SMBE 2025

ANNUAL MEETING OF THE SOCIETY FOR MOLECULAR BIOLOGY & EVOLUTION, 2025

July 20–24, 2025 · Beijing, China

Organized by Chinese Academy of Sciences

ABSTRACT BOOK

Contents

Oral **5**

Keynotes **5**

Keynote 1: Six impossible things to believe before breakfast, about gene flow 6

Keynote 2: The world of viruses and its evolution through the lens of metagenomics and metatranscriptomics 7

Keynote 3: Why do human germline mutation rates depend on sex and age? 8

Keynote 4: How and why should we sequence thousands of genomes? 9

Keynote 5: The genetic history of human populations and adaptations in East Asia since the Pleistocene 10

Graduate Student Excellence symposium **11**

Sex-biased duplicates are generated during metazoan tRNA repertoire expansion 12

Disentangling Complex Histories of Hybridisation: The Genomic Consequences of Ancient and Recent Introgression in Channel Island Monkeyflowers 13

The march to land: unearthing how gene repertoire evolution triggered terrestrialization across the Animal Tree of Life 14

A fresh approach for estimating selection on synonymous codons 15

Estimating effective population size from ARGs: An EM-based approach incorporating coalescent time uncertainty 16

How museomics can be used to reveal whether insect genomes experienced recent anthropogenic changes 17

Insights into the development of multiple sperm types and evolving sex chromosomes in drosophila species 18

Coupling AI and SMC-based algorithms: Inference of population structure from single genome sequencing analysis 19

Symposia **20**

Symposium 1: Crop evolution: from Darwin's legacy to genomics of admixture and climate resilience 21

Symposium 2: Evolution and regulation of sex chromosomes and sex determination 33

Symposium 8: Novel insights on genome architecture evolution 103
Symposium 10: Pangenome graphs and their applications in biodiversity genomics 123
Symposium 11: Phylogenomics in the presence of gene flow: opportunities and challenges 130
Symposium 12: Population genetics through time session A 138
Symposium 13: Synthetic approaches to evolutionary theory: A dialog between the past and the future 150
Symposium 14: The roles of hybridization and introgression in evolution 157
Symposium 16: Thematic symposium: evo-devo 186
Symposium 17: Thematic symposium: regulatory genomics and non-coding elements 195
Symposium 18: Thematic symposium: frontiers in single-cell level analysis 211
Symposium 19: Thematic symposium: conservation genomics 220
Symposium 20: Thematic symposium: evolution of immunity 234
Symposium 21: Thematic symposium: virus evolution and pathogen evolution 243
Symposium 28: Human history and past social organisation in the light of palaeogenomes: new methods, new findings session B 315
Symposium 29: Open symposium session A 326
Symposium 31: Open symposium session C 351
Symposium 34: Open symposium session D 367
Symposium 35: Open symposium session E 381

Poster 395

Session 1 396

Symposium 1: Crop evolution: from Darwin's legacy to genomics of admixture and climate resilience 397
Symposium 2: Evolution and regulation of sex chromosomes and sex determination 405
Symposium 4: Gene trees and beyond: genealogical methods and their applications in evolutionary genetics 421
Symposium 5: Human history and past social organisation in the light of palaeogenomes: new methods, new findings session A 443

Symposium 6: Mutation rate rules: on the origin of germline mutations and their evolutionary fate 451
Symposium 7: The Origin, Evolution, and Phenotypic Contributions of New Genes 458
Symposium 9: Origin and evolution of cell types 473
Symposium 10: Pangenome graphs and their applications in biodiversity genomics 485
Symposium 29: Open symposium session A 491
Symposium 31: Open symposium session C 509
Session 2 524
Symposium 35: Open symposium session E 525
Symposium 3: Experimental evolution: studying evolution in action 528
Symposium 8: Novel insights on genome architecture evolution 545
Symposium 11: Phylogenomics in the presence of gene flow: opportunities and challenges 579
Symposium 13: Synthetic approaches to evolutionary theory: A dialog between the past and the future 589
Symposium 14: The roles of hybridization and introgression in evolution 595
Symposium 15: Transposable elements in genome evolution and biotechnology development 622
Symposium 16: Thematic symposium: evo-devo 640
Symposium 16: Thematic symposium: evo-devo 647
Symposium 20: Thematic symposium: evolution of immunity 655
Symposium 28: Human history and past social organisation in the light of palaeogenomes: new methods, new findings session B 667
Session 3 674
Symposium 12: Population genetics through time session A 674
Symposium 27: Population genetics through time session B 704
Symposium 17: Thematic symposium: regulatory genomics and non-coding elements . 725
Symposium 19: Thematic symposium: conservation genomics 740
Symposium 21: Thematic symposium: virus evolution and pathogen evolution 756

Symposium 22: Thematic symposium: evolution theory and recent methodological advances 776
Symposium 23: Thematic symposium: Symbiosis 792
Symposium 24: Thematic symposium: colonizing new environments and invasiveness 799
Symposium 25: Advancing Inclusion, Diversity, Equity, and Accessibility in Molecular Biology and Evolution 810
Symposium 30: Open symposium session B 815

Oral

Keynotes

Keynote 1: Six impossible things to believe before breakfast, about gene flow

Ziheng Yang

University College London

Abstract Analysis of genomic sequence data in the past two decades has highlighted the pervasive nature of hybridization or introgression between species across the tree of life. There has been a keen interest in detecting gene flow and estimating its rate using genomic data, as this is important to reconstruction of species phylogenies, estimation of species divergence times, delineating species boundaries, and understanding speciation. There has been a burst of effort to develop methods for inferring gene flow, including simple methods based on data summaries that are computationally efficient and highly sophisticated parametric methods that are statistically efficient. In this talk I will describe our recent efforts to understand signals of gene flow in genomic data. I will describe a number of counter-intuitive results, concerning the direction of gene flow, gene flow involving sister and nonsister lineages, and the impact of gene flow on definition of species.

Keynote 2: The world of viruses and its evolution through the lens of metagenomics and metatranscriptomics

Eugene V. Koonin

National Institutes of Health, USA

Abstract Viruses and virus-like mobile genetic elements are ubiquitous parasites or symbionts of all cellular life forms and the most abundant biological entities on earth. The recent, unprecedented advances of comparative genomics, metagenomics and metatranscriptomics have led to the discovery of diverse novel groups of viruses and a rapid expansion of the chartered region of the virosphere. These discoveries provide for a vastly improved understanding of the evolutionary relationships within the virosphere. Arguably, we are approaching the point when the global architecture of the virus world can be outlined in its entirety, and the key evolutionary events in each of its domains can be reconstructed. I will present such an outline of the global organization of the virosphere and the corresponding megataxonomy, including 7 evolutionarily coherent realms of viruses, that has been recently approved by the International Committee on Taxonomy of Viruses, as well as additional candidate major taxa including new realms. The expansion of the prokaryotic virosphere that now includes many groups of viruses, particularly, those with RNA genomes, previously thought to be eukaryote-specific, will be emphasized. I will further discuss the position of viruses within the wider space of replicators and the recent dramatic expansion of the "alternative virosphere" that includes viroids and diverse viroid-like viruses that seem to have evolved on multiple, independent occasions. Finally, I will present the current scenario for the origin of viruses based on the analysis of the evolutionary provenance of the key proteins involved in virion structure formation and viral genome replication as well as the reconstruction of the viromes of Last Universal Cellular Ancestor (LUCA) and the Last Eukaryotic Common Ancestor (LECA).

Keynote 3: Why do human germline mutation rates depend on sex and age?

Molly Przeworski

Columbia University

Abstract Germline mutation is the source of all heritable differences and therefore of fundamental importance. In mammals, it has long been appreciated that mutation rates are higher in fathers, particularly older fathers. The textbook view has long been that these patterns reflect replication errors that accrue during spermatogenesis. I will present multiple lines of evidence that call this view into question. I will argue instead that the data are best explained by a much larger role of DNA damage in the genesis of germline mutations than previously appreciated, and draw implications for why mutation rates depend on sex and age and how they evolve over time.

Keynote 4: How and why should we sequence thousands of genomes?

Peter WH Holland

University of Oxford

Abstract In 2019, we proposed the Darwin Tree of Life (DToL) project, a major collaborative project aiming to generate a high quality genome sequence for every major species in the ecosystems of Britain and Ireland. This is a large collaborative project involving the Wellcome Sanger Institute plus research teams in Oxford, Cambridge, London, Edinburgh, Plymouth and Norwich. In Phase 1, from 2020 to 2025, the DToL collaboration successfully collected and identified 7000 eukaryote species; of these, our team in Oxford contributed 2000 arthropod species from one small woodland. DToL has sequenced the genomes of 1700 species so far; all raw sequence data and the assembled genomic data are freely available. I will discuss the project and some recent analyses focussing on new insights into genome evolution in insects.

Keynote 5: The genetic history of human populations and adaptations in East Asia since the Pleistocene

Qiaomei Fu

Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences

Abstract Advances in extracting and sequencing ancient biomolecules—particularly genomes and proteomes from archaeological remains—have opened unprecedented windows into the deep population history of East Asia. Despite the challenges posed by the region's limited preservation of biomolecules due to its relatively warm and humid climate, we've reconstructed thousands of ancient molecules from human fossils and sediments across China, dating back to at least 146,000 years ago, to illuminate a complex and dynamic evolutionary landscape of the East Asian populations since the Pleistocene. Our analysis includes the Denisovan proteomes and mitochondrial genomes from nearly complete scull, dental calculus and sediments (146ka-45ka) to unveil the morphology and spatiotemporal distribution of this archaic population, the nuclear genomes (40-11 ka) of diversified lineages of modern humans from northern and southern East Asia to elucidate their dynamics and adaptions to environmental changes during Ice Age, and the genome-wide data (11-0.1 ka) of populations associated with different cultures in Guangxi, Yunnan, Fujian, Shandong, Xinjiang and Tibetan Plateau to detailed untangle their migrations, replacements or interactions, as well as to reveal the origins of Austronesian and Austroasiatic speakers, Ryukyu Islanders and the Tibetans' "ghost ancestors". Drawing on these findings, this report will highlight a unique genetic history of East Asian populations.

Graduate Student Excellence symposium

Sex-biased duplicates are generated during metazoan tRNA repertoire expansion

Dylan Sosa

University of Chicago

Abstract Eukaryotic genomes encode hundreds of transfer RNA genes ostensibly to ensure efficient translation. How these seemingly redundant tRNA copies arise and are naturally selected in eukaryotic genomes and populations remains largely unexplored. Moreover, how recently evolved protein coding genes are regulated and influenced by tRNA repertoires is not well understood. In this work we systematically investigate the evolutionary and population genetics of the tRNA gene family expansion through comparative analysis of 24 Drosophilid species. We show that Drosophila tRNA genes have been continuously lost and gained throughout 60 million years of divergence leading to rapid taxonomic restriction of specific isodecoding and isoaccepting copies. By gene age dating analysis we infer that 1.03% of the D. melanogaster tRNA repertoire have originated within the last 10 million years with one copy being species specific. We assess the regulatory potential of these new tRNA genes through small RNA sequencing of D. melanogaster reproductive tissue and report age- and sex-dependent expression, fragmentation, and base modifications of distinct tRNA gene copies implicating the emergence of taxonomically restricted small non-coding genes as sources of phenotypic or functional diversity. Through targeted gene knock-out of young tRNA genes we investigate their contributions to Darwinian fitness and the phenotypic consequences of their ablation. Our findings provide evidence for translational conflict as a feature of gene origination, that the evolutionary process of eukaryotic tRNA gene repertoires is characterized by continuous changes in membership, and that the expansion of tRNA genes in eukaryotes is not generally driven by adaptive codon usage.

Disentangling Complex Histories of Hybridisation: The Genomic Consequences of Ancient and Recent Introgression in Channel Island Monkeyflowers

Aidan Short

University of South Carolina

Abstract The accumulation of genetic differences through time can lead to reproductive isolation between populations and the origin of new species. However, hybridization between emerging species can occur at any point before isolation is complete. The evolutionary consequences of this hybridization may vary depending on when it occurred. If hybridization occurred later during the process, when ecological and genetic differences have accumulated between diverging lineages, low hybrid fitness can result in selection against gene flow. If hybridization occurred earlier, when barriers present were too weak to limit introgression, then hybridization can lead to genetic swamping. Alternatively, adaptive introgression can occur at any point during speciation. Thus, by understanding the history and genomic consequences of hybridization at different points along the speciation continuum, we can begin to understand how variation present within populations translates to divergence between species. Here, we identified the genomic signals of introgressive hybridization at different points during the divergence of two monkeyflower taxa endemic to the Channel Islands of California. We found that both ancient and recent hybridization have shaped their genomes, but the impacts of selection on introgressed alleles varied. There was no signal of selection against ancient introgression, but we did find evidence for selection against recent introgression, potentially because there are more reproductive barriers in place now, reducing fitness in recent hybrids. Thus, this study reveals that hybridization can occur at multiple points throughout the divergence history of a radiation, but the processes shaping genome wide levels of introgression can change over time.

The march to land: unearthing how gene repertoire evolution triggered terrestrialization across the Animal Tree of Life

Gemma Martínez-Redondo

Institute of Evolutionary Biology (CSIC-UPF)

Abstract The transition from sea to land—terrestrialization—is one of the most extraordinary evolutionary events that has shaped Earth's biodiversity. This monumental shift required animals to overcome environmental challenges and adapt to life on land through profound biological changes. Terrestrialization occurred independently in nine animal phyla, often with multiple transitions, offering a unique window into the genomic basis of extreme convergent evolution. Despite its evolutionary importance, the genomic underpinnings of terrestrialization across the animal kingdom remain largely uncharted. Through the lens of comparative phylogenomics and machine learning, we analyzed the gene repertoire evolutionary dynamics of ~1,000 high-quality metazoan genomes and transcriptomes (~24 million genes), spanning virtually all animal phyla. Our findings reveal that gene loss was ten times more prevalent than gene gain in independently terrestrialized lineages, suggesting that adapting to the specific environmental pressures or constraints associated with life on land more often involved the reduction or removal of no longer advantageous genetic elements, rather than the acquisition of new genes. We also uncovered a dynamic 'terrestrialization genetic toolkit' formed by genes of ancient origin that were co-opted in parallel for key terrestrial functions, and the convergent emergence of new genes across different lineages. These adaptations included mechanisms for desiccation tolerance, structural support, odorant reception, DNA repair, xenobiotic processing, and nutrient acquisition. Our study provides unprecedented insights into the genomic foundations of terrestrialization, revealing how ancient genetic resources and convergent innovations have collectively shaped the remarkable evolutionary trajectories that allowed animals to conquer land.

A fresh approach for estimating selection on synonymous codons

Hannah Verdonk

Temple University

Abstract Selection on synonymous codon usage is a well-known and widespread phenomenon, yet existing evolutionary models often do not account for it or its effect on synonymous substitution rates. Multiclass Synonymous Substitution (MSS) models are codon evolutionary models that account for such selection by partitioning synonymous substitutions into two or more classes and estimating a relative substitution rate for each class, while accounting for important confounders like mutation bias. We use several MSS models to identify extensive heterogeneity among relative synonymous substitution rates in an empirical dataset of ~12,000 gene alignments from twelve Drosophila species. MSS rates are significantly correlated with other covariates of selection on codon usage such as population-level polymorphism data and tRNA abundance data, suggesting that MSS models can detect weak signatures of selection on codon usage. With MSS models, we can now better study selection on synonymous substitutions in diverse taxa, independent of any a priori assumptions about the forces driving that selection.

Estimating effective population size from ARGs: An EM-based approach incorporating coalescent time uncertainty

Kaiyuan Li

UC Berkeley

Abstract New methods for estimating Ancestral Recombination Graphs (ARGs) are promising to transform population genetics, and the inference of demographic parameters, by providing access to full likelihood functions for large whole-genome datasets. However, fully harnessing this potential requires the development of statistical methods that can effectively utilize ARGs while addressing the inherent uncertainty in coalescent time estimation - a critical factor for accurate population size inference. Simulation studies reveal that existing methods often are biased due to the misplacement of coalescence events in adjacent time intervals. To overcome these limitations, we introduce a novel method for inferring effective population sizes by integrating ARGs with an Expectation-Maximization (EM) algorithm, which explicitly incorporates coalescent time uncertainty into the estimation framework. Using msprime to simulate large sequencing datasets and SINGER and POLYGON to infer ARGs, we demonstrate that our new method provides accurate estimates of effective population size over time, effectively correcting biases observed in existing methods. This approach exhibits superior statistical properties in terms of Mean Square Error in the estimation of effective population size, establishing the method as a powerful and reliable tool.

How museomics can be used to reveal whether insect genomes experienced recent anthropogenic changes

Philipp Hummer

Vienna Graduate School of Population Genetics

Abstract Recently, it has become possible to sequence the genomes of historic museum specimens of small insects that are up to 200 years old. Such whole genome sequence data represents a great resource that can be used to tackle a diverse range of questions concerning the more recent evolutionary history of insect species. One such example is the question of whether horizontal transfer of transposable elements (TEs) has recently occurred. In D. melanogaster, this has happened a total of eleven times in 200 years.

Such a high rate is very unusual and is hypothesized to be caused by anthropogenic shifts in species distributions. It is thus plausible that recent cosmopolitans may display similarly inflated rates of TE invasions. We want to investigate this hypothesis by sequencing and investigating the genomes of historical specimens from a range of different species.

For this purpose, we have developed a protocol for non-destructive DNA extraction from whole pinned specimens that causes only minimal physical changes. The possibility of maintaining the physical voucher does not only help to make more specimens accessible but also enables an integration of morphology and genetics. Additionally, extracting from the whole specimen also allows for the retrieval of DNA from endosymbionts such as Wolbachia spp. At this point we have established collaborations with three institutions – the Natural History Museum of Vienna, the Smithsonian National Museum of Natural History and the lower-Austrian monastery Stift Seitenstetten – and extracted DNA from 86 individuals across ten species.

Insights into the development of multiple sperm types and evolving sex chromosomes in drosophila species

Huangyi He

Zhejiang University

Abstract Sperm heteromorphism is widely distributed throughout the animal kingdom, but its evolutionary significance and developmental mechanisms are unclear. Drosophila obscura group species produce two or three types of sperms, with one or two types of infertile parasperms protecting the fertile eusperms against the harsh environment in the female reproductive tract. Here we performed single-cell RNA-sequencing of larva and adult testes of Drosophila pseudoobscura whose male produce three types of sperm, and compared to the published testes scRNA-seq data of Drosophila melanogaster whose male only produce one type of sperm. We showed that the spermatogonia of D.pse have developed into two types of early spermatocytes, then the parasperm early spermatocytes further developed into two types of parasperm. Duplicated and de novo genes are enriched in differentially expressed genes between eusperm and parasperm and could play important role in the new sperm type. During the spermatogenesis, both ancient X chromosome and neo-X chromosome of D.pse might have meiotic sex chromosome inactivation and dosage compensation. We found new gene transport cases supporting MSCI and demasculinization of X chromosomes. On the ancient X chromosome, de novo genes are enriched, exhibited higher expression compared to other chromosomes and increased expression during meiosis, suggesting MSCI may not regulate new genes. Overall, we uncovered the developmental mechanisms underlying the different types of sperms of Drosophila species, and also the expression of evolving sex chromosomes, especially new genes in the male germline.

Coupling AI and SMC-based algorithms: Inference of population structure from single genome sequencing analysis

Alba Nieto

MNHN-PSL

Abstract The increasing availability of whole genome sequences has significantly advanced our ability to reconstruct demographic histories and evolutionary processes with greater resolution. These data have facilitated the application of powerful inference methods, many of which rely on Sequentially Markovian Coalescent (SMC) models to infer past population dynamics. However, while these models have been widely used to estimate changes in effective population size over time, their accuracy and reliability can be compromised when applied to populations with complex structures. In particular, structured populations introduce deviations from the assumptions of panmixia, potentially leading to biased inferences and misinterpretations of demographic history.

In this study, we systematically assess the impact of population substructure on SMC-based inference methods and explore how these biases arise under different demographic scenarios. To achieve this, we perform extensive coalescent simulations designed to capture the effects of structured demographic histories, enabling a detailed evaluation of the biases introduced by population stratification.

Our results demonstrate that the magnitude and direction of these biases depend on factors such as the sampling scheme and key parameters governing the underlying population structure. Furthermore, we show that demographic interpretations based on SMC-inferred coalescent rates are inherently constrained by the model's assumptions, which often overlook deviations from panmixia. These findings emphasize the need for caution when applying SMC-based methods to structured populations and highlight the importance of developing refined approaches for more accurate demographic inference.

Symposia

Symposium 1: Crop evolution: from Darwin's legacy to genomics of admixture and climate resilience

Plant physiological models: what are they and can they be used to help address questions evolutionary in nature?

Diane Wang

Purdue University

Abstract Crop species, along with their wild and weedy relatives, are equipped with an arsenal of strategies to cope with environmental variability. These strategies have different effects on survival and productivity of individuals that are influenced by temporal sequences of abiotic factors and their interactions with genetically-conditioned plant traits. Research in the Wang Lab aims to uncover the processes that give rise to genotype-by-environment interaction in plant species of agricultural significance using various approaches including the iterative development, testing, and evaluation of process-based physiological models across genetically diverse germplasm. This talk will present a primer on plant physiological models for the SMBE community, including recent developments from the lab on a more parsimonious generic representation of plant growth, with the goal of motivating new opportunities for collaboration to connect genetic adaptation with physiological processes.

Evolutionary genomics of climatic adaptation and resilience to climate change in alfalfa

Fan Zhang

Agricultural Genomics Institute at Shenzhen, CAAS

Abstract Given the escalating impact of climate change on agriculture and food security, gaining insights into the evolutionary dynamics of climatic adaptation and uncovering climate-adapted variation can empower the breeding of climate-resilient crops to face future climate change. Alfalfa (Medicago sativa subsp. sativa), the queen of forages, shows remarkable adaptability across diverse global environments, making it an excellent model for investigating species responses to climate change. In this study, we performed population genomic analyses using genome resequencing data from 702 accessions of 24 Medicago species to unravel alfalfa's climatic adaptation and genetic susceptibility to future climate change. We found that interspecific genetic exchange has contributed to the gene pool of alfalfa, particularly enriching defense and stress-response genes. Intersubspecific introgression between M. sativa subsp. falcata (subsp. falcata) and alfalfa not only aids alfalfa's climatic adaptation but also introduces genetic burden. A total of 1671 genes were associated with climatic adaptation, and 5.7% of them were introgressions from subsp. falcata. By integrating climate-associated variants and climate data, we identified populations that are vulnerable to future climate change, particularly in higher latitudes of the Northern Hemisphere. These findings serve as a clarion call for targeted conservation initiatives and breeding efforts. We also identified pre-adaptive populations that demonstrate heightened resilience to climate fluctuations, illuminating a pathway for future breeding strategies. Collectively, this study enhances our understanding about the local adaptation mechanisms of alfalfa and facilitates the breeding of climate-resilient alfalfa cultivars, contributing to effective agricultural strategies for facing future climate change.

A haplotype-based evolutionary history of barley domestication

Yu Guo

Leibniz Institute of Plant Genetics and Crop Plant Research

Abstract Barley is an old crop with a complex history. Its evolution and the molecular basis of domestication have been intensely studied. This research has ruled out a single origin of the crop and motivated a model of mosaic genomics ancestry. As more and better genome sequences have become available, this concept can be refined: where do the building blocks of the mosaic, that is, the haplotypes come from? Were all cultivated barleys created equal or did some wild barley populations contribute more to some domesticated lineages? To answer these and other questions, we pursue a haplotype-based approach to connect diversity and population structure in wild and domesticated barley. We sequence the genomes of 367 genebank accessions. Using these data, we infer the spatiotemporal origins of haplotypes and map the contributions of different wild barley populations either during the initial phase of domestication or through later gene flow. Our results indicate that an early domesticated founder population formed in the Fertile Crescent during an extended period of pre-domestication cultivation.

Molecular adaptation and the geographic spread of japonica rice in Asia

Ornob Alam

New York University

Abstract Rice (Oryza sativa), one of the world' s most important food crops, was first domesticated beginning ~9,000 years ago in China as subspecies japonica. Using whole genome re-sequencing data from 456 japonica landraces, including 47 herbarium specimens collected over the last 150 years, we reconstruct how japonica rice moved from its center of origin to other parts of Asia beginning ~4,000-4,500 years ago. We observe an enrichment of pathogen resistance loci in selective sweeps associated with distinct geographic populations, suggesting that biotic interactions may be a key driver of local adaptation. We also find that the majority of variants in a set of 70 functional mutations associated with japonica rice domestication and diversification, including those that may be important for modern breeding, had their origins in the Pleistocene >11,700 years ago. Flowering time mutations, however, appear to have originated during the Holocene in the period during which rice evolved. Around 41% of these mutations have significantly biased geographic distributions that may be associated with local adaptation. This genomic analysis provides a model for how this important global food crop diversified after its domestication.

