studies of post-Black Death Y. pestis genomes have identified a depletion of the key virulence factor pla, on the high copy number PCP1 plasmid, attempting to link this genomic event with the disappearance of plague from Europe. In this work, we challenge the current perspectives of pla depletion and take steps towards characterizing its genomic and functional characteristics. We

start by computationally screening ancient and modern Yersinia samples (n = 2,914), finding previously undescribed instances of this convergent depletion. We then perform targeted capture of pPCP1 on ancient Y. pestis samples from geographically and temporally distinct sites throughout Medieval Denmark, allowing for de novo assembly of pPCP1 (pla -). Taking a well characterized pla depleted sample we sequence it to a depth of 400M to evaluate the genomic movement of pla in second pandemic strains. Finally, our work integrates functional data providing the first insights into the in-vitro and in-vivo impacts of pla depletion on Y. pestis virulence.