Describing wild Brassica rapa accessions and their genetic variation across Eurasia

Uliana Kolesnikova

MPI for Plant Breeding Research

Abstract Brassica rapa is a widely cultivated species with many cultivar types. It originated in Central Asia and was domesticated several times. However, little is known about wild B. rapa and its distribution. A previous study described a few wild accessions from Italy, Caucasus and Siberia, whereas weedy accessions from Central and Northern Europe and America were confirmed to be feral.

To identify further wild accessions, we sequenced 100 herbarium samples of weedy B. rapa across Eurasia. Upon clustering whole-genome resequencing data, we identified 66 accessions that are genetically distant from all cultivars, which we classified as wild. This classification was then supported by admixture analysis, as wild accessions form their own unique cluster. Wild accessions have a wide geographical range across Eurasia, grow in various climatic conditions, and are genetically more diverse compared to the cultivars. Therefore, they can be used to study environmental adaptation in B. rapa and potentially serve as genetic resources to improve Brassica cultivars.

We ran GWAS for bioclimatic variables on wild B. rapa and identified several candidate loci that may confer adaptation to different environmental conditions. Following up on this, we will look for genomic signatures of selection and introgression from wild to local cultivated accessions and to related polyploid species B. napus and B. juncea.

Our study provides results on genetic variation in wild B. rapa and presents a dataset with highly diverse wild accessions that can be useful for studying natural adaptation and breeding of B. rapa cultivars and related polyploid Brassica species.

BACK to the WILD: deciphering the relaxed selection role during rice domestication

Yesenia Madrigal Bedoya

New York University

Abstract Relaxed selection is a reduction in the efficiency of both purifying and positive selection. As a result, it could produce different functional gene evolution changes, including gene loss or pseudogenization. It is known that during crop domestication there have had modifications in wild plant species by selecting favorable genetic variants to improve their productivity and nutrition. These causative mutations (i.e. SNPs, small indels and others), resulted in the dysfunction and/or alteration in the expression patterns of domestication-related genes and have been selected to improve crops. Loss of function mutations have been considered an important source of adaptive genetic variation, especially in domesticated species. However, no one have evaluated a "wild species" perspective of the relaxed selection and its relationship with the pseudogenization process. We used Oryza rufipogon (wild) and O. sativa subsp. japonica (domesticated rice) to understand genes that are needed to be wild. We compared the complete genomic data available for wild and domesticated rice populations and looked for pseudogenes and signals of relaxation of selection between homologous genes. We perform different population genomics strategies to annotate, classify pseudogenes and filter them by the strength of purifying selection. Our results let us to propose a set of candidate genes important to be wild in Oryza, and with transcriptomic validation will establish the genomic variation associated with relaxed selection during rice domestication.

Untangling the complex domestication of potato: A LD-based method for polyploid genomics

Sergio Tusso

LMU Munich

Abstract Polyploid species present unique challenges for genomic analysis, particularly in detecting introgression events. Traditional methods struggle to disentangle ancestry in tetraploid genomes, limiting our ability to trace evolutionary history. Here, we introduce a simple linkage-based approach using short-read sequencing to infer ancestry and identify introgressed regions in polyploids, using cultivated potato (Solanum tuberosum) as a case study. Potato domestication is particularly complex, involving admixture from multiple wild species spanning a broad geographical range from Mexico to Argentina. Introgression occurred both during initial domestication in the Andean region and through later breeding efforts. To resolve this history, we applied a window-based linkage disequilibrium (LD) approach to a dataset of 840 resequenced samples, including globally distributed cultivars and 84 wild species. By identifying clusters of correlated variants, we inferred ancestry blocks within cultivars and traced their origins to wild species. Validation using phased tetraploid assemblies confirmed strong concordance between LD-based and genetic distance-based ancestry inference. Results reveal extensive but heterogeneous introgression across cultivars and chromosomes, with some wild species contributing disproportionately. Despite this large admixture, haplotype diversity remains unexpectedly low, with introgressed blocks clustering into a limited number of haplotypes. Moreover, we find a strong correlation between genetic diversity and inferred ancestry, emphasizing the role of introgression in shaping potato's genomic landscape. This approach opens new avenues for investigating polyploid evolution and optimizing breeding strategies in agriculturally important species.

Archaeogenomics of maize evolution in the South American Andes

Shuya Zhang

The University of Warwick

Abstract Maize, a staple crop, evolved from the wild plant teosinte through human activities in Mexico approximately 9,000 years ago. Around 7,500 years ago, it was brought through Central America, and by 6,500 years ago, it had spread into South America.

While the timeline of maize's spread is well-documented, there remains debate about whether maize underwent complete domestication before reaching South America and how it transitioned from lowland Mexico to the high-altitude Andes.

To address this, we sequenced 137 archaeological genomes and 11 maize landrace genomes from the South American Andes. These genomes were analyzed alongside 600 previously published Zea mays genomes, with Zea luxurians and Zea diploperennis as outgroups. Additionally, we employed radiocarbon dating to determine the ages of the samples and used strontium isotope analysis to verify the geographical origins of ancient maize specimens.

Our findings provide insights into the evolution of maize between Andean cultures prior to subsequent waves of maize movement through South America and reveal an early diversification of lowland and highland genomes shedding light on how maize adapted to the high-altitude environments of the Andes.

Simulation-informed population-genomic analyses reveal the Western-Himalayan origin of the common walnut

Caijin Chen

Beijing Normal University

Abstract The common walnut, Juglans regia, has long been thought to originate from the Irano-Anatolian region in Central Asia, while oldest nut fossils come from Armenia, Kashmir, and Pakistan. We revisited the question of walnut origin with targeted sampling focusing on the Western Himalayan region in South Asia (especially Afghanistan, India, Nepal, Pakistan), with 28 newly re-sequenced and 371 published nuclear genomes, chloroplast haplotypes, and simulations exploring the effects of uneven population sizes and sampling on the inference of ancestral source populations. The combined results reveal that South Asian trees represent a distinct population despite exhibiting admixture, harboring the highest chloroplast haplotype diversity, heterozygosity, private allele frequency, and a low genetic load. Over the past 200 generations (~10,000 years), the effective population size of the South Asian source population seems to have been large and stable, while East Asian and European populations suffered severe bottleneck effects, consistent with their spread from the South Asia towards Central Asia and West Asia. Our simulations applied to different data types reveal that earlier inferences of J. regia originating in the Irano-Anatolian region likely were due to the inability of the STRUCTURE software to identify source populations when faced with imbalanced geographic sampling.

The other side of domestication: an arable weed as key model of adaptation

Célia Neto

University of Copenhagen

Abstract In parallel with crops, agricultural weeds are excellent examples of evolution under strong anthropogenic selection, making them ideal models for studying adaptation to agroecosystems. The evolution of weediness follows an opposing trajectory to domestication, offering unique insights into the genetic basis of rapid adaptation to human-influenced environments.

In northern and western Europe, blackgrass (Alopecurus myosuroides) has become a major arable weed, threatening crop yields. To understand its transition from wild to weed, we reconstructed its spread across European agricultural landscapes using contemporary agricultural populations and historical wild germplasm predating modern agricultural intensification.

We traced blackgrass history in space and time, identifying the Black Sea region as its native origin and following its northward expansion across Europe, via the Iberia Peninsula. Non-native weedy populations are more recent and genetically uniform, with reduced diversity, whereas past wild populations were larger and highly diverse. Our findings further indicate that weediness in blackgrass has a polygenic genetic architecture, with strong selection acting on genes related to tropism, growth, organ development, reproduction, and environmental response. These traits likely enabled blackgrass to thrive in disturbed environments, shaping a “weediness syndrome” with parallels to crop domestication.

This study reveals the genetic basis of adaptation to new environments and the evolutionary responses to anthropogenic pressures, highlighting polygenic adaptation’s role in weed evolution and its contribution to plant invasions and persistence in agroecosystems.

Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweet potato

Mengxiao Yan

Shanghai Botanical Garden

Abstract The hexaploid sweetpotato (Ipomoea batatas (L.) Lam.) is one of the most important root crops worldwide. However, its genetic origin remains controversial, and its domestication history remains unknown. In this study, we employed a range of genetic evidence and a newly developed haplotype-based phylogenetic analysis to identify two probable progenitors of sweetpotato. The diploid progenitor is likely closely related to I. aequatoriensis and has contributed the B1 subgenome, the IbT-DNA2 and the lineage 1 type of chloroplast genome to sweetpotato. On the other hand, the tetraploid progenitor of sweetpotato is most likely I. batatas 4x, which has donated the B2 subgenome, IbT-DNA1 and the lineage 2 type of chloroplast genome. Sweetpotato most likely originated from reciprocal crosses between the diploid and the tetraploid progenitor, followed by a subsequent whole genome duplication. Additionally, we detected biased gene exchanges between subgenomes; the rate of B1 to B2 subgenome conversions was nearly three times higher than the B2 to B1 subgenome conversions. Our analyses reveal that the genes involved in storage root formation, maintenance of genome stability, biotic resistance, sugar transport, and potassium uptake have been selected during the speciation and domestication of sweetpotato. This study sheds lights on the evolution of sweetpotato and paves the way for the improvement of the crop.

Symposium 2: Evolution and regulation of sex chromosomes and sex determination

Stepwise recombination suppression on fungal mating-type chromosomes and evolutionary dynamics of degeneration

Tatiana Giraud(tatiana.giraud@universite-paris-saclay.fr)

University of Paris Saclay-CNRS

Abstract Sex chromosomes and mating-type chromosomes can display large genomic regions without recombination. Recombination suppression often extended stepwise with time away from the sex- or mating-type-determining genes, generating evolutionary strata of differentiation between alternative sex or mating-type chromosomes. In anther-smut fungi of the Microbotryum genus, recombination suppression evolved repeatedly, and extended stepwise repeatedly. We found a dozen of independent events of linkage between the mating-type loci across the Microbotryum genus, and two independent events of linkage of mating-type loci to centromeres. In addition, we found dozens of independent events of stepwise recombination suppression away from mating-type loci, trapping different gene sets the Microbotryum genus, but also in other fungal genera, such as Neurospora, Podospora and Schizothecium. We further show that genomic degeneration occurred following recombination suppression. Leveraging on the 22 independent events of recombination suppression with ages ranging from 0 to 4 million years ago, we elucidate the dynamics of degeneration with time, in terms of genomic rearrangements, non-synonymous substitutions, transposable element accumulation and non-optimal codon usage. We further identified loss-of-function at mating-type genes in some lineages.

Complex evolution of sex chromosome within a haploid-dominant liverwort lineage

Bo Xu(boxu@ibcas.ac.cn)

Institute of Botany, Chinese Academy of Science

Abstract Sex chromosome have evolved independently in numerous lineages across the tree of life, in both diploid-dominant species, including many animals and plants, and the less studied haploid-dominant plants and algae. Strict genetic sex determination ensures that individuals reproduce by outcrossing. However, species with separate sexes (termed dioecy in diploid plants, and dioicy in haploid plants) may sometimes evolve different sex systems, and become monoicous, with the ability to self-fertilize. Here, we study dioicy-monoicy transitions in the ancient liverwort haploid plant lineage, using three telomere-to-telomere gapless chromosome-scale reference genome assemblies from the Ricciaceae group of Marchantiales. Ancestral liverworts are believed to have been dioicous, with U and V chromosomes determining femaleness and maleness, respectively. We confirm the finding that monoicy in Ricciocarpos natans evolved from a dioicous ancestor, and most ancestrally U chromosomal genes have become autosomal in this species. We furtherdescribe evidence suggesting the re-evolution of dioicy in the genus Riccia, with probable de novo establishment of a sex chromosome from an autosome, and further translocations of genes from the new sex chromosome to autosomes. Our results also indicated that m-chromosomes are consistent genomic features, and may have evolved independently from sex chromosomes in Ricciocarpos and Riccia lineages.

The peculiar evolution of a novel sex chromosome in an androdioecious species with two mating types

Vincent Castric(vincent.castric@univ-lille.fr)

University of Lille-CNRS

Abstract Androdioecy is a reproductive system characterized by the coexistence of male individuals along side with hermaphrodites within a given species. The conditions under which it can evolve and remain stable are extremely stringent, and accordingly androdioecy is extremely uncommon. The only known exception to this overall rarity is the Oleaceae family of plants, hosting more than half of all currently ascertained androdioecious species across the whole living world. A solution to this puzzle came from the discovery in this family of a highly conserved self-incompatibility system with only two cross-compatible groups in hermaphrodites that the males are able to by-pass, providing them an immediate reproductive advantage that can compensate for the loss of their female function. By combining genetic mapping with the annotation of a high-quality genome assembly, we recently showed that the two-groups self-incompatibility system has a hemizygous architecture and is controlled by a gene involved in catabolism of Giberellin hormones. Contrary to the classical case of species composed of males and females, in androdioecious species the distinction is betwen males and hermaphrodites. This modifies the scope of sexual antagonism between the sexual morphs, and thus potentially the evolutionary trajectory of sex chromosomes. To adress these issues, we genetically mapped, assembled and fully annotated the sex chromosome in Phillyrea angustifolia, revealing a relatively small but highly re-arranged sex-determination region. We are now exploring how it evolved from its autosomal progenitor, and how it established a series of complex epistatic interactions with the region controlling self-incompatibility.

Whole genome PoolSeq analysis reveals the coexistence and origins of three distinct levels of X-Y chromosome differentiation in the European common frog

Wen-Juan Ma (wen-juan.ma@vub.be)

Vrije Universiteit Brussel

Abstract In contrast to the highly degenerated Y/W chromosomes of many model organisms, reptiles, fish, and amphibians often exhibit undifferentiated sex chromosomes and frequent sex chromosome turnovers. The reasons behind the prevalence of early-stage sex chromosomes and their rapid evolution remain poorly understood. The European common frog (Rana temporaria) provides an exceptional system to investigate these intriguing evolutionary dynamics. Previous studies using sex-linked microsatellites have identified multiple Y haplotypes within R. temporaria populations, each exhibiting varying degrees of X-Y differentiation. In this study, we employed whole-genome sequencing on pooled male samples representing three distinct Y haplotypes and XX females to:
i) characterize fine-scale X-Y differentiation,
ii) assess historical recombination events potentially linked to XY sex reversals, and
iii) investigate the evolutionary origins and stability of the multiple Y haplotypes within populations.

Our analyses revealed one of the smallest non-recombining regions documented on a Y chromosome, while other Y haplotypes displayed extended recombination suppression across most of the sex chromosome. Interestingly, some males showed no detectable differentiation from XX females. Despite the early signs of degeneration, such as transposable element accumulation and an increase in non-synonymous substitutions, no gene loss was observed. Gene trees of sex-determining and sex-linked loci further unveiled the evolutionary history of these diverse Y haplotypes.

This study offers novel insights into the earliest stages of sex chromosome differentiation and degeneration. It also highlights the remarkable coexistence of multiple Y haplotypes within a single population, providing a valuable framework for understanding the dynamic evolution of sex chromosomes in non-model organisms.

Novel expression of Xist in the male white-footed deermouse Peromyscus leucopus

Sha Sun(shasun@uci.edu)

University of California, Irvine

Abstract *Xist* (X-inactive specific transcript) is an essential long noncoding RNA for X-chromosome inactivation, ensuring equal gene dosage between the XX female and XY male in eutherian mammals. *Xist* expression is known to be female-specific – it is transcribed in all female somatic cells to inactivate one of the two X chromosomes, but it is not expressed in male somatic cells so that the single X chromosome remains active. In the white-footed deermouse *Peromyscus leucopus*, however, we find novel expression of *Xist* in the male. Unlike the canonical *Xist* expression in female somatic cells, which is detected by RNA fluorescent *in situ* hybridization as a distinct 'cloud' in the cell nucleus, *Xist* expression in male cells appears as small puncta. We validate the sex-differential expression of *Xist* in *P. leucopus* by RNA-sequencing and reverse-transcription quantitative real-time PCR. Furthermore, we identify colocalization of the *Xist* RNA with the repressive histone mark H3K27me3 in both male and female *P. leucopus* cells, suggesting a role for *Xist* in chromatin silencing in both sexes. Our study provides new insight into the function of *Xist* RNA and the evolution of dosage compensation in mammals.

pSTAT3 activation of Foxl2 initiates the female pathway underlying temperature-dependent sex determination

Pengfei Wu(wupengfei@ioz.ac.cn)

Fudan University

Abstract Ovarian development was traditionally recognized as a “default” sexual outcome and therefore received much less scientific attention than testis development. In turtles with temperature-dependent sex determination (TSD), how the female pathway is initiated to induce ovary development remains unknown. In this study, we have found that phosphorylation of the signal transducer and activator of transcription 3 (pSTAT3) and *Foxl2* exhibit temperature-dependent sexually dimorphic patterns and tempo-spatial coexpression in early embryos of the red-eared slider turtle (*Trachemys scripta elegans*). Inhibition of pSTAT3 at a female-producing temperature of 31 °C induces 64.7% female-to-male sex reversal, whereas activation of pSTAT3 at a male-producing temperature of 26 °C triggers 75.6% male-to-female sex reversal. In addition, pSTAT3 directly binds to the locus of the female sex-determining gene *Foxl2* and promotes *Foxl2* transcription. Overexpression or knockdown of *Foxl2* can rescue the sex reversal induced by inhibition or activation of pSTAT3. This study has established a direct genetic link between warm temperature-induced STAT3 phosphorylation and female pathway initiation in a TSD system, highlighting the critical role of pSTAT3 in the cross talk between female and male pathways.

Sex determination via X chromosome elimination in sciarid flies

Robert Baird(rbaird@wi.mit.edu)

Massachusetts Institute of Technology

Abstract Sciarids (fungus gnats) are a speciose family of flies with highly unorthodox sex determination. Unlike most organisms, in which sexual karyotypes are established upon fertilization, all sciarid zygotes begin genetically sexless, with three X chromosomes. Sex is decided during early embryogenesis when either one or two X chromosomes is eliminated, producing XX females or X0 males, respectively. The number of chromosomes to be eliminated in embryos depends on the genotype of the mothers, who possess supergenes that control X elimination and turn over rapidly. X chromosome elimination is caused by a failure of X chromosomes to properly segregate, but the precise cellular and molecular mechanisms, as well as the genetic basis of downstream sex determination, remain unknown. Bizarrely, chromosome elimination only happens in somatic cells, such that male and female germ cells have identical sexual karyotypes throughout development yet somehow still produce sperm and eggs. Using a combination of genomic, transcriptomic, and cytogenetic techniques, we aim to uncover the mechanisms underlying the strange form of sex determination in this clade.

Y chromosome evolution of East Asian populations

Jing Liu(liujing1993@zju.edu.cn)

Zhejiang University

Abstract The Y chromosome is crucial for determining human sex and male-specific traits, yet but its highly repetitive sequences have complicated evolutionary studies. Here, we present near telomere-to-telomere assemblies of 160 Y chromosomes representing major East-Asian haplogroups, and reveal genetic diversity and differentiation, particularly in complex regions. Phylogenetic analysis of male-specific euchromatin suggests widespread phylogenetic discordance across subregions due to large rearrangements, including lineage-specific inversions in ampliconic regions. We observed significant differentiation of structural variations in East Asians, potentially linked to spermatogenesis, immunity, and psychiatric traits. Further, The TSPY-FAM197Y gene cluster, the largest tandem-duplicated gene cluster in the human genome, originated in common ancestor of humans and gorillas and exhibits extensive copy number variation among haplogroups. Notably, the O2a haplogroup, prevalent in East Asians, harbors a fixed amino acid mutation in the TSPY2 gene. We also resolved the complex palindromic structure of ampliconic region7 and identified recurrent segmental duplication, deletion, and conversion events. Finally, gap-free assemblies within the largest heterochromatic region allow us to reveal its evolutionary complexity, with peripheral-originated repeats (DYZ1 and DYZ2) expanding in several wares toward repeat-array center in a manner resembling centromere formation, while modern humans exhibit substantial size variation due to the whole-block duplications and deletions. Transcript evidence of DYZ repeats and size association with prostate cancer challenged their classification as genomic "dark matter". These fully resolved East-Asian Y assemblies offer a unique resource for understanding Y-linked genetic differentiation and provide new insights into the evolution and function of complex genomic regions.

Relative rates of evolution on autosomes and sex chromosomes under paternal genome elimination

Thomas Hitchcock(thomas.hitchcock@riken.jp)

Interdisciplinary Theoretical and Mathematical Sciences – RIKEN

Abstract Sex chromosomes differ from autosomes in both their ploidy and transmission genetics. Consequently, selection, mutation, and drift may act differently upon them, driving distinct patterns in genetic divergence, diversity, and gene content. Recently, researchers have begun to consider a wider set of organisms with non-standard inheritance and sex-determination systems, however in many cases we lack theory which extends to such organisms. One such example is paternal genome elimination (PGE), a genetic system where males receive but do not transmit a paternal genome. Some groups with PGE also have X chromosomes. This includes fungus gnats, gall midges, and globular springtails. Consequently, in these groups, autosomes and X chromosomes exhibit the same transmission genetics, but with different somatic ploidy, making them uniquely suited to test hypotheses about the role of haploid selection in males. Here we develop a series of theoretical models tailored to the genetics of these groups, considering the relative rates of evolution across different regions of the genome. We then compare these analytical results to simulations, and as well as recently collected data in fungus gnats and gall midges.

Unity in sex chromosome evolution hidden by diversity

Deborah Charlesworth(deborah.charlesworth@ed.ac.uk)

Institute of Ecology and Evolution, University of Edinburgh

Abstract The diversity of sex determination and sex chromosomes is often emphasised, but here I will focus on some unifying evolutionary concepts and processes that apply to sex chromosomes that may appear, at first sight, to be very different. One important unifying concept is the apparently repeated (though not inevitable) evolution of suppressed recombination, creating completely sex-linked regions, and sometimes expanding such regions. This probably generally involves situations with multiple genes involved in sex determination, or genes whose mutations interact with the sex determining genes (so that recombinants would be inter-sexes or sterile individuals). Lack of recombination initiates the second unifying process, allowing evolution of the well-known sex chromosome properties of repeat richness and loss of genes (genetic degeneration), sometimes creating hemizygous regions that can be deleted, leading to cytologically visible heteromorphism. In turn, this may favor the evolution of dosage compensation, and accumulation on the male- or female-determining sex chromosome of mutations favouring the respective sex, even if these are deleterious in the other sex. All these properties are seen in the sex chromosomes of widely different organisms, which can have sex-linked regions of very different sizes and ages. The different ages help to understand a considerable portion of the observed variability. I illustrate these ideas with several case-studies showing that diverse types of mutations can cause male and female sterility in plants, and their properties can help understand sex chromosome turnovers, a major source of diversity in the sizes and ages of sex-linked regions.

Symposium 3: Experimental evolution: studying evolution in action

Real and not just hopeful monsters: Evolution by very large jumps

Dmitri Petrov(dmitristanford@gmail.com)

Stanford University

Abstract Evolution is often thought to be slow, proceeding through the accumulation of a very large number of very small steps. This is partly due to theoretical expectations that it should be hard to make large and adaptive changes in already well adapted organisms. Nonetheless, there is now abundant empirical evidence that large-effect adaptive mutations are common and appear in a range of biological systems. I will discuss our studies into the nature of these large effect mutations and will present a hypothesis that evolution of phenotypic plasticity and environmental regulation creates functional systems capable of large adaptive shifts.

Terrestrial herbivory drives real-time evolution in an aquatic community via indirect effects

Shuqing Xu(shuqing.xu@uni-mainz.de)

University of Mainz

Abstract Indirect ecological effects, which occur when the impact of one species on another is mediated by a third species or the shared environment, are ubiquitous in nature. Given the complexity of natural systems, indirect ecological effects were thought to be important in driving eco-evolutionary processes across community boundaries. However, we know remarkably little about such effects. Here, by performing large experimental evolution outdoors, we show that indirect effects of terrestrial insect (aphids) herbivory on macrophytes (duckweed) drives adaptive evolution of water fleas (Daphnia) in large outdoor aquatic mesocosms. Aphid herbivory reduced macrophyte growth and increased the abundance of phytoplankton, which in turn increased the abundance of Daphnia. Whole genome pool sequencing and phenotypic assays revealed an impact on the genetic compositions of the Daphnia populations and transplant experiments indicated that these evolutionary changes were adaptive. Furthermore, these changes in the aquatic community altered the interactions of the aphids and the macrophytes. These results demonstrate that indirect ecological effects can shape eco-evolutionary interactions between different communities.

Resolving the Aneuploidy Paradox by Experimental Evolution

Jing Li(lijing3@sysucc.org.cn)

Sun Yat-sen University

Abstract Aneuploidy, characterized by unbalanced chromosome numbers, commonly shows low fitness like slow growth and poor viability. Paradoxically, it is one cancer hallmark, contributing to tumorigenesis and treatment resistance. The mechanisms underlying this “aneuploidy paradox” are largely unknown. Here we investigated the evolutionary dynamics of genomic, transcriptomic and phenotypic alterations of a systematically constructed panel of yeast aneuploidies (1N+1 and 2N+1) under scenarios of neutral and adaptive evolution respectively. Across ~3000 generations of neutral evolution, almost all constructed extra chromosomes were stably inherited, rather than returning to euploidy given the reduced fitness. We identified several aneuploidy-tolerant de novo mutations that contribute to the stable evolutionary maintenance of aneuploidy. The 2N+1 aneuploidies revealed higher rates of chromosome copy number alterations and structural rearrangements compared to euploid diploid, suggesting the existence of initial extra chromosomes fostered chromosome instability. Dosage compensation, whereby transcript abundance is adjusted to compensate for altered gene copy number, is almost absent for chromosome gains but exists for chromosome losses. Under adaptive evolution, several constructed chromosome gains confer fitness advantages to specific stresses, which was further supported by recurrent acquisitions of the same chromosome in other populations initially not carrying this chromosome, indicating chromosome-specific advantages of aneuploidies in the selected environments. Moreover, aneuploidy drives the initially clonal populations to become more heterogeneous, providing abundant genetic variation for selection to act upon towards adaptation. Taken together, using distinct experimental evolution settings, we comprehensively mapped the evolutionary trajectories of aneuploidies and provided a multi-layered understanding of the molecular basis underlying aneuploidy paradox.

The question of epimutation: How does epigenetic information become heritable?

Maximilian Fitz-James(maximilian.fitz-james@biology.ox.ac.uk)

University of Oxford

Abstract As well as specifying differences between cell types within an individual, epigenetic information is also a major source of variation between individuals within a population. While this epigenetic variation is often restricted to a single generation, it can in some cases become heritable by the process of "epimutation", raising questions about the role of such non-genetic information in adaptation and evolution.
Epimutation may occur spontaneously, or be induced by a specific event, such as an environmental trigger, but the mechanisms and prevalence of these processes are not clear. In a modification of a classic genetic method in evolutionary biology, our work seeks to answer these questions through "epimutation accumulation" experiments. We propose to follow populations of Drosophila over generations and under different conditions, profiling their epigenetic signals and gene expression to map the appearance of new epimutations and track them across generations. This will provide an unbiased reading of the epimutation rate under both natural conditions (i.e. "spontaneous" epimutations) and in response to potential environmental triggers (i.e. "induced" epimutations) to determine the broader contribution of epigenetic information to heritable phenotypes and adaptation.

Experimental evolution of evolvability

Michael Barnett(barnett@evolbio.mpg.de)

Max Planck Institute for Evolutionary Biology

Abstract Evolvability—the capacity to generate adaptive variation—can itself evolve through natural selection. However, the idea that mutation can become biased toward adaptive outcomes remains controversial. Here we report the evolution of enhanced evolvability via localised hyper-mutation in experimental populations of bacteria. The evolved mechanism is analogous to the mutation-prone sequences of "contingency loci" seen in pathogenic bacteria. Central to this outcome was a lineage-level selection process, where lineage success depended on capacity to evolve between two phenotypic states optimal for growth under alternating conditions. Drawing upon detailed knowledge of the evolutionary history of lineages, we reveal how successive mutations interacted to construct the hyper-mutable locus, including changes to the genotype-phenotype map that focus selection to a single regulatory gene, followed by mutations that enhance transcription and, concomitantly, rates of localized frame-shifting. We additionally show the locus is itself evolvable with respect to alterations in the frequency of environmental change. Moreover, we reveal an unanticipated evolutionary advantage to localised hyper-mutation, finding that lineages possessing this ability are more likely to acquire additional adaptive mutations. Our results demonstrate how evolvability can itself evolve adaptively and the highlight the importance of this process for microbial pathogens.

Dissect cancer genome evolution during transient and stable drug resistance

Liye Zhang(zhangly@shanghaitech.edu.cn)

ShanghaiTech University

Abstract The drug resistance is one of major cause of mortality in cancer, as most cancer cells eventually acquire drug resistance. Most of studies on drug resistance focus on stable drug resistance state. However, recently more studies suggested that drug resistance is gradual and multi-stages process. Our previous work on yeast led us to believe that compared with more diverse stable resistance state, transient resistance solution is more common and thus becomes an easier and better drug-combo target. Therefore, we adopted hepatocellular carcinoma cell lines as our model to characterize the dynamic genome evolution during the resistance to targeted receptor tyrosine kinase inhibitor drug lenvatinib. Indeed, we identified two major stages of drug resistance: transient and stable stages. In addition, by combining multi-omics profiling, we identified and validated the key factor and pathway that confer transient resistance and explored potential intervention strategies. Next, when we tried to understand how stable clones were evolved by dissecting the clonal structures by scRNA-Seq and paired DNA-Seq. Surprisingly, we found that stable resistance cells were not derived from dominant transient resistance clone, but from a distinct lineage. To conclude, our study supported that multi-stage drug resistance is common. We also provided a proof of principal of targeting the transient drug resistance mechanism as a rational strategy to design drug-combo to reduce resistance rate.

Deficiency of DNA polymerase subunit Pol32 leads to a distinct pattern of genome instability

Daoqiong Zheng(zhengdaoqiong@zju.edu.cn)

Zhejiang University

Abstract Being a shared subunit of DNA polymerases δ and ζ, the role of Pol32 in maintaining genome intergrity was not fully clarified. This study explored the global genome instability resulted from POL32 deletion in yeast model. Whole-genome sequencing of isolates from a diploid pol32 mutant demonstrated that the absence of Pol32 increased the rates of SNVs, InDels, LOH, chromosomal rearrangements, and aneuploidy. Chromosomal instability in pol32 mutant was moderately aggrevated by low temperature. The deletion of Pol ζ catalytic subunit encoding gene REV3 resulted in a very diverse effect on genome integrity from pol32, indicating the genome instability in pol32 mutant was not due to defunction of Pol ζ. Most LOH events in pol32 mutant were caused by DNA breaks generated in S/G2 phase. Low expression of POL1 and deletion of POL32 synergistically induce replication stress and genomic instability. However, the loss of Pol32 significantly impairs the cell's ability to repair DNA breaks induced by H2O2. In addition to elevated frequency of genome instability, POL32 knockout also resulted in signatured patterns of genetic events. In addition to an elevated frequency of genome alterations, the POL32 knockout resulted in distinct patterns of genetic events. The relative ratio of AT-to-TA increased significantly in the pol32 mutant, an effect that was dependent on Pol ζ. Moreover, the absence of Pol32 created a specific hotspot for mitotic recombination on the right arm of chromosome VII. Overall, our findings provided novel insights into how a structural subunit of DNA polymerase contributes to maintaining genome integrity.

Rapid adaptation in globally distributed evolution experiment in Arabidopsis thaliana

Tatiana Bellagio(tbellg@stanford.edu)

Stanford University

Abstract Understanding how and whether populations will adapt to rapidly changing environments is a central question in evolutionary biology, yet insights into rapid evolution in eukaryotic macroorganisms remain limited due to the lack of large-scale realistic experiments. Here, I present GrENE-net, a three-year globally coordinated evolution experiment with Arabidopsis thaliana across 30+ outdoor gardens in the Levant, Europe, and North America, generating 2,415 whole-genome sequenced samples. We observed rapid evolution over time, with FST increasing to 0.024 by year three, as well as divergent evolution across space, evidenced by significant FST differences within versus across gardens, suggesting environment-driven adaptation. We also found evolutionary repeatability, as indicated by high correlations in allele frequency changes within gardens (mean Pearson correlation = 0.293), alongside signals of evolutionary rescue, reflected by U-shaped population dynamics in multiple locations. A positive and significant relationship between repeatability and population rescue parameters suggests adaptation-aided survival. Since observing climate-driven adaptation and parallelism in similar environments, we used three different genome-environment association models and identified 44 genomic regions under selection, revealing known and novel genetic variants linked to climate adaptation traits, such as germination (CYP707A1), flowering time (TSF), and stress responses (CAM5). Finally, GEA-informed genomic offset predictions achieved a 0.341 Spearman's correlation in forecasting evolutionary outcomes, as validated by leave-one-out cross-validation of the experimental site samples. Our findings provide insights into the genetic basis of rapid adaptation and showcase the potential of genomics and evolutionary biology to predict and mitigate the impacts of climate change on wild populations.

Experimentally evolving cellular miniaturization in S. cerevisiae

Marco Fumasoni(marco.fumasoni@gimm.pt)

Gulbenkian Institute for Molecular Medicine (GIMM)

Abstract Cell volume is a fundamental determinant of cellular physiology, maintained through a process known as cell size homeostasis. Studies have shown that large deviations from the typical cell size often produce harmful effects. Nonetheless, cell sizes have dramatically diverged during evolution, ranging from micrometers to millimeters. How does cell size homeostasis evolve to produce such diversity without impacting cell physiology? To address this paradox, we designed an experimental evolution strategy to select progressively smaller cells from Saccharomyces cerevisiae populations. Over 1500 generations, we achieved a six-fold reduction in cell volume compared to wild-type cells, marking the largest decrease observed. We show that the new size homeostasis is stably maintained. Strikingly, evolved cells maintain robust fitness, suggesting that the dramatic change in cell volume did not severely impact physiology. We investigated the complex genetic basis of cellular miniaturization through whole-genome sequencing and genetic engineering. The interplay between known and novel factors influencing cell size regulation through cell cycle control and cellular metabolism will be discussed. Our results demonstrate the evolutionary plasticity of cell size homeostasis and offer novel insights into how variability in a complex cellular feature can be evolutionarily achieved.

Gene expression plasticity promotes cell state transition and cancer cell adaptation under nutrient deprivation

Can Liu(liucan@mail.kiz.ac.cn)

Kunming Institute of Zoology

Abstract Plasticity, the ability of organisms to adapt through non-genetic phenotypic changes, plays a significant role in cancer evolution within tumor microenvironments. However, in evolutionary biology, plasticity can sometimes impede adaptation. Whether all forms of plastic transitions promote or hinder cancer cell adaptation remains largely unclear. In this study, we exposed the HeLa cell line with nutrient deprivation for 3 months to establish an experimental evolution model to systematically dissect the contribution of plasticity during cancer cell adaptation. Through bulk RNA sequencing (RNA-seq), we identified a surprisingly broad set of genes displaying non-adaptive plasticity, with their ultimate evolutionary changes contradicting their initial plastic gene expression patterns. Whole genome sequencing (WGS) further revealed distinct genetic divergence between adaptively and non-adaptively plastic genes. To investigate the potentially heterogeneous cell states, we conducted time-series single-cell RNA sequencing (scRNA-seq), and identified 11 cell clusters, two of which became adaptive as the nutrient deprivation progressed. Intriguingly, trajectory analysis revealed that ancestor cells initially transitioned into stress-responsive or dormant states before reversing their gene expression profiles to achieve the final adaptive states. This reversible cell state transition trajectory aligned with the non-adaptively plastic genes and was linked to changes in gene-to-gene and cell-to-cell interactions. Our findings provide a striking example of how non-adaptive plasticity contributes to cancer evolution.

Uncovering the correlations between cell size and growth phenotypes

Alexander Sastokas(alexandersastokas@gmail.com)

Arizona State University

Abstract Experimental evolutions allow extensive study of how environmental changes shape fitness-related phenotypes. However, few evolution experiments have explored how biophysical constraints shape phenotypic outcomes. To address this gap, we conducted a large-scale experimental evolution to investigate and push the biophysical limitations of cell size, in order to study how cell size constrains other phenotypes. We performed a massively parallel experimental evolution using a barcoded library of over 300,000 unique Saccharomyces cerevisiae isolates, subjecting cells to size-selection pressure. Through this evolution, we uncovered a diverse array of mutations associated with increased cell size, including changes in ploidy, chromosomal duplications, and single-nucleotide mutations. These mutations predominantly affect membrane transport, vacuole size, and mitochondrial function. Despite differences in the genetic mechanisms driving cell size increase, we observed a consistent trend: as cell size increased, cell density decreased. Extending our analysis beyond S. cerevisiae, we found that cell density appears to be correlated with cell size across the tree of life, reinforcing our observed trend whereby cell density decreases in larger cells.Our study demonstrates how experimental evolution can reveal biophysical constraints in cellular evolution, demonstrating the power of laboratory evolution studies in uncovering fundamental trends of cell and evolutionary biology.

Symposium 4: Gene trees and beyond: genealogical methods and their applications in evolutionary genetics

TBA

Rasmus Nielsen

UC Berkeley

A pandemic-scale ancestral recombination graph of SARS-CoV-2

Shing Zhan(shing.zhan@ndph.ox.ac.uk)

University of Oxford

Abstract Widely used phylogenetic methods to study SARS-CoV-2 evolution assume that genetic recombination is negligible and viral relationships can be represented by a single tree. However, important SARS-CoV-2 recombinant lineages have emerged. A more suitable approach is to use ancestral recombination graphs (ARG), which jointly model recombination, mutation, and genetic inheritance in a cohesive framework. Recent computational advances have made ARG inference possible for datasets of millions of genome sequences. Building on these new advances, we have developed sc2ts, a scalable method to simultaneously detect recombinants and reconstruct ARGs from heterochronous SARS-CoV-2 genome sequences. Using sc2ts, we inferred an ARG of 2.48 million genomes sampled from January, 2020 to February, 2023. We find that this pandemic-scale ARG unifies known features of SARS-CoV-2 evolution—the backbone phylogeny, mutational spectra, and major recognised recombinants. This ARG enables us to explore the evolutionary origins of known and novel recombinants. Furthermore, the ARG is compactly stored in the tree sequence format using the tskit library and can be analysed efficiently and interactively using standard Python tools on a typical laptop. This combination of capabilities makes sc2ts a powerful tool for studying the recombinational and mutational processes driving SARS-CoV-2 evolution.

CASTER: Direct species tree inference from whole-genome alignments

Chao Zhang（chao.zhang@sund.ku.dk）

University of Copenhagen

Abstract Genomes contain mosaics of discordant evolutionary histories, challenging the accurate inference of the tree of life. While genome-wide data are routinely used for discordance-aware phylogenomic analyses, due to modeling and scalability limitations, the current practice leaves out large chunks of genomes. As more high-quality genomes become available, we urgently need discordance-aware methods to infer the tree directly from a multiple genome alignment. Here, we introduce CASTER, a theoretically justified site-based method that eliminates the need to predefine recombination-free loci. CASTER is scalable to hundreds of mammalian whole genomes. We demonstrate the accuracy and scalability of CASTER in simulations that include recombination and apply CASTER to several biological datasets, showing that its per-site scores can reveal both biological and artefactual patterns of discordance across the genome.

Spatial temporal evolutionary trajectory of solid tumor cells reveals clonal diversity changes during tumor progression and acquired drug resistance

Yue Hao(yhao@tgen.org)

Translational Genomics Research Institute

Abstract Glioblastoma remains one of the most challenging and aggressive brain tumors to treat, exhibiting significant heterogeneity in both tumor cells and the microenvironment (TME). Such heterogeneity is the direct results of the clonal dynamics shaped by evolutionary forces. Osteosarcoma (bone cancer) undergoes extensive chromothripsis leading to clonal diversification. To study clonal diversity in solid tumors, our strategy is to first identify copy number variation (CNV) from the sequencing data, then use CRUST and PyClone to statistically deconvolute clonal populations and phylogenetically reconstruct tumor clonal lineages spatially or longitudinally. Such approach was applied to three different datasets. The first dataset contains biopsies from both IDH-mutant and wildtype gliomas with recorded spatial coordinates, CNVs were estimated using the shallow whole genome sequencing data, genetic distances between samples were calculated compared with the spatial physical distances. Secondly, we longitudinally tracked an engineered normal human astrocyte (NHA) cell line that experiences dramatic structural variations and obtained Illumina short-read sequencing for these cells at multiple time points. Our preliminary clonal deconvolution results from the NHA dataset revealed rapid turnover of different subclones, suggesting that genetic drift underlies dynamic shifts in clonal populations in tumor cells. Similarly, we infer clonal changes that enable resistance to doxorubicin or methotrexate in an osteosarcoma cell line using CNV estimated from ctDNA data. Understanding clonal dynamics will help us elucidate how cancer cells evolve, adapt, and interact with their environments, and thereby identify key drivers of tumor occurrence, progression, and resistance to discover novel therapeutic opportunities.

Host-genetic interactions with symbionts: A cophylogenetic approach using ancestral recombination graphs

Rowan Hart(rowanh@uchicago.edu)

University of Chicago

Abstract Cophylogeny, the study of phylogenetic similarity between interacting organisms, provides insights into the shared evolutionary history and strength of symbiotic interactions. Traditionally, cophylogenetic analyses have focused on macroevolutionary relationships between species, revealing evidence of long-term relationships in host-symbiont pairs. Recent studies of human-microbiome codiversification have extended cophylogenetic analyses to within-species relationships, uncovering microevolutionary patterns of shared population structure. However, many approaches use a single evolutionary tree to represent host ancestry, capturing genome-wide signals of cophylogeny but lacking the resolution to detect patterns confined to specific genomic regions. Here, we leverage recent methodological advancements to infer ancestral recombination graphs in the host. This approach allows us to disentangle genome-wide signals of shared population structure from locus-specific host-symbiont genetic interactions. We simulate panmictic and structured populations to assess the power of this approach to detect cophylogenetic signals and to examine how selection pressure for matching host-symbiont alleles affects these signals. We then apply this framework to two datasets: (1) the 1000 Genomes Project to examine cophylogeny between nuclear genomic variation and mitochondrial DNA, and (2) the Human Microbiome Project to identify human genetic loci interacting with genetic variation in bacterial symbionts. Enrichment analyses reveal significant cophylogenetic signals in genomic regions that are relevant for microbial function and host regulation of the microbiome.

SNP2ARG: Inferring ancestral recombination graphs with deep learning

Marçal Comajoan Cara(comajoan@stanford.edu)

UC Santa Cruz & Stanford University

Abstract We present SNP2ARG, a deep learning approach for inferring ancestral recombination graphs (ARGs) from genomic data. Existing methods rely on probabilistic frameworks and predefined evolutionary models, often approximating the full coalescence process with recombination (CwR) using the refined sequentially Markovian coalescent (SMC') model, despite CwR's known non-Markovian nature. We introduce a new approach that instead leverages the power of neural networks in order to model the full CwR process. By representing an ARG as a sequence of coalescent trees that change at recombination breakpoints along the genome, we reframe ARG inference as a sequence-to-sequence learning task. Our model takes SNP sequences from multiple samples as input and directly learns to approximate the CwR process through end-to-end training on simulated data. By using modern transformer architectures, SNP2ARG captures complex evolutionary patterns, with attention mechanisms accurately identifying SNP patterns across the genome. We train the model on a diverse set of simulated scenarios covering a wide range of demographic histories and selective pressures. Through experiments on both simulated and real human genomic data, we evaluate SNP2ARG-generated ARGs in several downstream tasks against other methods. We assess model performance across varying dataset sizes and genomic scales, measuring both inference accuracy and computational efficiency. The trained model can be readily applied to new datasets. Through our open-source implementation and pre-trained models, we hope to make deep learning-based ARG inference accessible to the broader population genetics community.

Biological causes and impacts of tree space complexity in phylodynamic inference with measurably evolving populations

Jiansi Gao(jsigao@fredhutch.org)

Fred Hutchinson Cancer Center

Abstract Phylodynamic analysis has been instrumental in elucidating the spread and evolution dynamics of pathogens and cells. The limited genetic diversity in phylodynamic datasets and their sheer size lead to significant phylogenetic uncertainty. Bayesian inference accommodates phylogenetic uncertainty by integrating over the tree space; it does not, however, scale with modern genomic datasets, partly due to the vast and complex tree space. Here, we set out to characterize the tree space and assess its impacts on analysis difficulties and key biological inferences. By running extensive Bayesian analyses of 20 large phylodynamic datasets and carefully analyzing the posterior samples, we find that: 1) the tree space is diffuse and rugged, and a lack of topological convergence is widespread; 2) difficulties in tree-space exploration frequently stem from a small set of sequences that might have been subject to biological processes (e.g., recombination) that violate common phylogenetic assumptions, and; 3) impacts of the tree sampling problem are minimal on treewise phylodynamic estimates but more pronounced on the clade-specific ones, such as origin time and spread history. In addition, we demonstrate that heterochronous sampling dates contain considerable information about tree topology, which can be in conflict with genetic data at local scale, leading to further complexity in the tree space and systematic discrepancies between Bayesian and the commonly used stepwise phylodynamic approaches. We offer strategies and develop new diagnostics to mitigate impacts of the tree sampling problem. Our findings highlight new directions for efficient integration over the tree space, ultimately advancing scalable and reliable phylodynamics.

Historical biogeography of potamid crabs in Southern China: a phylogenomic perspective

Zechang Cheng(chengzch3@mail2.sysu.edu.cn)

Sun Yat-Sen University

Abstract Southern China, with highly heterogeneous natural habitats, supports exceptionally high biodiversity, but the underlying mechanisms remain incompletely understood. China hosts the world's highest species diversity of potamid crabs, with 52 genera and 320 species and subspecies, of which over 50% are confined to Southern China. The potamid crabs' strong dependence on specific environments and their relatively limited dispersal abilities make them ideal models for unraveling the drivers of this region's extraordinary diversity. However, such endeavor is hampered by the lack of a robust phylogeny for potamids. In our study, we gathered approximately 200 species of potamid crabs from Southern China and neighboring areas, ensuring representation from all Chinese potamid crab genera. We employed hybrid capture sequencing using in-house cDNA probes, resulting in the acquisition of about 600 orthologous sequences for phylogenetic analysis. Our analyses produced a phylogeny with robust bootstrap support, although some nodes exhibited significant gene tree discordance, potentially due to incomplete lineage sorting. Our preliminary findings suggest that Southern China serves as a convergence zone for diverse potamid crab lineages, providing support for the "Faunal Overlap Hypothesis."

High density cell lineage trees revealed distinct features of somatic evolution

Jian-Rong Yang(yangjianrong@mail.sysu.edu.cn)

Sun Yat-sen University

Abstract Recent technological advancements have enabled simultaneous determination of transcriptomes and phylogenies of single cells by using DNA lineage barcodes. With the rapid accumulation of such data, it has become imperative to develop new methods for reconstructing phylogenies of cells and downstream analyses that extract general rules of somatic evolution. To achieve the latter goal, we obtained transcriptomes and phylogenies of 10 to 50% of the cells within a colony. Similar to dense taxon sampling facilitating key phylogenetic insights for the evolution of species, our analysis of such densely sampled cell phylogenies and the gold standard lineage tree from Caenorhabditis elegans revealed two crucial features of somatic evolution distinct from species evolution. First, our novel algorithm designed for quantitative comparison (alignment) between cell lineage trees suggests the existence of multiple stereotyped subtrees with similar terminal cell types and topological structures. We showed that this feature is likely created by transcriptional regulatory programs and contributed to developmental robustness. Second, phenotypic divergence between cells is jointly determined by regulation and inheritance. We demonstrated that this feature allows stratified analyses of cell pairs at different phylogenetic distances that reliably quantify transition rates between different cell states. Combined, these somatic evolution-specific features on the one hand provide valuable information about development and disease progression, but on the other hand compromise independence between lineages, a common null expectation in phylogenetic analyses for species evolution.

Joint inference of cell lineage trees, genotypes and phylodynamic parameters from single-cell sequencing data.

Yao Xiao(yxia415@aucklanduni.ac.nz)

The University of Auckland

Abstract With the development of single-cell sequencing technology, phylogenetics can be applied to cell and developmental biology. For example, during development, cell variations can be viewed as markers of evolution. Consequently, evolutionary models offer new perspectives for understanding somatic development and cancer evolution. However, most existing methods handle data filtering, variant calling, and phylogenetic inference separately. This workflow does not fully leverage the information in the data, potentially leading to biased results. To address these issues, we propose a Bayesian inference–based model that simultaneously performs variant calling and phylogenetic tree inference directly from single-cell DNA sequencing read counts. Our model has the following features: it considers all 16 possible diploid genotypes and accounts for both sequencing errors and allele dropout that may occur during sequencing. Our research aims to provide a powerful new tool for studying cancer evolution and developmental biology.

Analyse the causality relations among genes that are associated with intellectual disability in primate brains

Jing Qin(qin@imada.sdu.dk)

University of Southern Denmark

Abstract Intellectual disability (ID) is defined by an IQ under 70, in addition to deficits in two or more adaptive behaviors that affect everyday living. Throughout history, individuals with ID have often been marginalized from society and continue to suffer significantly even in modern times. A varying proportion of ID cases are attributable to genetic causes. Deciphering their gene expression pattern of ID-associated genes during brain development process would gain us a fundamental understanding of the molecular basis of ID. Furthermore, when comparing across primates, genes associated with ID also hold promise for understanding how intellectual abilities has evolved. Here, we used Granger causality analysis to establish the dependence relations among genes that are known to be associated with intellectual disability. We utilized publicly available RNA-Seq data from developing human and gorilla brain organoids for this analysis. Based on our causality analysis, we created gene expression networks among these genes for both species and carried out network comparison between these two networks. In addition, we extended our technique to computationally provide a list of new candidate genes that are most likely to be associated with ID. Our work reveals how ID genes are differentially integrated in the molecular networks active during human and gorilla brain development, which might underlie functional evolutionary differences of the human brain.

Symposium 5: Human history and past social organisation in the light of palaeogenomes: new methods, new findings session A

Ancient pathogens as proxies of human movements and social interactions

María Ávila Arcos (mavila@liigh.unam.mx)

National Autonomous University of Mexico

Abstract Pathogens and humans have co-evolved in an arms race of immunity and immune evasion since the origin of our species. The spatiotemporal study of these human-pathogen interactions through ancient DNA has yielded remarkable insights into the evolution of pathogenicity but also into pathogen introductions to new territories through population movements. A textbook example is the introduction of devastating infectious diseases during the European colonization of the Americas between the 15th and 16th centuries. In this talk, I will describe how we use paleogenomic approaches to characterize the diversity of pathogens circulating during the colonial period of Mexico City. By leveraging genetic ancestry characterization of the hosts and phylogenetic analysis of the pathogens, we shed light into the geographical origin of the pathogens introduced by Europeans during colonization. These studies also illustrate how the phylogenetic analysis specifically of oral bacterial paleogenomes can reveal close social interactions during the first encounters of individuals from different continents. Our results illustrate the rich information paleogenomics can provide for the study of past humans and their pathogens and prompt necessary discussions on the multifaceted consequences of European colonization.

Ancient DNA reveals genetic barrier despite shared Avar-period culture in Eastern-Central Europe

Ke Wang（echo.ke.wang@gmail.com）

Fudan University

Abstract For a decade, DNA recovered from ancient bones and teeth has provided insights into the identification of a new Homo species (Denisovan) and the long-ago worldwide population movements through time. Now it also starts revealing past population secrets and social structure. Here we studied two Early Medieval communities associated with the historicaly known Avars. Initally it was a group of nomads with East Asian ancestry who arrived in East Central Europe in 567-568 AD and encountered groups with very different European ancestry. Focusing on 722 individuals from two neighboring cemeteries south of Vienna, we aim to address the centuries-long impact of this encounter by directly comparing their genetic origins, archaeological cultural traits, and pedigrees. We found that these two nearby sites show genetic distinctions with little biological affinity, despite sharing a distinctive Avar culture. In particular, we reconstructed six-generation pedigrees at both sites, including up to 450 closely related individuals at one site, allowing generation-by-generation demographic profiling of the communities. Despite diverse ancestry, these pedigrees, together with large networks of distant relatedness, reveal an absence of consanguinity, a patrilineal pattern with female exogamy, multiple reproductive partnerships (e.g. levirate), and a direct correlation of biological connectivity with archaeological markers of social status. Overall, we conclude that the perception of people as 'Avars' in the historical sources, the unification of cultural habitus, and genetic admixture did not align, resulting in contrasting genetic ancestry in two local communities within a radius of 20 km.

Ancient mastics offer new insights into oral microbiome evolution and health

Anna White(annawhite@palaeome.org)

Globe Institute, University of Copenhagen

Abstract The human oral microbiome comprises around 700 taxa and plays a crucial role in human health. Changes in oral microbiome composition have been linked with oral diseases but also systemic diseases such as cancer, cardiovascular disease, and type 2 diabetes. The oral microbiome is influenced by various factors including our diet and oral hygiene practices, which can alter its composition, making it essential to study these influences to better understand its evolution. Until now, studies of the human oral microbiome have primarily relied on ancient dental calculus, or calcified dental plaque. However, it remains unclear how representative dental calculus is of the oral microbiome, as it reflects a long-term deposit of the microbiome that accumulates over a lifetime. In this study, we analysed DNA extracted from over 100 ancient mastics from archaeological sites in Europe to characterise the oral microbial composition of Mesolithic and Neolithic populations. Using this novel substrate, we characterised the oral microbiome and performed functional analyses on a community level. We found that the oral microbiomes derived from ancient mastics resemble modern salivary profiles more closely than those derived from ancient dental calculus in both diversity as well as functional potential. As the microbial composition of ancient mastics appears to be more comparable to modern oral microbiomes, this material opens up new possibilities for studying the evolution of the human oral microbiome and its impact on human oral and systemic health.

Ancient metagenomics reveal human-microbial interactions over an evolutionary timescale

Yichen Liu(yichen.liu@ivpp.ac.cn)

IVPP, CAS

Abstract The interaction with microbes has long been an indispensable piece in human history. Despite over 10,000 years history of utilizing and manipulating microorganisms to produce fermented food, little is known about the origin of the techniques and history of its development and spread. In addition to the scarcity of organic residues preserved over thousands of years, a major challenge is to obtain molecular evidence from the ancient remains. Here, through accessing several Bronze Age cheese samples from the Xiaohe Cemetery, we were able to interrogate the DNA information of probiotic bacteria that were regularly propagated and consumed by humans over thousands of years. We found that the Xiaohe population actively adopting techniques from western steppe populations and integrating them into their own culture. In addition, while kefir is a traditional fermented milk beverage that is still widely consumed today, the origin and spread of it is unclear. Through the reconstruction of the phylogeny of Lactobacillus kefiranofaciens – a key fermenter in kefir, we found there are two distinct spread routes in East Asia, with one primarily spread along the coastalline and the other spread within the mainland. We also identified dual-role gene clusters gained by Lactobacillus kefiranofaciens through horizontal gene transfer over their thousand years of interactions with humans. Through retrieving the first ancient DNA study of dairy residue, our study reveals the histroy of dairy fermentation by a Bronze Age population as well as dynamic interactions between humans and microbes over thousands of years.

Late Neandertals in focus: a fine-resolution paleogenetic study of Neandertal groups in North-Western Europe

Alba Bossoms Mesa(alba_bossoms@eva.mpg.de)

MPI - EVA

Abstract The analyses of Neandertal genomes are often limited by poor DNA preservation in Pleistocene remains. Much of what we know comes from a handful of individuals that do not fully represent the geographic and temporal distribution of Neandertals. Some of their genomes, such as those from Chagyrskaya and Mandrin, suggest that Neandertals lived in small, isolated groups with reduced genetic diversity. It is currently unclear whether this pattern applies to all Neandertals.

To explore the fine-scale genetic structure in late Western Neandertals, we analyzed 35 skeletal remains, dating to 39,950 - 49,800 years ago, from seven archaeological sites in the Meuse Basin (Belgium) and two sites in France. Of these, 25 specimens contained sufficient endogenous DNA to produce genome-wide data, including a new high-coverage genome (24-fold) from Goyet Q56-1.
The late Western Neandertals in this study had greater genetic diversity and less evidence for recent consanguinity than the Eastern Neandertals living ~120,000 - 60,000 years ago in the Altai mountains. Moreover, we detect no evidence of additional genetic substructure or close kinship among these late North-Western Neandertals. Interestingly, none of the analyzed genomes show signs of recent introgression from modern humans, although evidence for gene flow in the opposite direction has been widely reported. These findings highlight the variability in community structure across the Neandertals' geographic and temporal range.

Paternal Genetic Contributions from Millet and Rice Farmers to the Genetic Makeup of Southern Chinese and Southeast Asians

Yunhui Liu(gianthuihui@gmail.com)

Chongqing Medical University

Abstract The genetic diversity of southern Chinese and Southeast Asian populations has long intrigued researchers, yet their roles in early East Asian peopling, Holocene expansions, and paternal contributions to southern minorities and Southeast Asians remain underexplored. We present comprehensive paternal genomic datasets comprising 7,877 microarray-based genotyping variants and 2,605 whole Y chromosome sequences from ancient and contemporary Eurasian populations, including 2,636 males from 40 ethno-linguistically underrepresented groups. Admixture modeling reveals minimal genetic differentiation between northern and southern Han populations, underscoring notable genetic homogeneity. In contrast, ethnically diverse southern minority groups exhibit distinct paternal substructures, including coastal groups linked to Tai-Kadai speakers, southwestern populations represented by Hmong-Mien speakers and highland groups predominantly associated with Tibeto-Burman speakers. Haplogroup composition analyses tied to subsistence strategies highlight differential paternal contributions from northern millet farmers and southern rice farmers to East Asian and Southeast Asian populations. These findings emphasize southern China's central role in shaping paternal genetic diversity across these regions. Our results support the demic diffusion model, whereby ancient millet farmers and their Han Chinese descendants influenced both southern Han Chinese and indigenous groups in southern China and Southeast Asia. A strong paternal connection between southern China and Southeast Asia, mediated by rice farmer-related lineages, further elucidates the integration of these lineages into the Southeast Asian paternal gene pool. This study provides robust genetic evidence to inform future whole-genome sequencing and ancient DNA research, offering critical insights into the Neolithic contributions of early farming communities to the genetic landscape of East Asia.

The dynamic genetic landscape of indigenous South Americans

Marcos Araujo Castro e Silva

Universitat Pompeu Fabra

Abstract Indigenous American populations, despite their rich cultural and environmental diversity, remain understudied genetically, particularly regarding how historical, environmental, cultural, and evolutionary factors shaped their variation. To address this, we analyzed whole-genome sequences from 128 individuals representing 45 ethnic groups, 28 linguistic families, and 8 countries—the most diverse dataset assembled to date. We found minimal post-colonial European and African admixture and contemporary genetic structure correlates strongly with geography, with exceptions showing evidence of recent inbreeding.

Our analysis reveals four major genetic clusters in contemporary South America: Southern Cone, Western Lowlands, Eastern South America, and Chaco. Genetic differentiation among populations exceeds expectations under an isolation-by-distance model, likely influenced by isolation and heightened genetic drift, as evidenced by high homozygosity and intra-population genetic similarity. Integrating ancient and modern genomes, we identify remarkable genetic continuity across multiple regions, sometimes spanning from the Early Holocene to the present.

Additionally, we detect evidence of three major population dispersals into South America. The most recent, includes almost all present-day Indigenous South Americans and Ceramic period ancient Caribbeans and suggests a complex ancestry, combining early South American lineages with later Mesoamerican-related contributions and, in some cases, influence from Southern Cone populations.

Lastly, we refine our understanding of the enigmatic excess Australasian affinity inferred in some Indigenous groups, revealing evidence of a broader spatiotemporal presence and identifying genomic regions with signatures of positive selection, suggesting a potential adaptive role. This study sheds light on the complex interplay of demography, selection, and admixture in shaping Indigenous American genetic diversity.

Ancient genomes reveal complex genetic history of ancient populations in the western Tibetan Plateau

Fan Bai(baifan@ivpp.ac.cn)

IVPP, CAS

Abstract The western Tibetan Plateau is the crossroads connecting the Tibetan Plateau to Central Asia, South Asia and other regions, and may play an important role in the ancient human migrations. However, due to the harsh natural environment, coupled with the barrier of the Himalayas mountain, it is uncertain whether the cultural cross-regional integrity corresponds to interactions of populations, and it is impossible to answer this question by relying solely on traditional archaeological methods. In recent years, ancient DNA sequencing technology has developed rapidly, and provided important clues for understanding the origin and evolution of populations, interactions between different populations, and other issues. This report will introduce the first systematic study of the genetic characteristics of ancient populations in the Ngari region of the western Tibetan Plateau. This study newly reports the ancient human genetics data of 6 sites in the Ngari region. Combined with the data of 2 published sites, the age span of the sites ranges from 3,500 to 150 years ago. The results of this study have expanded our understanding of the changes in the genetic structure of ancient populations in the western Tibetan Plateau since 3,500 years ago, as well as the genetic connections and population exchanges between the ancient populations in the Ngari region and ancient populations in neighboring northern Pakistan, southern and central Tibet, and northern Nepal.

Improved palaeoproteomic methods allow for identification of a Pleistocene tooth from Khudji, Tajikistan, as a probable Neanderthal

Zandra Fagernäs(zandra.fagernas@sund.ku.dk)

Globe Institute, University of Copenhagen

Abstract When archaeological specimens lack morphologically diagnostic features, analysis of ancient biomolecules, such as DNA, may help determine taxonomic identity and phylogenetic relationships. However, for Pleistocene specimens, DNA preservation is often too poor for palaeogenetic analyses. Ancient proteins, on the other hand, have been found to preserve over much longer time scales, even millions of years, in mineralized tissues. Here, we apply palaeoproteomic analysis to a hominin tooth from the archaeological site of Khudji (Tajikistan), dated to approximately 40,000 years before present, which lacks diagnostic morphological features to determine if it stems from a modern human, a Neanderthal, or a Denisovan. Initial analyses showed heavy contamination by modern skin proteins. We therefore compare commonly used decontamination methods by artificially contaminating a Pleistocene bone fragment with modern dog saliva. After finding that a brief bleach (sodium hypochlorite) wash efficiently removes modern contamination without damaging the ancient proteome, we reconstruct the sequences of 12 proteins from the Khudji tooth dentine, covering over 3,000 amino acids. Phylogenetic analyses show that with currently available reference genomes, the Khudji hominin is most likely a Neanderthal. This identification contributes significant knowledge to the timing and distribution of Neanderthals in Central Asia, as they have thus far only been conclusively identified from a handful of archaeological sites. The results show both the applicability of palaeoproteomics to the study of the history of hominin taxa during the Pleistocene, as well as the importance of continued method development to uncover improved ancient proteomes.

Human adaptation to farming through DNA methylation

Youssef Tawfik(ytawfik@uni-mainz.de)

Johannes Gutenberg University of Mainz

Abstract The emergence of agriculture approximately 10.000 years ago marked the beginning of the Neolithic period when humans first transitioned from a hunting-gathering to a farming lifestyle, representing the major environmental transition in humans' evolutionary history. The field of palaeogenomics has produced and studied ancient human genomes, shedding light on the genetic origins of the world's first farmers and tracing the migration patterns and the genetic exchange between hunter-gatherers and early farmers. Recent methodological advances made it possible to infer DNA methylation patterns from post-mortem cytosine deamination that is characteristic in ancient DNA. As an epigenetic mechanism, DNA methylation mediates the interaction between genetics and environment which constitutes a fast adaptive response when facing new and challenging environments. By offering an additional layer of biological insight, it expands genomic analyses beyond mere genetic variation, to study ancient human evolution.

Here, we investigate how humans adapted to an agricultural lifestyle through comparative analysis of the ancient DNA methylomes of 18 hunter-gatherers and 24 farmers from western Eurasia. We've identified 458 differentially methylated CpG sites (DMCpGs) between the two lifestyles and have annotated 185 target genes that highlight biological functions associated with metabolic and cellular processes.

Moreover, we've found evidence of genetic variants associated with 168 of the DMCpGs. Altogether, we elaborate on how these variations likely underpin adaptive responses to the new agricultural lifestyle.

Symposium 6: Mutation rate rules: on the origin of germline mutations and their evolutionary fate

Evolution of germline mutagenesis as a function of genetics, demography, and reproductive longevity

Kelley Harris(harriske@uw.edu)

University of Washington

Abstract Mutation rates vary across the tree of life by many orders of magnitude, with fewer mutations occurring each generation in species that reproduce quickly and maintain large effective population sizes. A compelling explanation is that large effective population sizes facilitate selection against weakly deleterious "mutator alleles" which modulate DNA damage, proofreading, and repair; such mutator alleles may also explain the pervasive mutation spectrum variation that exists among vertebrates, other eukaryotes, and even virus strains. I will discuss the mutator alleles that have so far been identified in mice and humans, and I will describe how selection against such mutator alleles is jointly impacted by both population size and reproductive age. I argue that a long generation time likely amplifies the fitness consequences of a mutator allele, leading to the counterintuitive prediction that the species with the highest mutation rates per generation are also the species with the most effective mechanisms for DNA damage avoidance and repair. These insights into germline mutagenesis parallel recent findings that long-lived species have the lowest mutation rates in somatic tissues, uniting predictions made about germline mutagenesis and the biology of aging.

Mutagenic Treatments Create Male-biased Germline Mutations in Mice

Danqi Qin

Zhejiang University

Abstract The male germline mutation rate is generally higher than that of females across vertebrates, a phenomenon referred to as the male mutation bias. This disparity has traditionally been attributed to the higher number of cell divisions in the male germline, but this explanation has been challenged. An alternative explanation, the damage-induced hypothesis, posits that males have weaker DNA protection and repair mechanisms, leading to more damage-induced mutations. To test the validity of the damage-induced hypothesis, we conducted an experimental study using laboratory mice. Accordingly, we predicted that male mice would exhibit more germline mutations than female mice when treated with equal levels of mutagenic stress. We exposed male and female mice to five mutagenic chemicals and a PBS control, then employed parent-offspring whole-genome sequencing to detect de novo mutations. Compared to the PBS control, mutagenic treatments resulted in altered mutation rates and spectra. Notably, significant differences in mutation rates between sexes were observed in two of the five treatments, with males exhibiting higher rates in both cases. Furthermore, under one treatment (N-methyl-N-nitrosourea), only males showed an increased mutation rate and significantly altered mutation spectra relative to the control while these features remained unchanged in females. Comparative analysis with established mutational signatures revealed that a specific signature associated with deficiencies in the base-excision repair pathway was enriched in males but not females following mutagenic treatment. Collectively, our results provide direct evidence supporting the damage-induced hypothesis and elucidate its underlying molecular mechanisms.

Interactions between sequence context and methylation shape the germline mutation rate of CpG sites in the human genome

Ziyue Gao(ziyuegao@pennmedicine.upenn.edu)

University of Pennsylvania

Abstract The single-nucleotide mutation rate varies tremendously along the human genome, strongly correlated with the local sequence context. Notably, the mutation rate of hypermutable CpG sites also varies across sequence contexts, and this variation cannot be fully explained by methylation levels. In this study, we quantify the mutation rate of CpG sites in different sequence contexts and investigate the underlying mechanisms driving this variation.

Using a regression framework designed to account for recurrent mutations, we analyzed data from the 1000 Genomes Project and the gnomAD dataset to estimate context-specific CpG mutation rates. Our findings revealed that CpG mutation rates vary significantly across sequence contexts in both methylated and unmethylated states, though in distinct orders, indicating a context-dependent mutagenic effect of cytosine methylation. In addition, the upstream and downstream flanking bases exert largely independent effects on mutation rate. Interestingly, we observed an asymmetry between the two C:G base pairs within the same methylated CpG site: the C:G base pair with the cytosine flanked by an adenine on the same strand—either upstream or downstream—was more prone to mutation. To explain these observations, we tested three potential mechanisms: interference of DNA repair by transcription factor binding, demethylation by ten-eleven translocation enzymes, and local double-stranded DNA structural dynamics (i.e., “DNA breathing”).

Overall, our findings highlight the context-dependent impact of methylation on CpG mutation rates. By refining mutation rate estimates, this work provides a foundation to investigate molecular mechanisms underlying CpG hypermutability and improve mutation models for inferring evolutionary constraints and detecting disease mutations.

Detection of de novo SNV and microsatellite mutations in primate testis and human sperm samples from HiFi sequencing

Mikkel Heide Schierup(mheide@birc.au.dk)

Aarhus University

Abstract Trio studies have taught about de novo germline mutations in humans and other primates, but each trio only contributes 50-100 mutations. With very accurate sequencing, it should be possible to call de novo mutations directly from reproductive tissues. We use high-coverage Pacbio HiFi sequencing of a total of 20 samples, including human sperm samples and testis biopsies from humans, chimpanzees, gorillas, gibbons, baboons, macaques and marmosets. We call mutations by making a diploid de novo assembly from each individual, then map reads back to these and call de novo mutations as high-quality differences from the reference after extensive quality control. We then filter positions that could be caused by single-strand damage. In human sperm, the number of SNV mutations we identify is close to the number expected from trio studies. In testis biopsies, the number of called mutations is much larger because these include somatic cells which are known to have higher mutation rates. We, therefore, use the inferred CpG methylation pattern for each sequence read to classify it as having either somatic or germline origin, and we show how most SNV mutations of the testis thus can be assigned to germline or somatic origin. We then compare the mutational spectrum across primate species. Finally, we also call SNV mutations in microsatellites. Depending on age, we call between 1000 and 5000 SNVs and a similar number of microsatellite de novo

Rapid and highly variable mutagenesis of complex tandem repeats identified across four generations of a large Utah pedigree

Michael Goldberg(megoldberg2@gmail.com)

University of Utah

Abstract Variable number tandem repeats (VNTRs) are hotspots of rapid mutation driven by replication and recombination-based mutagenesis. VNTRs serve as ancestry markers and are implicated in disease and trait evolution. Their structure is often complex, including multiple motifs or interspersed with short tandem repeats; in extreme cases, orthologs comprise entirely different motifs. This complexity, and their length, obscures their evolutionary history and complicates genotyping. Understanding how VNTRs evolve and contribute to traits requires precise mutation rate estimates and quantifying rate variance between sites, thus elucidating the activity of different mutational pathways. We studied VNTR mutagenesis by sequencing 28 members of a four-generation pedigree with PacBio HiFi, and genotyped millions of VNTR loci to identify de novo mutations. VNTRs mutate at a rate of 2*10^-6 mutations/generation*locus, >100x higher than SNVs in non-repetitive regions. Eight loci mutate multiple times; at one complex VNTR, we identified 10 mutation events across 22 transmissions, making it one of the most hypermutable loci identified to date.

What explains these dramatic mutation rates? Interestingly, at complex hypermutable loci, an allele's mutation rate depends on the motifs it contains. Within the same locus, certain motifs mutate quickly while others remain comparatively stable. Similarly, conditional on locus, motifs that mutate in the pedigree are more historically polymorphic. Thus, while these loci are hypermutable, it is clear that locus composition matters over different evolutionary timescales. We will present our work on how changes in motif incite large changes in instability, an essential step toward unraveling mutational mechanisms.

An adaptive deep learning framework for high-resolution prediction of somatic mutations in normal and cancerous tissues

Kun Wu(haoxuan@zju.edu.cn)

Zhejiang university

Abstract Somatic mutations play a pivotal role in the pathogenesis of cancer and aging. Despite advancements in sequencing techniques, detecting somatic mutations remains a significant challenge, primarily due to genetic heterogeneity and low allele frequencies within mutated cells. Deep learning approaches have shown promising results in predicting germline mutation rate in humans. However, their applications to somatic mutations in both normal and cancerous tissues remain underexplored. In the study, we introduce the Human Mutation Rate Learner (HuMR), a deep learning framework developed for the accurate prediction of somatic mutations based on mutation data from 25 normal tissues and 30 cancerous tissues. Compared to previous models, we adopt the more efficient and accurate ConvNeXt architecture, facilitating the integration of genomic sequences and additional features. By employing larger convolution kernels and strides, the model enhances the capture of broader contextual information and optimizes the extraction of mutation-related features. Our results demonstrate that HuMR outperforms current state-of-the-art methods in predicting mutation rates at both single-base and larger genomic regions. Using a transfer learning strategy, we enable users to achieve highly accurate mutation rate predictions with a small set of tissue-specific mutation data. Additionally, our model proves applicable to the prediction of germline mutations, with results surpassing those of existing models. Based on this model, we created genome-wide fine-scale mutation maps across 25 normal tissues and 30 cancerous tissues. The open-source pipeline and generated somatic mutation rate maps hold great potential to significantly advance research in human pathogenic mutations and cancer genetics.

Mutation accumulation lines in mice to study mutational spectra, transcriptional phenotypic variance, and effects on fitness reduction

Diethard Tautz(tautz@evolbio.mpg.de)

Max-Planck-Institute for Evolutionary Biology, Ploen

Abstract We present results from a large mutation accumulation experiment using up to 59 inbred mouse lines across up to 21 generations, with a cryopreserved control to account for environmental influences. We used Illumina whole-genome sequencing and long-read PacBio sequencing to call the complete spectrum of new mutations, including structural mutations. On average, there are ~45 mutations per haploid genome per generation, about half of which (54%) are insertions and deletions shorter than 50 bp (indels). The remainder are single nucleotide mutations (SNMs, 44%) and large structural mutations (SMs, 2%). We found that the degree of DNA repetitiveness is positively correlated with SNM and indel rates. We uncover a rich landscape of active TEs and show that gene retroposition contributes to SMs at a high rate. Analysis of transcriptomic mutational variance shows a substantially higher transcriptome evolution rate in the brain than in the liver. A differential expression analysis revealed only minimal contributions to this variance from genes with significant expression changes. We have also studied the impact of spontaneous mutation accumulation on the mean and genetic variation for quantitative and fitness-related traits. We show that variation for morphological and life history traits accumulates at a sufficiently high rate to maintain genetic variation and selection response. Fitness proxy measures (litter size and surviving offspring) decrease on average by about 0.2% per generation. We also discuss these findings' implications in the context of concerns about a possible fitness reduction in humans due to reduced selection.

Determinants of de novo mutations in extended pedigrees of 43 dog breeds

Shaojie Zhang(1993zsj123@163.com)

Kunming Institute of Zoology, Chinese Academy of Sciences

Abstract Intensive breeding of dogs has had dramatic effects on genetic variants underlying phenotypes. To investigate whether this also affected mutation rates, we deep-sequenced pedigrees from 43 different dog breeds representing 404 trios. We find that the mutation rate is remarkably stable across breeds and is predominantly influenced by variation in parental ages. The effect of paternal age per year on mutation rates is approximately 1.5 times greater in dogs than humans, suggesting that the elevated yearly mutation rate in dogs is only partially attributed to earlier reproduction. While there is no significant effect of breeds on the overall mutation rate, larger breeds accumulate proportionally more mutations earlier in development than small breeds. Interestingly, we find a 2.6 times greater mutation rate in CG Islands (CGIs) compared to the remaining genome in dogs, unlike humans, where there is no difference. Our estimated rate of mutation by recombination in dogs is more than 10 times larger than estimates in humans. We ascribe these to the fact that canids have lost PRDM9-directed recombination and draw away recombination from CGIs. In conclusion, our study sheds light on stability of mutation processes and disparities in mutation accumulation rates reflecting the influence of differences in growth patterns among breeds, and the impact of PRDM9 gene loss on the de novo mutations of canids.

Distinct temperature-dependent patterns of molecular rates across vertebrates

Tianlong Cai(zhenying@westlake.edu.cn)

Westlake Laboratory

Abstract The relationship between molecular evolution rates and temperature has not been extensively explored across vertebrates. In this study, we analyzed molecular rates from mitochondrial (5,424 species) and nuclear (1,512 species) genomes across major vertebrate groups, including fishes, amphibians, reptiles, mammals, and birds. We observed distinct temperature-dependent patterns in absolute synonymous substitution rates (dS), influenced by thermoregulatory strategies. In ectotherms, dS increases with ambient temperature and decreases with latitude, whereas no significant correlation was found in most endotherms, except birds, which exhibited higher dS at higher latitudes. These patterns can be explained by the basal metabolic rate (BMR) theory. In ectotherms, body temperature is determined by ambient temperature, with higher temperatures driving higher mass-specific BMR and mutation rates, resulting in increased dS. In endotherms, although body temperature is independent of ambient temperature, increased heat production to adapt to cold environments, along with higher BMR in migratory birds breeding at high latitudes, likely contribute to higher dS in colder conditions. Interestingly, mitochondrial absolute nonsynonymous substitution rates (dN) increase with temperature in both endotherms and ectotherms. This pattern may be driven by the combined effects of mutation, selection, and drift, with mutation rates and purifying selection shaping the patterns in ectotherms and endotherms, respectively. Our study provides new insights into the temperature-dependent molecular evolution rates across vertebrates.

Reduced germline mutation rate in repeatedly evolved inbreeding social spiders

Jilong Ma(aujilongm@bio.au.dk)

Aarhus University

Abstract a large range of vertebrates has pointed to life history being an important determinant of the mutation rate and types of mutations. Invertebrates have received far less attention despite rapid changes in life history often observed. In spiders of the genus Stegodyphus, three independent transitions from solitary to social behaviour have occurred over the past 300 thousand years associated with inbreeding, sex ratio changes, reduced effective population size and larger and fewer eggs. We sequenced trios from three social and three solitary species for a total of 149 trios representing 25 families. We found that the mutation rates in solitary species are more variable among families for a mean rate of 3.29e-09 whereas mutation rates in social species are much reduced with a similar rate for each of the three social species, with a mean of 1.64e-09. This reduced rate is also reflected in a reduced synonymous branch length on each of the branches to the social species suggesting that it has been reduced throughout the time that species have been social. This observation may be caused by (a) adaptive evolution, (b) sheltered mutagenic environments associated with sociality, or (c) reproductive process shifts with sociality. As independent and repeated adaptive evolution with reduced Ne are considered less likely in a short time, the two other hypotheses are further explored and tested with differences in sibling-shared DNMs, paternal bias and mutational spectrum.

Symposium 7: The Origin, Evolution, and Phenotypic Contributions of New Genes

Functional innovation through new gene origination is a general evolutionary process

Manyuan Long(mlong@uchicago.edu)

The University of Chicago

Structures and functions of genes are the most basic features biologists are concerned with, whereas their origination and evolution are the fascinating subjects to understand for molecular evolutionary biologists. Over recent two decades, new genes evolved as new structure through various molecular mechanisms were widely documented. However, nothing in gene evolution makes sense except in the light of functional understanding while it has continuously proved so difficult to detect functional effects of molecular evolutionary changes. It has been a serious challenge to detect functional innovation brought by the new genes due to requirement of cross-disciplinary efforts for functional and evolutionary analyses. Nevertheless, the past decade saw the advancement of genomic scale functional analyses with advent of various highly variable knockdown and knockout libraries in model organisms and gene-based molecular analyses in medicine and agriculture. Our understanding of how functional innovation in evolution through new genes has advanced substantially, overturning various static views of gene functions in textbooks or views of luminaries in the history of biology and evolution. These efforts and findings offer a new and general paradigm for researchers from all biological fields from evolution to molecular biology to biochemistry, or medicine/agriculture to explore in various organisms. Exemplar findings are discussed to interpret gene functional innovation. New genes have been found to integrate into and modify pre-existing gene networks to form biochemical multimers. New genes may arise de novo from noncoding DNA sequences creating de facto novel proteins. These newly evolved genes have important roles in phenotypic and functional evolution across diverse biological processes and structures, making evolution functionally interpretable from processes to forces.

Evolutionary mechanisms of new gene birth and protein sequence convergence

Xuan Zhuang(xz036@uark.edu)

University of Arkansas

Determining the origins of novel genes and the mechanisms driving the emergence of new functions is crucial for understanding evolutionary innovation. Antifreeze proteins (AFPs), specifically the near-identical type I AFP (AFPI) found in four divergent fish taxa, provide an exceptional model to study these processes. This study tested the hypothesis of protein sequence convergence beyond functional convergence in three unrelated AFPI-bearing fish lineages. Using newly sequenced genomes of winter flounder and grubby sculpin, alongside existing high-quality genomes of cunner and related species, the analysis revealed that near-identical AFPI proteins originated independently in each lineage from distinct genetic precursors. Each lineage evolved a de novo coding region for the novel function while repurposing ancestral fragments as regulatory elements, exemplifying partial de novo origination—a process that bridges de novo gene formation and the neofunctionalization of duplicated genes. The findings support existing models of new gene origination and introduce new models: the Innovation-Amplification-Divergence (IAD) model, where novel changes precede gene duplication; the Duplication-Degeneration-Divergence (DDD) model, describing new functions arising from degenerated pseudogenes; and a DDD gene fission model, where sibling genes evolve distinct functional domains through differential degeneration and renovation. These insights underscore the diverse evolutionary pathways that lead to the emergence of novel functional genes with convergent protein sequences across divergent species.

This work advances understanding of the mechanisms driving new gene formation and was recently published in Molecular Biology and Evolution (MBE).

New genes enable protein structural innovation and function in the brain

Victor Luria(victor.luria@yale.ed)

Yale University

How genuinely new protein-coding genes originate is a central question in biology. Many new genes arising from non-copying mechanisms such as de novo from intergenic genomic DNA long considered to be “junk”, or from long non-coding RNAs, were recently found in Eukaryotes. New genes are taxon-restricted, may encode structurally novel proteins with new protein domains, and illuminate the emergence of protein structure. To understand how new genes arise, we built a mathematical model based on genomic parameters and dynamic factors including mutation. Combining gene age evaluation and proteogenomics, we identified taxon-restricted genes in >100 eukaryotic genomes from human to paramecium and evaluated their predicted biophysical properties. Compared to ancient proteins, new proteins are shorter, more fragile, disordered and promiscuous, yet less prone to forming toxic aggregates. We experimentally measured structure content and protease resistance of new human proteins and showed new genes function in vivo in zebrafish brains. We showed new human genes have fewer regulatory elements than ancient genes but more than control intergenic open reading frames. Our GTEx RNA expression analysis shows new human genes are expressed in many tissues. Our gnomAD mutational constraint analysis shows some new human genes are functional. Using single-cell RNA-Seq and proteomics, we found new genes are expressed in human brains at multiple ages. Thus, genomic sequence turnover generates many new genes encoding short proteins with distinct structural features and functioning in the brain. Variation in large eukaryotic genomes having large intergenic “dark matter” regions continuously generates new protein structures and new functions.

Solving a genetic paradox: from an essential gene to a killer

Polina Tikanova(polina.tikanova@gmail.com)

IMBA, VBC

Toxin-antidote elements (TAs) are selfish genes that ensure their survival by selectively killing the offspring that do not inherit them. TAs typically consist of two genes: a toxin and its cognate antidote, which must be in tight genetic linkage to drive in populations. Despite their simplicity, TAs evolutionary origin is perplexing: the toxin imposes extreme fitness costs without the antidote, and the antidote has no apparent function beyond neutralizing the toxin. This raises the question: what comes first, the toxin or the antidote?

We addressed this question by studying the evolutionary origin of three newly discovered TAs in Caenorhabditis tropicalis. The three toxins evolved from the same gene: fars-3, an essential and conserved subunit of the phenylalanine-tRNA synthetase. Their respective antidotes are also evolutionarily related and belong to a novel family of fast evolving F-box proteins, which, despite lacking sequence similarity to known F-box proteins, mediate the degradation of cognate toxins by recruiting SCF E3 ubiquitin ligase complex.

In C. wallacei, the sister species of C. tropicalis, as well as more distantly related nematodes, the closest orthologs to antidotes were located in the same chromosomal arm as fars-3, indicating the genetic linkage between their ancestral genes. Our data suggests that the duplication of an ancestral module consisting of fars-3 and an F-box protein drives the recurrent evolution of fars-3-derived TAs. This finding indicates that selfish behaviour can emerge in just one step from duplication of linked gene pairs coding for a protein degrader and its substrate.

Moderate presence and evolutionary causes of paralogous compensation in humans

Jieyu Shen

Institute of Zoology, CAS

Paralogous compensation is widely considered essential for genetic robustness, yet its in vivo manifestation and evolutionary maintenance remain largely unexplored in humans. To address this, we compiled a dataset of compensated paralogs based on published and newly generated CRISPR-Cas9 screening data. Compensated pairs were defined as those exhibiting negative genetic interactions (GIs), where single-gene knockouts had negligible fitness effects, but dual knockouts significantly reduced fitness. We classified paralog pairs into core (housekeeping/robust) GIs, conditional GIs, and non-interacting pairs and examined their distribution across in vivo and evolutionary contexts. Core GI pairs showed higher expression correlation during prenatal development, suggesting greater compensatory potential. In tumors, core pairs actively compensated for gene loss through upregulation of the remaining duplicate. Evolutionarily, ancient whole-genome duplicates (WGDs), particularly those with high pleiotropy, were enriched for core GIs, likely constrained by functional entanglement, which slowed sequence divergence. In contrast to prevailing views, recent small-scale duplicates (SSDs) showed limited compensation due to incomplete duplication and low or tissue-specific expression. However, out-of-X retrogenes, which evolve broad expression profiles, exhibited strong compensation for their parental copies. Together, these findings support paralogous compensation in early development and tumorigenesis and illuminate the evolutionary forces shaping both ancient and young paralogs.

Evolution of retrocopies of protein-coding genes in the context of HUSH complex repression.

Izabela Makalowska

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 retrocopies' 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.

Using high-quality gene annotations on complete primate genomes and human pangenomes to study gene evolution

Prajna Hebbar

University of California Santa Cruz

Vertebrate genomics is heading into an era where the release of hundreds of new reference genomes a year is going to be the norm. If we are to effectively engage with the new biology, we need to be able to annotate genes on these genomes accurately and efficiently. Unfortunately, the current state of gene annotation software requires the use of multiple tools and hours of manual curation. In order to remedy this, I am working on CAT2.0, an improved version of the Comparative Annotation Toolkit. CAT2.0 performs both evidence-based and de novo gene annotation. It can simultaneously annotate many closely related genomes using an annotated reference genome while incorporating information like RNA-seq/Iso-seq data or transcripts and protein homology. CAT2.0. was used to perform gene annotation of protein-coding and non-coding genes for all the Telomere-To-Telomere non-human primate (NHP) genome assemblies. We are also using CAT2.0 to annotate genomes for several large-scale genome projects- such as the Human Pangenome and Vertebrate Genomes Projects. In my talk, I will present current developments in the tool and the various gene annotation modules that can be used depending on the assemblies in hand. I will also describe some exciting results from using CAT2.0 annotations to study the evolution of a few gene families throughout the primate phylogeny, especially focusing on gene fusions and novel gene copies.

Gene expansion and evolution in detoxification and lignocellulose degradation enzymes in mangrove herbivorous crabs

Kwok Lun Hui

The Chinese University of Hong Kong

Mangrove herbivorous Sesarmid crabs were shown to have diversified with the emergence of mangrove plants. Their ability to utilize mangrove leaf litter as their major food source establish their ecological significance as primary consumer and is pivotal to the terrestrialization of these crabs. A versatile molecular toolkit of digestion and detoxification enzymes is required for crabs to degrade alkaloid deterrents and recalcitrant lignocellulosic compounds in plants. With herbivorous crabs conquering estuarine and freshwater environments, we anticipate gene expansion and/or diversifying selection occurred on detoxification and lignocellulolytic genes to enhance plant digestion capacity.

To test this hypothesis, we generated 20 de novo transcriptome assemblies and retrieved three RefSeq reference genomes of crabs, six of which are under Sesarmidae to unveil their gene evolution patterns. We clustered orthologous/paralogous coding sequences for gene expansion analyses and generated codon-aware multiple sequence alignment for non-synonymous to synonymous substitutions ratio (ω) analyses. Although results showed a lack of significant expansion or episodic diversifying selection ($\omega>1$) in most genes, endoglucanase GH9 and cytochrome P450 families, which involved in primary stages of cellulose digestion and detoxification respectively, were the only gene families found to experience both gene expansion and relaxed purifying selection/intensified positive selection in Sesarmidae. The expansion of lignocellulolytic and detoxification genes in Sesarmidae could provide redundancy as well as potential for positive selection for higher expression and enzyme activity of endoglucanase presented in previous studies and provide diverse detoxification enzymes that adapt to diverse plant deterrents, contributing to the adaptation in semi-terrestrial environments.

The gene regulatory networks and transcription factors controlling the expression of de novo originated genes

Li Zhao

The Rockefeller University

It has been demonstrated that, across the tree of life, novel genes can emerge from ancestrally non-coding sequences. Such genes are called de novo genes. However, the regulatory programs governing the expression of de novo originated genes remain unknown. To tackle this question, we used a cured list of de novo genes found in the Drosophila melanogaster lineage and used computational methods to single-cell RNA sequencing data, to identify the transcription factors regulating the expression of de novo genes in the testis. We found that all transcription factors are not equal regarding to regulating de novo genes: among the 83 transcription factors we studied, 3 accounted for the regulation of 28% of de novo genes. Leveraging genetic and genomic tools available in Drosophila, we further examined experimentally the role of two key transcription factors and the regulatory architecture of novel genes. Our findings identify key transcription factors associated with the expression of de novo genes and provide new insights into how modifications in existing transcription factors enable the emergence, maintenance, and regulation of de novo genes.

Massive introgression and gene duplications drove sexual system transition in an algal genus Vaucheria

Seok-Wan Choi

Sungkyunkwan University

The evolution of sexual systems has been a central theme in evolutionary biology since Darwin's time, yet many aspects remain unresolved. Developing a comprehensive understanding of sexual system evolution requires more insights from non-model eukaryotic lineages, beyond well-established plant and animal model species. In this study, we investigated the genomic changes associated with sexual system transitions in the algal genus Vaucheria (class Xanthophyceae, Heterokontophyta), which exhibits various sexual systems (hermaphrodite, monoecy, dioecy) within a single genus. We aimed to uncover the evolutionary trajectory of sexual systems in Vaucheria through phylogenomic analyses, using plastid genomes from 31 species and high-quality nuclear genomes from ten species. The phylogenomic analyses revealed that the transition from an ancestral hermaphroditic state to monoecy and dioecy occurred independently, coinciding with the adaptation from freshwater to brackish environments. Notably, we found that genome-wide introgression happened from dioecious species into monoecious species. Additionally, transcriptomic analyses and in situ hybridization experiments revealed that gene duplication events were significantly enriched among female-related genes in the introgressed species, implying that these duplications may have contributed to the rewiring of sexual development. Our findings suggest that the emergence of new sexual systems can be driven not only by gradual mutations and natural selection but also by more drastic, non-Darwinian processes such as introgression and gene duplication. This study provides valuable insights into the ecological and genomic mechanisms driving adaptation and sexual system transitions.

Monkeyflower (Mimulus) uncovers the evolutionary basis of the eukaryote telomere sequence variation

Jae Young Choi

University of Kansas

Telomeres are nucleoprotein complexes with crucial role of protecting chromosome ends. Because of its vital functions, components of the telomere, including its sequence, should be under strong evolutionary constraint. Yet across the tree of life there are numerous examples of telomere sequence variation and the evolutionary mechanism driving this diversification is unclear. Here, we studied the telomeres in Mimulus by investigating the noncoding telomerase RNA (TR), which is a core component of the telomere maintenance complex and determines the telomere sequence in eukaryotes. We conducted de novo transcriptomics and genome analysis of 18 species, and discovered Mimulus has evolved at least three different telomere sequences: (AAACCCT)n, (AAACCCG)n, and (AAACCG)n. We discovered several species with TR duplications, implying functional consequences that could influence telomere evolution. For instance, M. lewisii harbored two sequence-divergent TR paralogs while its sister species the paralog had pseudogenized. Nanopore-sequencing and fluorescence in situ hybridization indicated M. lewisii had a sequence heterogeneous telomere, and Telomeric Repeat Amplification Protocol combined with Terminal Restriction Fragment analysis confirmed the telomerase can use both TR paralogs for telomere synthesis. Interestingly in closely related species M. cardinalis, TR was also duplicated and both paralogs were expressed but its telomere consisted of a single telomere repeat. Evolutionary analysis indicated the TR paralogs arose from an ancient transposition-mediated duplication, which also underlies the evolutionary origin of multiple Mimulus species with divergent telomere sequences. We propose sequence variation in eukaryotic telomeres arises from an evolutionary process involving TR duplication, sequence divergence, and loss of TR paralog.

Using house mouse system to trace the evolutionary fate of newly originated gene copies

Wenyu Zhang

Northwestern Polytechnical University

The birth of new genes with novel functions has been considered to be a major contributor to adaptative evolutionary innovation. However, most of the previous studies were conducted in comparisons between species, and less attention has been paid to the evolutionary dynamics and patterns of very recently derived new gene copies that are still polymorphic within natural populations. Herein, we used house mouse as the study system to trace the origin of evolutionary fate of two types of newly arrived gene copies: retrocopies and alternative isoforms. We found the primary rate of both types of gene copies is orders of magnitude higher than the long-term rate. Comparisons with single-nucleotide polymorphism distribution patterns in the same populations show that most retrocopies and alternative isoforms are deleterious. Transcriptomic profiling analysis shows that new retrocopies become easily subject to transcription and have an influence on the expression levels of their parental genes, especially when transcribed in the antisense direction. These results imply that the impact of new gene copies on the mutational load has been highly underestimated in natural populations. Despite of their largely deleterious effects, we also found a small subset of new retrocopies and alternative isoforms with rapid increasing frequency in populations. They were likely involved in a recent adaptation process, evidenced by the strong signature of positive selection. Overall, our study shows the dual roles of new gene copies at population scale.

Symposium 8: Novel insights on genome architecture evolution

Adaptive structural variation in insect genomes

Chris Jiggins

University of Cambridge

There is considerable interest in the role of structural variation in evolution and I will present some examples demonstrating the adaptive importance of such variants in insect genomes. Wing colour patterns in Lepidoptera have long been an exemplar system for studying adaptive evolution. We have shown that TE insertions influencing regulatiory variation have been responsible for adaptaive change in both Heliconius and Hypolimnas butterflies. In the case of H. erato we could verify the adaptive role of the TE insertion by CRISPR recapitulation of the natural mutation. In Hypolimnas the mutation caused a reversion of phenotype to an ancestral form, which was mimetic to a toxic model species, Danaus chrysippus. In contrast, in the wood tiger moth, we have shown that a gene duplication was responsible for a wing pattern polymorphism, with an additional copy of a yellow family gene, which we named valkea, being responsible for a male-limited wing polymorphism. Finally, I will present evidence for introgression of inversions between two crop pest species, Helicoverpa armigera and H. zea, which hybridise in Brazil. Inversions have introgressed from zea into armigera on several chromosomes, and in all cases containing genes putatively involved in resistance to Bt toxins, extensively deployed in GM soybean and other crops in Brazil. These examples demonstrate that insertion/deletion and inversion variants, especially mediated by TEs, are likely are a common mechanism for adaptive evolution.

Genome scrambling in the Tree of Life

Charles Plessy

OIST

Genome scrambling is the phenomenon where gene order is massively reshuffled on homologous pairs of chromosomes in related species. While it is often observed at long evolutionary distances, we recently reported extreme genome scrambling in cryptic species of the globally distributed zooplankton Oikopleura dioica (phylum: chordates, sub-phylum: tunicates). Our quantitative analysis confirmed it to take place faster than in tunicates (Ciona), nematodes (Caenorhabditis) or fruit flies (Drosophila). To extend our finding to non-model organisms, we are now screening the whole Tree of Life in more than 30 branches containing 100~200 chromosome-level genome assemblies representing animals, plants, fungi, bacteria and archaea. Using the LAST software, we computed all possible pairwise genome alignments, for a total of more than 300,000 comparisons. Our results hint to the possibility that genome scrambling is the norm in the whole Tree of Life and that its speed at the evolutionary scale is independent from karyotype stability. After extracting statistics from each alignment to quantify genome scrambling and evolutionary distance, we screen for clades that show a significantly high scrambling or high genome stability compared with the baseline in their branch. By searching in these outlier clades for opposite molecular and phenotypic features, we aim at identifying the core principles governing the evolution of gene order, with the hope of applying them to synthetic genome design.

Chromosomal fusions linked to adaptive evolution in an invasive copepod species complex

Zhenyong Du

University of Wisconsin–Madison

Genome architecture evolution is a rapidly advancing field in evolutionary biology, fueled by the surge of new genomic data revealing unexpected patterns and mechanisms of adaptation. Chromosomal rearrangements are increasingly recognized as contributors to evolutionary adaptation, yet direct empirical evidence linking chromosomal evolution with response to selection remains scarce. This study investigates the role of chromosomal fusions in adaptive evolution within the copepod *Eurytemora affinis* species complex (Crustacea). Within this species complex, certain clades repeatedly give rise to invasive populations that cross habitat boundaries. High-quality chromosome-level genome assemblies from three clades revealed remarkable karyotypic evolution driven by up to eleven independent chromosomal fusion events within this species complex. The Atlantic clade (*E. carolleeae*), which is the most invasive, exhibited the fewest chromosomes (four) due to numerous fusion events. Comparative analyses between saline and freshwater populations of *E. carolleeae* uncovered significant enrichment of selection signatures at the chromosomal fusion sites, particularly at the centromeres, coinciding with regions of reduced recombination. At these fusion sites, collections of ion transport-related genes form potential "supergenes." These particular sets of ion transporter genes repeatedly show population genomic signatures of natural selection in both wild populations and laboratory selection lines during salinity transitions. This work provides novel empirical support for the hypothesis that chromosomal fusions might facilitate adaptation by clustering coadapted alleles. This emphasizes the potential role of chromosomal variation in driving evolutionary novelty and speciation. This study represents a rare demonstration that genome architecture evolution might be directly associated with adaptation to environmental change.

Horizontal transmission of functionally diverse transposons is a major source of new introns

Russ Corbett-Detig

UC Santa Cruz

Since the discovery of spliceosomal introns in eukaryotic genomes, the proximate molecular and evolutionary processes that generate new introns have remained a critical mystery. Specialized transposable elements (TEs), introners, are thought to be one of the major drivers of intron gain in diverse eukaryotes (1–8). However, the molecular mechanism(s) and evolutionary processes driving introner propagation within and between lineages remain elusive (1). Here, we analyze 8716 genomes, revealing 1093 introner families in 201 species spanning 1.7 billion years of evolution. Introners are derived from functionally diverse TEs including families of terminal-inverted-repeat DNA TEs, retrotransposons, cryptons, and helitrons as well as mobile elements with unknown molecular mechanisms. We identify eight cases where introners recently transferred between divergent host species, and show that giant viruses that integrate into genomes may facilitate introner transfer across lineages. We propose that ongoing intron gain is primarily a consequence of TE activity in eukaryotes, thereby resolving a key mystery of genome structure evolution.

Chromosomal Chaos: Exploring a Dramatic Shift in the Evolution of Genome Organization in Mucoromycota

Helle Baalsrud

The Norwegian University of Life Sciences

The evolution of species is influenced by changes in individual genes, but also by alterations in their chromosomal organization. The connection between the evolution of biodiversity and the order and content of genomic features on chromosomes, known as synteny, remains an underexplored frontier in biology. Using chromosome-scale assemblies from 25 species in the understudied Mucoromycota fungi we detected an intriguing lineage-specific shift in the rate and mode of synteny evolution. This shift entails a higher overall rate of rearrangements as well as a distinct mode known in fungi as “mesosynteny”, which is characterized by scrambling of gene order within chromosomes and a paucity of inter-chromosomal rearrangements. Interestingly, this shift coincides with loss of essential centromere protein (CENP) genes which are essential for meiosis and genome stability and are highly conserved across Eukaryotes. This loss could have affected how new chromosomal rearrangements navigate meiosis as heterozygotes. The mesosyntenic lineage also has a significantly higher repeat content as well as multiple instances of whole genome duplications. 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.

An episodic burst of massive genomic rearrangements and the origin of non-marine annelids

Carlos Vargas-Chavez

Institute of Evolutionary Biology (CSIC - UPF)

Genome-level analyses have revealed a striking pattern of conserved macrosynteny within chromosomes across huge phylogenetic distances in animals, yet progress in many lineages has been hampered by the absence of genome-level data. Here, we show that the tectonics of genome evolution in clitellates, a clade composed of most freshwater and all terrestrial species of the phylum Annelida, is characterised by extensive genome-wide scrambling that resulted in a massive loss of macrosynteny between marine annelids and clitellates, to the point that ancient bilaterian linkage groups are fully disrupted. These massive rearrangements included the formation of putative neocentromeres with newly acquired transposable elements, and preceded a further period of genome-wide reshaping events including whole-genome duplications and massive macrosyntenic reshuffling between clitellate lineages, potentially triggered by the loss of genes involved in genome stability and homeostasis of cell division. Notably, while these rearrangements broke short-range interactions observed between Hox genes in marine annelids, they were reformed as long-range interactions in clitellates. These rearrangements, relocating genes and creating chimeric elements, may have contributed to the adaptation of clitellates to freshwater and terrestrial environments. Our findings reveal extensive genomic reshaping in clitellates at both the linear (2D) and three-dimensional (3D) levels, suggesting that, unlike in other animal lineages where synteny conservation constraints structural evolution, clitellates exhibit a remarkable tolerance for chromosomal rearrangements. Our study thus suggests that the genomic landscape of Clitellata resulted from a rare burst of genomic changes that ended a long period of stability that persists across large phylogenetic distances.

Genomic evidence for rediploidization and adaptive evolution following the whole-genome triplication

Xiao Feng

Sun Yat-sen University

Whole-genome duplication (WGD), or polyploidy, events are widespread and significant in the evolutionary history of angiosperms. These events have been proposed to contribute to plant adaptation, even survival in the face of drastic environmental changes. However, empirical evidence for rediploidization, the major process by which polyploids give rise to modern diploid descendants, is still lacking at the genomic level. Here, we present the assembly of two chromosome-scale genomes from Lythraceae plants, the mangrove tree Sonneratia alba and the related inland plant Lagerstroemia speciosa. Their common ancestor experienced a whole-genome triplication (WGT) approximately 64 Myrs ago coinciding with a period of dramatic global climate change (K-Pg boundary). Subsequently, Sonneratia adapted to the unique habitats of mangroves. We hypothesize that the rediploidization process in mangroves differs from that of non-mangrove relatives. Following the recent WGT event, this mangrove underwent extensive chromosomal fractionations, rearrangements, and fusions that have shaped its current genome structure. We observe the WGT retentions display sequence and expression divergence, suggesting potential neo- and sub-functionalization. Additionally, we observe strong selection acting on three-copy retentions following the polyploidization–rediploidization process, indicating the potential adaptive value and accelerated evolutionary adaptation to new environments. Several retentions are associated with root development and salt tolerance, potentially contributing to surviving in intertidal zones. To elucidate the role of ploidy changes in genome evolution, we improve a model of the polyploidization–rediploidization process based on genomic evidence, aiming to enhance our understanding of adaptive evolution during periods of global climate change.

Genetic diversity and evolution of rice centromeres

Dongya Wu

Zhejiang University

Understanding the driving force of centromere dynamics is crucial for deciphering the complexity of eukaryotic evolution and speciation. Despite their widely recognized characteristics of conserved function in cell division, centromeres exhitbit extremely high diversity in composition and structure between species. The mechanism underlying this paradox remains poorly understood. Here, we assembled 67 high-quality rice genomes from the Oryza AA group, encompassing both Asian and African rice species, and analyzed over 800 nearly complete centromeres. Through de novo annotation of centromeric satellite CEN155 sequences and employing a progressive compression strategy, we quantified the local homogenization and multi-layer nested structures of rice satellite arrays. Our results indicate that genetic innovations in rice centromeres primarily arise from internal structural variations and centrophilic retrotransposon insertions, along with a certain number of non-canonical satellite repeats. Despite these rapid structural alterations, the single-base substitution rate in rice centromere regions appears relatively lower compared to the chromosome arms. Contrary to the KARMA model proposed for Arabidopsis centromere evolution, our rice-based study suggests that retrotransposon invasion probably contributes to the decline of progenitor centromeric satellite arrays, and facilitates the repositioning of centromeres, with extended CENH3 ChIP-seq enrichment beyond the native satellite arrays. In summary, our findings provide novel insights into centromeric divergence among rice species and subspecies, and advances our understanding of plant centromere evolution.

Inversions contribute disproportionately to parallel genomic divergence in dune sunflowers

Kaichi Huang

Sun Yat-sen University

The probability of parallel genetic evolution is a function of the strength of selection and constraints imposed by genetic architecture. Inversions capture locally adapted alleles and suppress recombination between them, which limits the range of adaptive responses. In addition, the combined phenotypic effect of alleles within inversions is likely to be greater than that of individual alleles; this should further increase the contributions of inversions to parallel evolution. We tested the hypothesis that inversions contribute disproportionately to parallel genetic evolution in independent dune ecotypes of Helianthus petiolaris. We analysed habitat data and identified variables underlying parallel habitat shifts. Genotype–environment association analyses of these variables indicated parallel responses of inversions to shared selective pressures. We also confirmed larger seed size across the dunes and performed quantitative trait locus mapping with multiple crosses. Quantitative trait loci shared between locations fell into inversions more than expected by chance. We used whole-genome sequencing data to identify selective sweeps in the dune ecotypes and found that the majority of shared swept regions were found within inversions. Phylogenetic analyses of shared regions indicated that within inversions, the same allele typically was found in the dune habitat at both sites. These results confirm predictions that inversions drive parallel divergence in the dune ecotypes.

Symposium 9: Origin and evolution of cell types

Exploring cell type evolution through single-cell transcriptomics

Jordi Solana Garcia

Living Systems Institute, Department of Biosciences, University of Exeter

Cell types are the fundamental units of multicellular organisms, yet their origins and evolutionary trajectories remain largely unresolved. While some cell types, such as neurons and muscles, are conserved across metazoans, others, like cnidarian cnidocytes or vertebrate immune cells, are lineage-specific innovations. How do new cell types emerge? Addressing this question requires a comparative framework spanning diverse lineages.

Recent advances in single-cell transcriptomics have transformed our ability to classify and compare cell types across species. In this talk, I will present insights from single-cell atlases of planarians, annelids, and cnidarians – three groups with striking regenerative capacities and diverse developmental modes. In the asexually reproducing annelid Pristina leidyi, we uncovered a rich cellular diversity and identify a piwi+ stem cell population with a pluripotency signature, suggesting deep evolutionary roots of adult stem cell systems. In the colonial cnidarian Hydractinia symbiolongicarpus, we revealed a novel stolon-specific cell type linked to biomineralization, highlighting how cell type combinations shape colonial organization. Comparative single-cell analyses of five planarian species further expose interspecies variation in secretory and digestive cell types, revealing patterns of cell type evolution in closely related species.

Despite these advances, major challenges persist in cross-species comparisons, particularly in defining homologous cell types when gene duplication and regulatory divergence obscure direct equivalences. I will discuss how integrating molecular atlases with new computational approaches can refine our understanding of cell type evolution and provide a roadmap for future comparative studies. By leveraging diverse animal models and large-scale single-cell data, we can begin to decode the evolutionary mechanisms underlying cell type diversity and their role in shaping complex life forms.

A novel cell type enabled the transition from egg-laying to livebearing and then to placentation in fishes

Diego Safian

Crick Institute

Biological complexity emerges step-by-step from simpler forms. A fascinating example is the evolution of different reproductive modes in the fish family Poeciliidae. Unlike egg-laying fish, all Poeciliids retain the nutrient-rich eggs in the ovary, with the embryos developing inside the mother until birth (egg-laying to livebearing transition). Remarkably, several Poeciliids additionally evolved a placenta, an organ allowing nutrient transfer from the mother to the embryo throughout development. Using single-cell RNA sequencing, we have characterized the cellular composition of the ovary in eight species (seven Poeciliids and an outgroup), covering the transition from egg-laying to livebearing and four independent origins of a placenta. We discovered that a Poeciliid-specific cell type is the main component of the ovarian follicle that surrounds each developing embryo. This novel cell type does not appear to respond to ovulation signals and likely enabled Poeciliids to transition from egg-laying to livebearing. By tracing the evolution of ovarian cell types across the Poeciliid phylogeny, we also found that the new Poeciliid cell type was fundamental in four independent transitions to placentation. In each of the four transitions, the Poeciliid-specific cell further diversified and acquired genetic programs associated with nutrient transport, enabling the evolution of a placenta. Our work shows that the evolution of a new cell type was fundamental for the evolution of live birth and placentation.

Visual system cell type differences underlying divergent butterfly mate preference

Wei Lu

University of Chicago

Closely related species often evolve unique mating cues and the corresponding mate preference. The genetic mechanisms underlying such quick turnovers of mate preference remains largely unknown. Here we focus on the visual mate preference variations in Heliconius, a group of poisonous butterflies showing diverse warning colors, which also serve as mating cues. Previous GWAS of male mate preference for yellow wing colorations has identified a major effect locus on chromosome 1. To further understand the cellular level changes underlying diverse mate preference, we conducted single-nuclei RNAseq (snRNAseq) of adult compound eyes in both sex from multiple populations with varied mate preference for yellow. We found that while most cell type abundances are similar across species and sex, the abundance of three subtypes of pigment cells are highly variable in males and yellow preferring males have higher proportions of these pigment cells. These pigment cells are all marked by the expression of a transcription factor sens-2, which is located 15kb away from one of the top mate preferences GWAS peaks. This GWAS peak falls within a single topologically associated domain (TAD) identified through micro-C in adult male brains and sens-2 is the only gene inside this TAD, which suggests a potential regulatory role. Finally, we showed that pigment cells in Drosophila don't express sens-2. Overall, our results suggest a few subtypes of pigment cells gain novel expression of sens-2 in yellow preferring males and that leads to divergent color preference during mate choice.

A daphnia single-cell atlas of tissue-specific gene expression: insights into the evolution of cell types in reproductive systems

Yue Hao

Translational Genomics Research Institute

Key questions in evolutionary genetics center on the origins and maintenance of reproductive-system diversity across phylogenetic lineages. In animals, transitions between sexual and asexual reproduction represent a critical area of study, yet most research has focused on obligate reproductive strategies and molecular pathways, limiting insights into the mechanisms driving cell-type differentiation and mating-system evolution. Here, we focus on Daphnia, a model organism that undergoes cyclical parthenogenesis. Under optimal conditions, Daphnia females reproduce asexually for numerous generations, but environmental changes induce male production, and entry into sexual reproduction. This reproductive shift is likely controlled by signals from various tissues, including sensory organs like the eyes and antenna (which detect environmental cues), the brain (which processes and transmits these signals), and the reproductive organs. To identify unique cell types within each tissue and uncover the molecular signatures essential for their functions, we constructed a single-cell transcriptomic atlas of Daphnia tissues. Using split-pool ligation-based transcriptome sequencing technology (Parse Biosciences), we sequenced 120,000 cells from 12 dissected tissues, as well as the entire head and body. By comparing these data to previously published single-cell datasets from other pancrustacean species, we aim to uncover the genetic mechanisms underlying cell-type and mating-system evolution. This work contributes to the growing body of tissue-specific single-cell atlases, offering novel insights into the evolution and persistence of reproductive systems across diverse species.

Comparative single-cell transcriptomics reveals variable degrees of cell-type conservation between two distantly related nematodes

Yinan Wang

Max Planck Institute for Biology, Tübingen

Different cell types perform distinct functions and might be under different levels of evolutionary constraint. With known cell homology and large genetic divergence (estimated divergence time 100-250MYA), the free-living nematodes Pristionchus pacificus and Caenorhabditis elegans represent a promising model system to investigate transcriptional evolution between homologoues cell types.
We applied scRNA-seq to P. pacificus and identified all major cell types including muscle, neuron, hypodermis, intestine, and germline cells. Comparison of one-to-one orthologues expression between homologous cell-type pairs in these two species indicated variable degrees of transcriptome divergence, with neurons showing low expression similarity. To further support the expression divergence between orthologous pairs in neurons, we confirmed individual cases experimentally. We extended our analysis to evolutionary young genes and quantified their contribution to cell-type specific transcriptomes by calculating their transcriptome age index (TAI). This identified some reproductive cell types including male germline and amphid neurons as cell types having the youngest transcriptomes. The fast transcriptomic evolution of the amphid neurons is consistent with the previously identified morphological divergence of this class of sensory neurons between the two species. Thus, the comparison between two nematodes with similar cell composition and long divergence showed varying degrees of evolutionary changes among cell types and confirmed the fast evolution of the reproduction system and sensory neurons.

Single-cell profiling of the amphioxus digestive tract reveals conservation of chordate endocrine cells

Yichen (Serena) Dai

Fudan University

Despite their pivotal role, the evolutionary origins of vertebrate digestive systems and key endocrine cell types remain enigmatic. We explored the cellular characteristics of the amphioxus (Branchiostoma floridae) digestive tract, a model for the presumed primitive chordate digestive system, using bulk tissue companioned with single-cell RNA sequencing. Our findings reveal segmentation and a rich diversity of cell clusters, and we highlight the presence of epithelial-like, ciliated cells in the amphioxus midgut and describe three types of endocrine-like cells that secrete insulin-like, glucagon-like, and somatostatin-like peptides. Furthermore, Pdx, Ilp1, Ilp2, and Ilpr knockout amphioxus lines revealed that, in amphioxus, Pdx does not influence Ilp expression. We also unravel similarity between amphioxus Ilp1 and vertebrate insulin- like growth factor 1 (Igf1) in terms of predicted structure, effects on body growth and amino acid metabolism, and interactions with Igf-binding proteins. These findings indicate that the evolutionary alterations involving the regulatory influence of Pdx over insulin gene expression could have been instrumental in the development of the vertebrate digestive system. Furthermore, cross-species comparison between amphioxus and zebrafish gut cell types suggest early emergence of endocrine cell types secreting glucagon-like and somatostatin-like peptides in the gut of the last chordate ancestor.

Single-cell analysis of the amphioxus hepatic caecum and vertebrate liver reveals genetic mechanisms of vertebrate liver evolution

Baosheng Wu

Institute of Zoology, Guangdong Academy of Sciences

The evolution of the vertebrate liver is a prime example of the evolution of complex organs, yet the driving genetic factors behind it remain unknown. Here, we study the evolutionary genetics of liver by comparing the amphioxus hepatic caecum and the vertebrate liver, as well as examining the functional transition within vertebrates. Using in vivo and in vitro experiments, single-cell/nucleus RNA-seq data, and gene knockout experiments, we confirm that the amphioxus hepatic caecum and vertebrate liver are homologous organs and show that the emergence of ohnologs from 2R-WGD significantly contributed to the functional complexity of the vertebrate liver. Two ohnologs, kdr and flt4, play a significant role in the development of liver sinusoidal endothelial cells. Additionally, we found that liver-related functions such as coagulation and bile production evolved in a step-by-step manner, with gene duplicates playing a significant role. We reconstructed the genetic footprint of transfer of heme detoxification from the liver to the spleen during vertebrate evolution. Together, these findings challenge the previous hypothesis that organ evolution is primarily driven by regulatory elements, underscoring the importance of gene duplicates in the emergence and diversification of a complex organ.

Evolution of novel cell types in the mammalian placenta

Marcin Falis

The Francis Crick Institute

The emergence of novel cell types is an important mechanism underlying the evolution of organismal diversity. Yet, we lack empirical evidence of how new cell types arise and diversify across species. The mammalian placenta, an organ that supports embryo development during pregnancy, offers a unique opportunity to investigate the evolution of novelty. This is because the placenta arose recently and has since diverged rapidly, resulting in an organ that is remarkably diverse across mammals. Morphological and functional differences amongst placentas are often attributed to trophoblasts - foetal cell types, which include unusual multinucleated and polyploid cells. Trophoblasts are diverse and may harbour examples of novel cell types, but their evolutionary history remains unexplored.
Here, we integrate evolutionary and developmental analyses of single-cell transcriptome data to reconstruct the evolutionary relationships between trophoblasts. Using placentas from murids (mouse, rat), guinea pig, rabbit, human, and marmoset, we profile nearly 380,000 cells and compare the resulting atlases across species. Whilst most placental cell types are conserved across species, we discover a remarkable lineage-specific expansion of trophoblasts. We identify three novel trophoblast cell types that evolved in murid placentas and show that they took up functions that were ancestrally performed by a single cell type. We also observe that invasive trophoblasts, a long-standing candidate of species-specific cell types, show transcriptomic similarity suggestive of homology. Our results provide a rare empirical example of cell type innovations in mammals and shed light on how new cell types evolve.

Evolution of lung somatic cells' life beyond orgnismal death - transitioning from multicellular organisms to unicellular entities

Bingjie Chen

Guangzhou Medical University

The mechanisms governing cellular adaptation to organismal death remain poorly understood, particularly how differentiated cells transition from multicellular regulation to autonomous survival following systemic collapse. Here, we demonstrate that in human and murine lung tissues, over 30% of cells retain viability beyond 20 days post-mortem (PMI) at near-freezing temperatures, with subsets exhibiting proliferative capacity and heightened metabolic activity. These surviving cells predominantly display stem-like properties, enabling environmental adaptation under extreme post-mortem conditions. Through isolation and cultivation, stem cells from PMI-extended tissues (≤20 days) were successfully grow into functional alveolar organoids. Integrated single-cell transcriptomic (scRNA-seq) and chromatin accessibility (scATAC-seq) analyses revealed PMI-dependent molecular reprogramming, characterized by upregulated stress response, cytoskeletal maintenance, and protein synthesis pathways. Notably, surviving cells exhibited activation of embryonic developmental transcription factors and re-engagement of pluripotency-associated regulatory networks. Furthermore, extended PMI-derived cells further demonstrated lineage plasticity, generating alveolar type I/II pneumocytes and mesenchymal cells within organoids. Our findings uncover a previously unrecognized resilience of somatic cells, which can decouple from multicellular organismal constraints post-mortem and adopt unicellular survival strategies. This paradigm shift highlights the dynamic adaptability of differentiated cells under systemic posmorterm samples, offering insights into somatic cell evolution after death and potential applications in regenerative medicine.

Symposium 10: Pangenome graphs and their applications in biodiversity genomics

Pangenomes of scrub-jays (Aphelocoma) across a gradient of effective population sizes reveal fitness effects of structural variation

Scott Edwards

Harvard University

In this talk I will present an analysis of 45 genomes from three species of Scrub-Jay (Aphelocoma), each assembled with PacBio HiFi reads and summarized in a multi-species pangenome graph. By comparing the diversities of SNPs and structural variants (SVs) across a gradient of effective population sizes, as well as by estimating the fitness effects of SVs directly, we conclude that on average, SVs are deleterious and that the extent of deleteriousness increases with SV length. Our data also sheds light on the population dynamics of inversions, telomeres and satellites, revealing rapid shifts in genomic abundances between species.

Pangenomics for De-Extinction: Towards the Great Passenger Pigeon Comeback

Simona Secomandi

The Rockefeller University

The passenger pigeon (Ectopistes migratorius) was once the most abundant bird species in North America, numbering between 3 and 5 billion individuals before being driven to extinction by human activities in the early 1900s. Thousands of museum specimens exist, and its closest relative, the band-tailed pigeon (Patagioenas fasciata), is still extant. Thus, the passenger pigeon is one of the few species amenable for de-extinction, the ecological replacement of an extinct species by adapting a living organism to fulfill its ecological role through phenotype alteration with the goal of promoting ecosystem resilience and biodiversity. Restoring the passenger pigeon's ecological role will particularly affect Eastern North American forests, which once benefitted from the disturbances generated by its large flocks. Indeed, the passenger pigeon was the only truly colonial species in the Columbidae family, with its sexual dichromatism and graduated tail possibly evolving from this behavior, both rare traits in the family. These three traits are the focus of this project, which involves state-of-the-art sequencing technologies and bioinformatic pipelines to reconstruct the genomic blueprint of the passenger pigeon. We generated high-quality genomes for multiple band-tailed pigeon individuals sampled across the species' geographic range and assembled a pangenome reference to help map the passenger pigeon's ancient DNA and mitigate reference biases. This approach will enable more confident variant calling between the two species, facilitate the reconstruction of the extinct species' genome, and pinpoint the "de-extinction mutations" required for editing the band-tailed pigeon's genome to restore the passenger pigeon's phenotypes and ecological role.

Exploring the impact of structural variation on adaptive evolution in Quercus based on super-pangenome

Ruirui Fu

Zhejiang University

The genus Quercus (Oaks) is the most widely distributed woody plant in the world and one of the largest tree genera in the forest ecosystem of the Northern Hemisphere. It has a wide range of phenotypic diversity and environmental adaptability, and interspecific hybridization and introgression are obvious. However, the connection and potential mechanisms between its genome evolution and adaptability are still unclear. Our previous studies have shown that adaptive gene introgression between two closely related oaks can be maintained by suppressing recombination caused by chromosomal inversion. In the study, we preliminary constructed high-quality chromosome-level genomes of nine oaks, with annotated genes ranging from 39,466 to 44,803 and repetitive sequences accounting for 49.21% to 54.47% of the total. Together with the 15 published oak species, we present a genus-level super-pangenome consisting of 24 oak genomes, in which core genes, softcore genes, dispensable genes, and private genes account for 29.13%, 26.28%, 40.93%, and 3.67%, respectively. Further systematic analysis of the characteristics and functions of structural variation in Quercus at the genome level will provide new insights into the formation of species diversity and adaptive evolution mechanisms in Quercus.

Assembly-based pangenome graphs illuminate the plasticity of mammalian gene repertoires

Landen Gozashti

Harvard University

Pangenome methods enable the unbiased characterization and representation of all forms of genomic variation from multiple individuals within a species, and are thus especially useful for species with high levels of genetic diversity. Here, we constructed a pangenome for the deer mouse (Peromyscus maniculatus), a model for studying local adaptation and a potent reservoir for diverse zoonotic diseases, using long-read assemblies from 14 ecologically and biogeographically diverse populations and 2 outgroup species (totaling 20 haplotypes). Using this pangenome, we identified on average a remarkable ~1.4 Gb of sequence not represented within any single haplotype, and 138 million single nucleotide polymorphisms (SNPs), 25 million indels (variants < 50 bp), and 2 million structural variants (variants > 50 bp). Recent endogenous retroviral invasions account for much of the structural variation across deer mouse populations and SVs generally exist at lower frequencies than SNPs and indels, consistent with strong purifying selection on larger variants. Furthermore, we uncover widespread gene copy number variation, including genes that vary in copy number by an order of magnitude across deer mouse haplotypes, in addition to several gene-rich loci that display high levels of structural diversity. Copy number variable genes are enriched for functions in immunity, sensory perception and environmental interaction. Several of those display signatures of positive selection at the molecular level, suggesting functional diversification. Together, our work provides important insight into the utility of pangenomes for genetically diverse species and illuminates the plasticity of mammalian gene repertoires.

Complete assemblies and pangenome reference reveal unique features in the complex genomic regions of Tibetan highlanders

Bing Su

Kunming Institute of Zoology, Chinese Academy of Sciences

Pangenome references and complete genome assemblies provide unprecedented opportunities to gain in-depth insights into the human genomic complexity and diversity. Here, we present a pangenome reference of Tibetan populations by sequencing 105 Tibetan genomes from 35 complete trios. We integrated 70 haplotype-resolved assemblies in the pangenome graph, including 8 nearly complete genomes and 27 high-quality diploid genomes. We identified ~15 million small variants and 223,151 structural variants (SVs), of which 5 million small variants and 53,829 SVs have not been reported in the current global pangenome reference. We added 122 million base pairs of euchromatic polymorphic sequences relative to the reference genome of T2T-CHM13. Furthermore, we characterized a large number of novel genomic variants and architectures in the complex genT2Tomic regions using fully-sequence-resolved Y chromosomes and T2T autosomes of the Tibetan genomes, especially the distinctive repeat motifs in telomeres and centromeres. Additionally, we uncovered novel inversions and variable number tandem repeats (VNTRs) enriched in Tibetan populations with potential contribution to high-altitude adaptation. We also identify 1,911 archaic-introgressed segments and reconstructed the evolutionary history of the Denisovan-EPAS1 haplotype in the Tibetan genomes. Our study demonstrates the efficacy of pangenome analysis in resolving complex regions of the human genome shaped by population history and biological adaptation.

Evolution and maintenance of a transpecies color polymorphism

Ming Li

University of Konstanz

Color polymorphisms are often associated with speciation, especially sympatric speciation1–3. Theoretically, color polymorphisms that mediate assortative mating along with disruptive selection could lead to sympatric speciation. In the Midas cichlid fish species complex a golden color morph exists across the entire species flock4,5. Yet, despite disruptive selection6–10 and strong assortative mating6,11,12 according to color, they have not (yet) genetically differentiated appreciably4. We generated 32 chromosome-level haplotype-resolved assemblies for 16 phenotypically and genetically diverse Midas cichlid individuals, and constructed a graphed pangenome reference, and combined 453 re-sequenced individuals in an effort to dissect the genetic basis of these two color morphs. This high-quality pangenome approach allowed for the discovery of a novel TE-derived complex structural variant that underlies the color polymorphism. Reconstructing the evolutionary trajectory of the TE-insertion history revealed that the origin gold morph is older than the sympatric speciation in crater lakes and is an ancient polymorphism that was maintained across new species boundaries. We found that the ontogenetic transition from the dark, all individuals start as “dark” morphs, to the gold morph is dosage-dependent and heterozygote individual loose the black pigments later in life. In mate choice experiments the two color morphs exhibit nearly complete assortative mating, but, importantly, individuals that reach sexual maturity before color-transition were found to mate according to their still dark phenotype not their gold genotype. This “anti-genetic” mate choice facilitates gene flow between the two color morphs and contributes to the maintenance of this color polymorphism, counteracting speciation according color morphs.

Symposium 11: Phylogenomics in the presence of gene flow: opportunities and challenges

TBA

Anne D. Yoder

Duke University

The power of coalescent methods for inferring recent and ancient gene flow in endangered Bactrian camels

Tianqi Zhu

Academy of Mathematics and Systems Science, CAS

Genomic sequence data harbour valuable information concerning the history of species divergence and interspecific gene flow, and may offer important insights into conservation of endangered species.
However, extracting such information from genomic data requires powerful statistical inference methods. A recent analysis of genomic sequence data found little evidence for gene flow from domestic Bactrian camels into the endangered wild Bactrian species. Nevertheless, the methods used to infer gene flow are based on data summaries and lack the power and precision to represent the complex phylogenetic history of the species with gene flow. Here we apply newly developed Bayesian methods to genomic sequence data to test for both recent and ancient gene flow among the three species in the genus Camelus, and to estimate the strength and timing of gene flow. We detect strong signal of gene flow from domestic into wild Bactrian camels, confirming early evidence based on mitochondrial DNA and the Y chromosome. Overall gene flow appears to affect the autosomal genome uniformly, with similar effective rates of gene flow for exonicand noncoding regions. Estimation of species divergence times is seriously affected if gene flow is not accommodated in the analysis. Our results highlight the power of the coalescent model in analysis of genomic data and the utility of the coding as well as noncoding parts of the genome in elucidating the evolutionary history of modern species.

Palaeogenomic study of the extinct Bubalus mephistopheles shed light on the evolutionary history of the Bubalus Genus and domestic water buffaloes

Xueyuan Liang

Peking University

The water buffaloes (Bubalus spp.), once widespread across Eurasia in Pleistocene, are now mostly restricted to tropical South and Southeast Asia. The extinct B. mephistopheles was the last water buffalo species found in East Asia during the Holocene. Its close relationship with human makes it once suggested to be the wild ancestor of the swamp buffalo. However, this species disappeared from around 3,000 years ago, with its taxonomy status, genetic relationship with other buffalo species especially the domestic water buffalo, remain largely unclear. To address these questions, we generated 30 mitochondrial genomes and six nuclear genomes of B. mephistopheles, together with one Late Pleistocene Bubalus wangsjoki. By integrating with modern water buffalo genomes, we revealed the taxonomic relationship of B. mephistopheles with other Bubalus species, and suggested that it was conspecific with the B. wangsjoki which was found in Pleistocene Northeast China. We further detected gene flow between B. mephistopheles and modern swamp buffaloes, indicating that the genetic diversity of domestic water buffaloes were also shaped by post-domestication introgressions from wild species, as described in domestic cattle. This study provides new molecular insights for understanding the evolution of the Bubalus genus, and the dispersal of swamp buffaloes.

Extensive introgression obscures phylogenetic relationships among major clades of the walnut family (Juglandaceae)

Bowen Zhang

Beijing Normal University

Despite numerous phylogenetic analyses, the phylogeny of the walnut family (Juglandaceae) has not been confidently resolved due to its complex evolutionary history shaped by ancient allopolyploidy, among-species rate variations, and interspecies gene flows. Recent advancements in phylogenomic studies, facilitated by the assembly of more complete genomes, present an opportunity to tackle these challenges. In this study, we analyzed 14 chromosome-level genomes spanning all nine genera of Juglandaceae, which included two newly assembled genomes from Alfaroa costaricensis and Oreomunnea mexicana. To address post-polyploid evolution and cross-species introgression simultaneously, we carefully curated orthologous gene sets, considering the dominant and recessive subgenomes separately to mitigate biased gene loss. By employing a gene-content tree and applying filtering based on the complexity hypothesis, we established a robust species branching order that aligns with microsynteny-based topology. To avoid the influence of among-species rate variations, we investigated introgression events based on this species branching order, and constructed an optimal phylogenetic network using a multispecies coalescent model implemented by curated orthologous gene trees. Our results not only elucidated the reticulated evolutionary relationships within Juglandaceae but also unveiled multiple origins of nuts from winged fruits within the family. Furthermore, our study underscores the efficacy of gene content and micro-synteny in deciphering complex phylogenetic relationships.

Pervasive incomplete lineage sorting and ancient hybridization during the early divergence of major angiosperm lineages

Bojian Zhong

Nanjing Normal University

The vast majority of extant angiosperms belong to Mesangiospermae (eudicots, monocots, magnoliids, Chloranthales, and Ceratophyllales). Resolving evolutionary relationships of the five lineages of Mesangiospermae is important for understanding the early diversifying process of angiosperms. However, the early rapid diversification of angiosperms in a short geological period complicates untangling the phylogenetic relationships among these Mesangiospermae lineages. Here, we used 177 publicly available angiosperm genomes to reconstruct the phylogeny of Mesangiospermae using multiple orthology inference approaches, character coding schemes, and data filtering criteria. Importantly, we investigated the potential causes of phylogenetic discordance and inferred phylogenetic networks to explore reticulating events among the five lineages of Mesangiospermae. Coalescent simulation analyses suggested that incomplete lineage sorting and hybridization best explain the high levels of discordance among nuclear genes in the backbone of Mesangiospermae. Cytonuclear discordance among the five Mesangiospermae lineages was observed, potentially from hybridization. Our study demonstrated that ancient hybridization largely accounts for deep phylogenetic discordances among the five Mesangiospermae lineages, but ILS cannot be entirely excluded. The phylogenetic network might better represent the relationship among the five Mesangiospermae lineages.

Phylogenetic analysis of tetraploid species using polarized genomic sequences in the presence of introgression

Martin Lascoux

Uppsala University

Phylogenetic analysis of polyploid species has long posed a formidable challenge. We developed an iterative approach based on phasing to produce separate nucleotide sequences that capture the reticulate evolutionary history of a polyploid. We tested the method extensively using simulated and real data and showed that the method allows for the correct identification of both parental species of an allotetraploid with up to 97% certainty in phylogenies with moderate incomplete lineage sorting (ILS) and 87% in phylogenies with high ILS. We then extended the method to incorporate the possibility of introgression from diploid progenitors or from other unrelated diploid or polyploid lineages. High levels of introgression can profoundly reshape a polyploid's genomic composition. In such complex polyploids, it is difficult to determine which taxa were the progenitor species and which taxa provided additional introgressive blocks through subsequent hybridization. To identify the most likely scenario we modeled polyploidization and introgression events under the multispecies coalescent and then used an approximate Bayesian computation to evaluate and compare competing polyploidization models. We then applied the method to reconstruct the history of Betula pubescens (downy birch), a tetraploid species often assumed to be of allopolyploid origin and which is known to hybridize with at least four other birch species. We show that B. pubescens is an autoploid from the common ancestor of B. pendula and its extant sister species, B. platyphylla, that originated approximately 178,000–188,000 generations ago. Extensive hybridization with at least three other species followed autopolyploidization.

Phylogenomic analyses reveal widespread hybridization events across the Potentilla s.l. genus (Rosaceae)

Tiantian Xue

Institute of Botany, Chinese Academy of Sciences

Phylogenomics has greatly advanced the reconstruction of the tree of life, but resolving evolutionary relationships among groups with frequent hybridization and incomplete lineage sorting (ILS) remains challenging. Based on 824 single-copy nuclear genes and complete plastid and mitochondrial coding genes from 187 species, this study presents the most taxon-rich phylogenomic analysis of Potentilla s.l. to date. We evaluated gene tree conflicts within the nuclear dataset and investigated the causes of incongruences, focusing on hybridization and ILS. All three datasets consistently divided Potentilla s.l. into eight highly supported clades. However, their relationships differed significantly across datasets. Multiple pieces of evidence supported a hybrid origin for the Ancistrifolia clade, explaining the observed cytonuclear discordance. Additionally, substantial conflicts between nuclear gene trees and the species tree were identified at the stem nodes of four basal clades within Potentilla s.s. and within the Argentea clade. Substantial hybridization events and significant introgression signals were detected at the conflicting nodes, suggesting gene flow as the primary driver of genealogical discordance. Mitochondrial capture associated with hybridization has occurred at least twice in Potentilla s.s., resulting in discordance between the two organellar genomes. Ancestral range estimation suggests that the most recent common ancestor of Potentilla s.l., along with the six early-diverging clades, originated on the Qinghai–Tibetan Plateau or in East Asia, likely serving as evolutionary museums for global Potentilla.

Symposium 12: Population genetics through time session A

TBA

Ali Akbari

Harvard University

Parallel adaptation in two sibling Drosophila species across global environments

Weixuan Li

Peking University

The repeatability of evolution is a central question in evolutionary biology. Comparative population genomics offers a powerful approach to reveal shared and species-specific mechanisms of adaptation. Using globally distributed Drosophila melanogaster and D. simulans as models, we conducted large-scale comparative genomic analyses to elucidate the genetic basis of environmental adaptation and assess the repeatability of adaptation.

We first assembled a high-quality reference genome for D. simulans (156 Mb, N50: 53.73 Mb), achieving superior completeness and accuracy compared to existing assemblies. Subsequently, we sequenced 544 strains (426 D. simulans and 118 D. melanogaster) from 58 locations in China and integrated these with global datasets, totaling 681 D. simulans and 1,335 D. melanogaster genomes. Population genomic analyses revealed similar population structures, subpopulation differentiation, and migration histories for both species. Notably, D. simulans exhibited a more recent migration timeline, indicating a shorter period of adaptation compared to D. melanogaster.

Natural selection signals were detected by integrating four methods, identifying both shared and species-specific genetic mechanisms of adaptation. While genes under selection were enriched in distinct pathways reflecting local adaptation, convergent evolution was evident at the pathway, gene, and variant levels. In particular, 21 genes, including Ace and Cyp6g1, showed convergent adaptation to insecticide exposure, with shared missense mutations in particular genes highlighting the genetic repeatability of environmental adaptation. In all, our findings provide insights into the interplay of determinism and contingency in evolution, underscoring the universality of genetic mechanisms underlying adaptation.

A generalized structured coalescent for purifying selection without recombination

Stefan Strütt

University of Bern, Switzerland

Purifying selection is a critical factor in shaping genetic diversity. Current theoretical models mostly address scenarios of either very weak or strong selection, leaving a significant gap in our knowledge. The effects of purifying selection on patterns of genomic diversity remain poorly understood when selection against deleterious mutations is weak to moderate, particularly when recombination is limited or absent. In this study, we extend an existing approach, the fitness-class coalescent, to incorporate arbitrary levels of purifying selection in haploid populations. This model offers a comprehensive framework for exploring the influence of purifying selection in a wide range of demographic scenarios. Moreover, our research reveals potential sources of qualitative and quantitative biases in demographic inference, highlighting the significant risk of attributing genetic patterns to past demographic events rather than purifying selection. This work expands our understanding of the complex interplay between selection, drift, and population dynamics, and how purifying selection distorts demographic inference.

Genetic signature reveals the impact of colonial history on Mtb evolution in Taiwan

Ruo-ya Bai

National Tsing Hua University

Taiwan has experienced multiple waves of colonial events over the past several centuries. Given the restricted cross-country mobility during these periods, colonial episodes may have played a significant role in facilitating pathogen transmission between nations, including Mycobacterium tuberculosis (Mtb). Previous studies have suggested a potential link between Mtb transmission and historical events, as supported by the distinct prevalence of Mtb lineages among specific ethnic populations in different regions of Taiwan. To investigate this connection, we analyzed whole-genome sequencing data from 2,655 Mtb samples collected in Taiwan between 2012 and 2021. We estimated changes in the effective population size (N_e) of Mtb over time and observed a sharp increase in N_e for lineage L2 after 1875, during the Qing dynasty—a period marked by large-scale Han immigration following the lifting of an immigration ban. Similarly, other lineages exhibited increases in N_e following colonization by various nations. Furthermore, coalescent events in phylogenetic analyses and directionality indices supported the timing and direction of transmission, respectively. These findings illustrate how colonial history shaped the evolution of Mtb in Taiwan, highlighting an interesting connection between historical events and pathogen dynamics.

Ancient Crow DNA Reveals Phenotypic Genes Underpinning a Stable Hybrid Zone

Chyi Yin Gwee

LMU Munich

Advances in ancient DNA (aDNA) technology have revolutionized our ability to extract and analyze highly degraded genetic material, offering direct snapshots of past evolutionary processes. Using modern and aDNA spanning up to 20,000 years, we investigated the evolutionary dynamics of the European hybrid zone between the carrion crow (Corvus corone corone) and the hooded crow (C. corone cornix), distributed across Eurasia. Despite stark differences in plumage—the carrion crow's all-black feathers and the hooded crow's grey-and-black pattern—these populations are genetically almost identical, with key differences localized to a single locus on chromosome 18 spanning approximately 2.5 million base pairs.

While plumage color differentiates the two forms, our analyses reveal high levels of introgression in putatively neutral genomic regions. Through aDNA analysis, we provide direct evidence that this morphological hybrid zone has remained stable for at least 5,000 years. Moreover, we identify genes on chromosome 18 that underpin plumage variation and facilitate assortative mating, shedding light on the genetic mechanisms and evolutionary forces maintaining this hybrid zone. Our study demonstrates the power of aDNA in resolving genetic changes over time, enabling the identification of key genes despite the challenges posed by high linkage disequilibrium within this region.

3,023 human genomes from mainland Southeast Asia disclose hidden genetic diversity and signatures of natural selection

Yaoxi He

Kunming Institute of Zoology, Chinese Academy of Sciences

Mainland Southeast Asia (MSEA) harbors rich ethnic and cultural diversity with a population of nearly 300 million. However, MSEA people are underrepresented in the current genomic database of global populations. Here we present the SEA3K genome dataset (Phase I), consisting of the deep short-read whole-genome sequencing (WGS) data of 3,023 individuals from 30 MSEA populations, together with the long-read WGS data of 37 representative MSEA individuals. We identified 79.59 million small variants and 96,384 structural variants (SVs), among which 22.83 million small variants and 24,622 SVs are novel. We observed a high level of genetic heterogeneity across MSEA populations, reflected by the varied combinations of genetic components. Remarkably, we discovered 44 genomic regions with strong signatures of Darwinian positive selection, covering 89 genes involved in varied body systems such as physical traits and immune response. Furthermore, we observed different patterns of archaic Denisovans introgression in MSEA populations, supporting the proposal of at least two distinct instances of Denisovan admixture into modern humans in Asia. Importantly, we detected the genomic regions indicating “borrowed” fitness in MSEA populations, due to either Neanderthal or Denisovan adaptive introgressions. The reported multi-million novel genomic variants in MSEA highlight the necessity of studying regional populations that can help answer the key questions on prehistory, genetic adaptation and complex diseases.

Population structure, dynamics, and evolution of vibrio parahaemolyticus

Jiancheng Wang

Shanghai Institute of Immunity and Infection, Chinese Academy of Science

Vibrio parahaemolyticus is a halophilic, gram-negative bacterium that thrives in warm marine and coastal environments. Recently, it has become a significant cause of food-borne illnesses worldwide, including acute gastroenteritis in humans. Characterized by a high recombination rate, V. parahaemolyticus experiences frequent homologous recombination, which intertwines variation sites introduced by recombination with those of spontaneous mutations, making it difficult to distinguish vertical genetic signals from horizontal gene transfer. Consequently, conducting population genetics studies on V. parahaemolyticus is of particular significance for both clinical and evolutionary research.

In this study, we created a global dataset comprising over 12,000 strains of V. parahaemolyticus, by which we identified five populations within this species, including a novel population named VppAfrica. Additionally, we found that population VppAsia has dispersed globally and demographic analysis indicated a recent expansion of VppAsia, which is consistent with its worldwide distribution. Subsequently, we observed several highly differentiated genes between VppAsia and other populations, which are associated with sulfur metabolism and low-temperature adaptation. Our results provide a reasonably complete picture of the population structure and genomic patterns of global V. parahaemolyticus and suggest that the ability to utilize organic sulfur may contribute to the local adaptation of this bacterium.

Somatic evolution and clonal competition in ageing human haematopoiesis

Nathaniel Mon Pere

Queen Mary University

For most organisms, the soma participates little (if at all) in its species' evolution. Yet, despite the brevity of a single lifetime, it is not exempt from the forces of evolution. Its transformation forms a fascinating counterpart to that of the species. For somatic cells the increasing mutational load is inexorably tied to the host's ageing, and unchecked progression along the path of increasing fitness ultimately brings about death in the form of cancer. Recent advances in sequencing have emphasised the ubiquity of fitness-driven evolution in human tissues. Blood presents an excellent example of this, where large clonal expansions are observed in healthy elderly individuals – a condition termed clonal haematopoiesis. How precisely this evolution unfolds remains to be understood. Not all expansions progress to disease, and the behaviours of known cancerous mutations appear highly heterogeneous across individuals.We investigated the evolutionary dynamics of the haematopoietic stem cell pool by analysing two large scale datasets of different type: one with clone trajectories from multi-timepoint sequencing, and the other single stem cell-derived colony expansions. Through population genetics modelling we showed that clonal competition plays a significant role in shaping distinctive fitness landscapes. Applying Bayesian statistics, we quantified the distribution of fitness effects and the occurrence rate of fit mutants in blood. We then used these calibrations to estimate the fitness effects of 2000 haematopoietic clone trajectories across three distinct datasets. This revealed a multi-stage mode of clone evolution, where the fittest clones only occur through accumulation of multiple beneficial mutations.

Preliminary evidence for natural selection in time-series data from Central and East Asia over the past 10,000 years

Alison Barton

Harvard University

Recent advances in methods for detecting natural selection have allowed for the discovery of orders of magnitude more targets of adaptation in West Eurasia over the past 10,000 years. This new approach utilizes a generalized linear mixed model that includes a genetic relatedness matrix as a covariate to account for population structure without the requirement of explicitly modelling demography and takes advantage of imputed ancient data to test variants genome wide. In a preliminary study, we apply this new approach to a sample set from Central and East Asia (N=587) dating from ~10,000 BP to present. Our scan recovers several previously reported hits such as positive selection on rs1229984 in the ADH1B gene (s=4.0%, P=5E-12) that decreases alcohol processing and at variants spanning FADS1/2 (s=1.7%, P=1E-10) that affect fatty acid metabolism. Another region spanning the HLA-DQB1 locus was also under positive directional selection (s=2.6% P=2E-11), which remains genome-wide significant after strict correction for genomic inflation, and is associated in Biobank Japan today with an increased risk of rheumatoid arthritis. Several sites which pass genome-wide significance only without correction are suggestive, including one spanning multiple genes including ANO8 and PLVAP and associated with monocyte count in Biobank Japan (s=1.5%, P=1E-9). We are currently expanding this study to N~3500 samples from across a broader region of Central and East Asia; based on our pilot analysis and the results from West Eurasia, we expect to find many additional targets of selection.

Fine-scale genetic structure and recent demographic history in the Ryukyu Archipelago

Masatoshi Matsunami

University of the Ryukyus

The Ryukyu Archipelago is located at southwest of the Japanese Archipelago, and is composed of dozens of islands, such as Miyako, Yaeyama, and Okinawa Islands. The principal component analyses (PCA) of Japanese using genome-wide SNPs have been shown two main clusters, Hondo (Japanese mainland) and Ryukyu. However, a detailed population structure of the Ryukyu Archipelago has not been elucidated yet. Previously, we identified the genetic background of Ryukyu population are diverse depending on island groups: Miyako, Yaeyama, Okinawa Islands, and Kerama/Kume-jima. Then, we showed Miyako population can be divided into three subpopulations based on haplotype structures, and coalescent simulation suggested these differentiations are mainly because of at least two migration wave which are dating back to about 41 and 16 generation ago using 1,240 genomic data living in Miyako Islands. In this study, we genotyped over 12,000 individuals registered in Okinawa Bioinformation Bank for 629,792 SNPs using the Asian Screening Array (Illumina, CA, U.S.A.). These data are utilized for population genomics analysis. Our preliminary haplotype-based analysis showed complex migrations within and among each island group. Especially, within Yaeyama Islands, we observed frequent migrations between Ishigaki island and other islands: Taketomi, Aragusuku, Hateruma, Iriomote, Kohama, Kuroshima, Hatoma, and Yonaguni Island. This information will shed a light on recent demographic history and is useful to infer disease associations and phenotypic prediction in focal region.

The genetic origins and impacts of historical Papuan migrations into Wallacea

Gludhug Purnomo

The University of Adelaide

The Wallacean archipelago was first settled by anatomically modern humans (AMH) by 50 thousand years ago (kya), with descendent populations thought to have remained genetically isolated prior to the arrival of Austronesian seafarers around 3.5 kya. Modern Wallaceans exhibit a longitudinal counter-gradient of Papuan- and Asian-related ancestries widely considered as evidence for mixing between local populations and Austronesian seafarers, though converging multidisciplinary evidence suggests that the Papuan-related component instead comes primarily from back-migrations from New Guinea. We present a comprehensive study of the genetic history of Wallacea and West Papua, analyzing 254 newly sequenced genomes, primarily from previously undocumented populations. Integrating genetic, linguistic, and archaeological evidence, we demonstrate that Wallacean societies were transformed by the spread of Papuan genes and languages over the past 3,500 years—coinciding with interactions between Austronesian seafarers and Wallacean-Papuan groups. Mixing between Papuan and local Wallacean lineages appear to have been confined to the western and central parts of the archipelago and likely occurred contemporaneously with the widespread introduction of genes from Austronesian seafarers—which now comprise between ~40–85% of modern Wallacean ancestry—though dating historical admixture events remains challenging due to mixing continuing into the Historical Period. In conjunction with archaeological and linguistic records, our findings point to a dynamic Wallacean population history that was profoundly reshaped by the spread of Papuan genes, languages, and culture in the past 3,500 years.

Symposium 13: Synthetic approaches to evolutionary theory: A dialog between the past and the future

Reconstructing ancestral genotype–phenotype maps: lineage-specific biases caused the evolution of functional diversity in a protein family

Joseph Thornton

University of Chicago

A classic question in our field is whether biases in the production of phenotypic variation – mutation biases, developmental biases, etc. – play a meaningful role in the outcomes of evolution. The modern synthesis assumes that the answer is no – that all phenotypes are produced with equal propensity, leaving selection as the only explanation for directional trends in evolution. But a basic understanding of biology suggests that phenotypic biases may be strong and therefore could matter to evolution. Comprehensively characterizing biases imposed by the genotype-phenotype (GP) map and their effects on evolutionary history is challenging, especially for high-level biological phenotypes that evolved in the distant past. We addressed this problem at the level of proteins and their biochemical functions by applying comprehensive combinatorial deep mutational scanning to the protein-DNA interface of two related reconstructed ancestral transcription factors. Both ancestral GP maps were strongly anisotropic (the distribution of phenotypes encoded by genotypes is highly nonuniform) and heterogeneous (the phenotypes accessible around each genotype vary dramatically among genotypes), but the extent and direction of these properties differed dramatically between the maps. In both cases, these properties steered evolution toward the lineage-specific molecular phenotypes that evolved during history. Our findings establish that ancient properties of the GP relationship were causal factors in the evolutionary process that produced present-day patterns of functional conservation and diversity.

The evolutionary potential of bacterial gene regulatory elements: insights from synthetic biology and adaptive landscapes

Cauã Antunes Westmann

University of Zürich

Transcriptional regulation is fundamental to life's diversity, yet the origins and evolution of transcription factor binding sites (TFBSs) remain poorly understood. In prokaryotes, transcription factors (TFs) bind to TFBSs to control gene expression, forming gene regulatory networks that coordinate cellular responses. Here, we investigate the evolutionary potential of TFBSs by integrating massively parallel reporter assays with adaptive landscape theory. To study the potential for the de novo evolution of TFBSs, we mapped genotype-to-phenotype relationships for hundreds of thousands of TFBS variants across five bacterial TFs with distinct regulatory scopes, from local regulators that control a few genes to global regulators that govern many. While all adaptive landscapes were highly rugged with a few accessible high peaks, key differences emerged based on regulatory scope. For instance, transcriptional robustness - how stable gene expression levels remain despite single mutations in a TFBS - progressively decreased from local to global regulators, likely reflecting the need for global TFs to maintain stable control over diverse targets. Beyond de novo evolution, we explored whether TFBSs can be exapted from one landscape to another, allowing recognition by a different TF. We show that a strong TFBS for one TF can evolve through a few adaptive mutations to a site recognized by another TF, with intermediate states recognized by both. These findings provide the first in vivo assessment of de novo and exaptive TFBS evolution in bacteria, offering insights into broader evolutionary questions such as adaptation on rugged landscapes and functional conflicts arising during exaptation.

Connecting distant proteins with similar structures: a path perspective

Kaiyu Qiu

Max Planck Institute for Biology Tübingen

The evolutionary distinction between homologous (divergent) and analogous (convergent) proteins remains challenging when both exhibit low sequence identity yet high structural similarity. We present a path-based framework combining protein generative modeling and structure prediction to map connectivity between distant proteins, addressing critical gaps in our understanding of protein sequence-structure relationships.

Our analysis progresses through two stages. First, we systematically compared natural protein pairs with matched sequence divergence, generating paths using family-specific variational autoencoders. Paths connecting analogs showed distinct signatures: transient destabilization (lower pLDDT), reduced evolutionary plausibility (lower language model likelihoods), and impaired structural decodability (lower inverse folding probability). Notably, some analog paths maintained the global fold despite passing through low-confidence regions, suggesting potential cryptic homology.

Building on these insights, we further evaluated whether de novo-designed proteins represent true analogs against their natural counterparts by developing general path-generation methods that bypass MSA requirements. Generating paths guided by protein language models or by backpropagating through AlphaFold, nearly all designed-natural pairs (even with <10% sequence identity) were connected by smooth, structurally conserved intermediates. This demonstrates that most de novo proteins reside within naturally connected regions of protein space rather than forming structural outliers.

Our path-centric approach reveals previously hidden connectivity patterns that challenge conventional homology detection thresholds, provides operational criteria for distinguishing evolutionary relationships from structural convergence, and recontextualizes de novo protein design within the protein space continuum. By enabling systematic navigation through dark regions of protein space, this framework advances our understanding of protein evolution while offering new paradigms for synthetic protein engineering or design.

Functional optimization in distinct tissues constrains the rate of protein evolution

Dinara Usmanova

Columbia University

Understanding the main determinants of protein evolution is a fundamental challenge in biology. Despite many decades of active research, the molecular and cellular mechanisms underlying the substantial variability of evolutionary rates across cellular proteins are not currently well understood. It also remains unclear how protein molecular function is optimized in the context of multicellular species and why many proteins, such as enzymes, are only moderately efficient on average. Our analysis of genomics and functional datasets reveals in multiple organisms a strong inverse relationship between the optimality of protein molecular function and the rate of protein evolution. Furthermore, we find that highly expressed proteins tend to be substantially more functionally optimized. These results suggest that cellular expression costs lead to more pronounced functional optimization of abundant proteins, and that the purifying selection to maintain high levels of functional optimality significantly slows protein evolution. We observe that in multicellular species both the rate of protein evolution and the degree of protein functional efficiency are primarily affected by expression in several distinct cell types and tissues. Specifically, in developed neurons with upregulated synaptic processes in animals and in young and fast-growing tissues in plants. Overall, our analysis reveals how various constraints from the molecular, cellular, and species' levels of biological organization jointly affect the rate of protein evolution and the level of protein functional adaptation.

Temperature dependence of protein fitness landscape topography

Shizhe Tang

Westlake University

Temperature influences life at all levels, yet its effect on the genotype-fitness relationship remains poorly understood. We combined fluorescence-activated cell sorting with deep sequencing to quantify the fluorescence of 109,856 yellow fluorescent protein variants and the folding stability of 33,280 functional variants across temperatures ranging from 15 to 42°C. We found that high temperatures lower protein stability, and amplify both the stability effects of mutations and the fitness effects of stability changes. Consequently, the same mutations induce larger fitness changes and tend to interact additively at high temperatures. This "low-stability large-effect" smooths the landscape by reducing sign epistasis and the number of fitness peaks. Our findings suggest that temperature can shape the fitness landscape topography and thus influence the exploration of sequence space.

Can a nonadaptive genetic system become indispensable?

Yun-Ju Lee

University of Michigan

RNA editing is a post-transcriptional process that modifies RNA sequences, resulting in insertions, deletions, or base substitutions. This can change the original genetic information from DNA. One prominent type of RNA editing, the A-to-I editing which converts adenosine to inosine, is performed by the ADAR (adenosine deaminase acting on RNA) enzymes. The ADAR gene family was found in almost all metazoan lineages, and it is likely that the ADAR gene emerged in the common ancestor of the animals. Selectionists would explain that ADAR is beneficial so that it is conserved, because it increases transcriptomic and proteomic diversities or alters gene regulation without disrupting the genome. However, an alternative hypothesis was proposed that ADAR may be constrained by nonadaptive manner. This hypothesis suggests that once A-to-I RNA editing emerged, it allows deleterious G-to-A mutations (ribosomes recognize I as G on mRNAs) to be fixed in the genome by drift. When such substitutions accumulate and exceed a certain level, the cell becomes fully dependent on the RNA editing mechanism to survive. To test the hypothesis, I introduced the mallard duck ADAR gene into the baker yeasts, Saccharomyces cerevisiae. The yeasts were further treated with several rounds of ethyl methanesulfonate to facilitate G-to-A mutations in the genomes. Once the ADAR gene became indispensable to yeasts, the RNA editing profile of yeasts was examined. The results may support the idea that an initially neutral or even deleterious genetic system can be constrained after it has been maintained for many generations.

Symposium 14: The roles of hybridization and introgression in evolution

Hybrid speciation over space and time in the Amazon rainforests

James Mallet

Harvard University

Heliconius elevatus and H. pardalinus are sister species of butterflies that overlap geographically over a huge range across the entire Amazon basin. They still hybridize occasionally, which has led to homogenization of ~98% of their genomes. However, the strongly diverged ~2% of the genome determines colour pattern differences and other speciation traits such as pheromones and mate choice. We have shown that H. elevatus is a hybrid species obtained these divergent loci via hybridization and back-crossing with a third species, H. melpomene. Here I use biogeographic analysis of our genomic data to support a scenario of hybrid speciation and geographic spread over space and time. In doing so, we provide a rare example of how population spread is likely involved during the history of any speciation event.

Inferring introgression based on deep learning approach

Yubo Zhang

Peking University

Gene flow is widespread in nature and strongly influences the tree of life. Identifying introgression signals between taxa remains a challenge in the genomics era. Machine learning, especially deep learning, has provided a powerful framework for solving population genetic problems. To increase the accuracy and flexibility of introgression detection, and reduce the computational complexity, we have developed a deep learning-based approach, ERICA (Evolutionary Relationship Inference using a CNN-based Approach). The approach takes multiple sequence alignments as input, and directly outputs evolutionary relationships based on the pre-trained convolutional neural networks, and identifies local introgression signals based on the topological discordance. Evaluation of multiple simulated datasets with different evolutionary scenarios showed the ERICA approach was effective in detecting gene flow, particularly in adaptive introgression, although the accuracy was also affected by the demographic history, like other methods. We further applied ERICA on the real genomic data of animals and plants, including the Heliconius butterflies and the wild and cultivated rice. We identified the introgressed regions across the genome and revealed the important role of introgression in adaptation. In summary, we have provided a new method for evaluating evolutionary relationships and detecting gene flow using genomic data, which may contribute to a deeper understanding of hybridization and introgression.

Complex motility adaptation in a Vibrio parahaemolyticus ecospecies driven by introgression

Sarah Svensson

Shanghai Institute of Immunity and Infection

Some bacterial species exchange DNA frequently enough that they have well-mixed geographic gene pools as in outbreeding animals and plants. Nevertheless, natural selection can generate genome-wide genetic structures that we have termed "ecospecies" (Tourrette et al. 2024, Nature). In Vibrio parahaemolyticus (VP), ecospecies EG1a is defined by over 40 genes coinherited on otherwise diverse genetic backgrounds. Ecospecies differentiation has occurred both through point mutation and acquisition of genetic material from sympatric Vibrio species. Laboratory investigation of the EG1a phenotype shows that unlike other VP, it cannot swarm on surfaces but instead induces and uses its differentiated lateral flagella in viscous liquids, where it swims better than non-EG1a strains. Transposon-sequencing (Tn-seq) and RNA-seq in this swimming condition showed that the majority of genes that differentiate EG1a strains are involved in this phenotype, which thus seems to have driven ecospecies divergence. Combining evolutionary and phenotypic analysis, including knock-in of an EG1a accessory gene, we propose that EG1a divergence was precipitated by introgression of genes for flagella glycosylation, and culminated with remodelling of the cell surface by acquisition of an S-layer. EG1a divergence also involved concomitant/subsequent differentiation of core genes, including an antibacterial Type 6 secretion system coregulated with, and affected by, lateral flagella-driven swimming. Our “reverse ecology” approach has shed light of the evolution of a key Vibrio trait and revealed the advantages of bacterial inheritance mechanisms for adaptation, compared to meiosis. VP has been able to maintain coherent gene pools and species structure, while occasionally importing evolutionary novelties from related species.

Origin and maintenance of a shared sexual mimicry polymorphism

Tristram Dodge

Stanford University

The importance of long-term balancing selection [LTBS]—which maintains variation over long evolutionary timescales—remains controversial. Because neutral polymorphisms rarely coalesce deeper than speciation, genetic trans-species polymorphisms are often used to identify loci under LTBS. While this model assumes that variation arose in an ancestor and has been maintained in the daughter lineages after speciation, several alternative scenarios—including introgression, convergent evolution, or turnover of genetic control—could also explain the presence of shared polymorphic phenotypes and influence their detectability at the genetic level. We studied a shared male-limited coloration trait, the "false gravid spot," segregating in a dozen Xiphophorus fish species diverged by ~6 million generations. This trait mimics the female pregnancy spot, and our behavioral experiments suggest it is balanced by male aggression and female preference. In four species, we map the false gravid spot to cis-regulatory elements upstream of kitlga, a conserved pigmentation gene. These haplotypes show phylogenetic discordance suggesting multiple introgression events. Notably, not all species share this pattern. In one clade, the false gravid spot maps to a repetitive region on a distinct chromosome enriched with zinc-finger proteins. Instead of representing a convergent origin, our data suggest this locus interacts with kitlga in trans to repress its expression, causing turnover of genetic control. Together, our findings indicate that a simple model of LTBS on ancestral variation is inaccurate. Instead, introgression and turnover contribute substantially, which may influence the ability to detect LTBS using traditional molecular approaches, and potentially led to a misestimate of its importance.

Evolutionary insights from nine-spined stickleback genomes: demography and admixture

Xueyun Feng

University of Helsinki

Pleistocene glaciations shaped the genetic structure of Northern Hemisphere organisms by isolating populations during cold periods and facilitating gene flow during warm phases. Using 888 whole-genome sequences, we explored the complex phylogeography and demographic history of nine-spined sticklebacks (Pungitius pungitius), focusing on two divergent European lineages and their broad admixture zone across the southern Baltic Sea. Our findings reveal a complex history characterized by multiple waves of migration and introgression, with freshwater isolates providing unique insights into intermediate evolutionary stages and revealing previously unrecognized population dynamics. Demographic analyses uncovered glacial contractions across all populations, with post-glacial expansions in marine populations contrasting with declines in freshwater populations. Differences in effective population size trajectories suggest that demographic features have substantially influenced current patterns of genetic diversity. A closer analysis of the evolutionary dynamics of hybrid genomes across the Baltic Sea uncovered varying selection pressure on the introgressed variation. Functionally important regions had resisted introgression, while promoter regions showed enrichment. Despite widespread negative selection against introgressed variants, specific genomic regions showed signals of adaptive introgression, suggesting introgression plays a role in the adaptation to the conditions of the Baltic Sea. Selection against functional changes was strongest near the secondary contact zone, diminishing with distance. The results provide a detailed understanding of the demographic history of nine-spined sticklebacks across northern Europe and offer new perspectives on how hybridization and admixture shape the evolutionary trajectories of natural populations. This study introduces the nine-spined stickleback as a versatile model for investigating the evolutionary impacts of hybridization.

A 120-year time-series of genomes reveals the consequences of closed breeding in German Shepherd Dogs

Lachie Scarsbrook

University of Oxford

In the last few centuries, dogs have experienced significant declines in overall genomic health, including reduced heterozygosity and increased load, largely linked to breeding practices. Lack of historical genomes, however, has made it difficult to disentangle the impacts that breed-specific demography, artificial selection, and crossbreeding have had on the genomic health of modern breed dogs. Here, we sequenced genomes from nine 20th-century museum specimens (1906–1993) of pedigree German Shepherd Dogs (GSD) to directly assess the consequences of contemporary breeding practices on genomic diversity and health. The GSD offers an excellent case study, given the breed's fluctuations in global popularity, differing selective regimes, and the GSD's use in the establishment of other breeds, including wolfdogs. Genome-wide heterozygosity showed significant reductions after the Second World War, coincident with an increase in both the frequency of runs of homozygosity and load. We also detected repeated population bottlenecks linked to the use of popular sires throughout the 20th century. While most contemporary GSDs traced their ancestry to individuals born in the mid-1960s, identical-by-descent allele sharing with historical GSDs was both geographic (e.g. America vs. Switzerland) and lineage (e.g., working vs. conformation) specific. Finally, although hybridisation with wolves has led to genome-wide increases in heterozygosity in wolfdog breeds, non-admixed ancestry blocks (dog or wolf) were severely depleted in diversity due to the limited number of founders involved in their establishment. Combined, our results indicate that declines in GSD genomic health and related breeds occurred throughout the last century, rather than during breed formation.

Evolution of viviparity promotes species divergence：The genetic mechanism in common lizards

Hongxin Xie

University of Glasgow

The innovation of viviparity is associated with faster species diversification in multiple vertebrate lineages. However, the underlying genetic mechanism is largely unknown. We made used of a reproductively bimodal species, Zootoca vivipara, to understand the evolution of viviparity and its association with lineage divergence using whole genomes and geographically wide-spread sampling. We show that viviparity originated once in this species complex and is associated with subsequent lineage diversification and demographic expansion that is not mirrored in the co-existing oviparous clades. Gene flow is ongoing among lineages including between parity modes, and we found regions across the genome that resist homogenisation and strongly reflect the evolutionary history of parity mode. Viviparity occurred by divergent selection on many genomic regions over a long time in the ancestral viviparous population. Gene flow barrier regions are enriched for functional viviparous genes differing in regulatory elements between oviparous and viviparous lineages, with effect on gene expression, and not in any protein-coding sequences. Our results reveal sustained contribution of viviparity to lineage divergence by selection on many loci in the genome.

Subgenome-level phylogeny reveals the reticulate effects of allopolyploidization on angiosperm evolution

Yu Cao

Huazhong Agricultural University

Whole-genome duplications (WGDs) have played a crucial role in the adaptation and diversification of angiosperms. Most angiosperm lineages have experienced at least one ancient WGD event after the divergence of angiosperms, except for Amborella from the ANA grade and Aristolochia from the magnoliids. Unlike autopolyploidization, allopolyploidization entails both duplication and hybridization, which complicatesthe dichotomy of lineage divergence. Currently, the impact of allopolyploidization on the reticulate origin ofangiosperms remains largely unexplored. In this study, we construct a multilabel phylogeny of major angiosperm lineages at the subgenome level by resolving genes from different subgenomes. For the five major lineages that have undergone one additional WGD (eudicots, monocots, Chloranthales, Austrobaileyales and Nymphaeales), we selected one subgenome per lineage to reconstruct the phylogeny. Our phylogentic results uncovered two distinct topologies within Mesangiospermae. The discrepancy lies in the position of Chloranthales, which is sisiter to either eudicots or magnoliids. Monocots are consistently recovered as the sister group to the last common ancestor of all the other Mesangiospermae. Further analysis of gene trees at the subgenome level shows that, unlike other lineages, the two subgenomes of Chloranthales rarely form a sister group with each other. These findings suggest that the allopolyploidization history of Chloranthales may be reticulated with other major angiosperm lineages. This study is the first to highlight the significance of incorporating allopolyploidization into the reticulate evolution of angiosperms.

The role of introgression in the sympatric speciation of Quercus acutissima and Q.variabilis

Mingming Zhang

Zhejiang University

The likelihood of speciation in the face of gene flow is one of the most debated topics in evolutionary biology. Introgression can be an important source of new mutations for host species, but also a homogenizing force that prevents populations from speciation and diverging into separate species. While many studies have discovered speciation under gene flow, its evolutionary mechanisms still remain unclear. In this study, we used resequencing data of 796 individuals to investigate the sympatric speciation between two sympatric widespread Asian oak species, Quercus acutissima and Q. variabilis. Asymmetric gene flow were detected from Q. variabilis to different populations of Q. acutissima, led to highly differentiated genetic structures in latter species. At introgressed and highly genetic differentiated genomic regions together with effects of selective sweep, we identified speciation related candidate genes of flowering time controling, flower morphology and pollen morphology development which had most likely contributed to the establishment of pre-zygotes reproductive barriers. Suppressed recombination due to chromosomal inversion prevented introgressd genes being breaken down and might lead to speciation. Finally, results from comprehensive speciation models confirmed the importance of introgression in shaping SFS with excessive high frequency alleles. In summary, this study has explored the evolutionary mechanisms of how two sympatric and wide distributed oak species diverged under the complicated evolutionary scenario combined with genetic drift, natural selection, and more importantly, asymmetric introgression.

Machine learning-based population genetics analyses reveal historical and genomic impacts of introgressed variants on grape breeding

Hua Xiao

Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences

Grapevine is among the earliest domesticated crops. Wild grapes are frequently used in breeding as sources for key agronomic traits. However, gene flow (introgression) from wild grapes to cultivars can introduce both beneficial and deleterious variants. Our understanding of the consequences, genetic architecture, and traits affected by introgression remains limited. To address this, we analyzed resequencing data from wild (Vitis vinifera ssp. sylvestris) and cultivated (Vitis vinifera ssp. vinifera) grapes. Machine learning-based population genetic analyses predicted the timing, mode, genomic pattern, and biological effects of introgression events. We found continuous gene flow between European wild grapes (EU) and cultivated grapes over the past ~2000 years, particularly from EU to wine grapes. Gene pathways related to aromatic compound synthesis were enriched in regions both selected and introgressed, indicating that EU wild grapes have been crucial for enhancing grape flavor. Despite these benefits, introgressed fragments increased the deleterious burden, with most deleterious SNPs and structural variants (SVs) hidden in a heterozygous state, raising the genetic load in wine grapes and complicating future breeding efforts. Forward simulations of introgressed allele dynamics suggest that introgression has historically been significant for the improvement of clonal systems like domesticated grapes. Our study of the dual effects of introgression is vital for leveraging wild resources in genomic breeding of grapevines.

Multiple divergent Denisovan ancestries and adaptation of introgressed copy number variants in mainland Southeast Asian populations

Yini Xiao

Fudan University

Mainland Southeast Asia (MSEA), situated at the crossroads of South Asia, East Asia, and Island Southeast Asia, provides substantial genetic sources due to its environmental distinctiveness for investigating evolutionary history and its phenotypic implications. East Asian and South Asian populations are known to have interbred with multiple Denisovan groups, yet the landscape of archaic introgression in MSEA populations remains poorly explored. In this study, we analyzed 2110 individuals across 30 MSEA populations to detect archaic introgressed variants, including single nucleotide variants (SNVs) and copy number variants (CNVs). We identified Denisovan introgressed alleles and observed different patterns of Denisovan gene flow across MSEA populations. Specifically, three populations from Laos

exhibit two separate Denisovan introgressed components, while a Vietnamese population shows three distinct pulses of Denisovan admixture. We further estimated the timing of these introgression events, traced their origins of the Denisovan lineages, and discovered the significant impact of introgressed alleles on adaptive evolution. Moreover, leveraging multiple CNV detection methods and high-coverage whole-genome sequencing data, including long-read technologies, we identified population-stratified CNVs shared with archaic genomes in MSEA populations. These introgressed CNVs are shown to play a pivotal role in the adaptive evolution of MSEA populations, particularly in genes associated with nervous system development and immune function. We provide novel insights into the complex population history of MSEA, emphasizing the role of archaic introgression in shaping both functional adaptations and phenotypic diversity in modern humans.

The role of introgression in driving repeated adaptation between species of section Costatae and Betula platyphylla

Nian Wang

Shandong Agricultural University

Polyploid species can acquire novel genetic materials from their diploid progenitors or closely related species via hybridization and introgression, which may facilitate adaptation to similar environmental conditions. In this study, we utilized population-level genomic data to reconstruct speciation event of species within section Costatae of the genus Betula in China. Sect. Costatae includes two tetraploid species B. utilis and B. ermanii and three diploid species B. ashburneri, B. costata and B. buggsii. We modeled polyploidization and introgression events under the multispecies coalescent framework and applied a simulated annealing algorithm to evaluate and compare various polyploidization models. Additionally, we investigated the genetic repeatability of adaptive introgression within Section Costatae species and B. platyphylla through genotype-environment association analysis and ABBA-BABA analysis. Our results provide statistical evidence for genetic repeatability across diploid and tetraploid species within Section Costatae and B. platyphylla. Some of this repeatability appears to have originated from putative introgression from other closely related species. These introgressions are associated with climatic adaptation phenotypes, such as cold- and thermo-tolerance, as well as flowering time regulation. Taken together, our study supports the role of adaptive introgression from related species in driving parallel local adaptation to environmental conditions

The genomic insights of the early stages of sympatric homoploid hybrid speciation in crater lake cichlid fishes

Axel Meyer

University of Konstanz

In vertebrates sympatric homoploid hybrid speciation is very rare but also the mechanisms for generating the reproductive isolation are disputed. Midas cichlid fishes in Nicaraguan crater lakes are an example for sympatric speciation. Here we analyze more than two hundred resequenced genomes of a presumed species species as well as multiple morphological dataset and conducted laboratory mating experiments to investigat the underlying mechanisms of homoploid hybrid speciation. We show that the hybrid lineage is genetically and behaviorally isolated from both parental species. This can be shown although it is of very recent (a few hundred generations coalescent time) origin and although backcrosses exist in nature. The morphological traits that differentiate the hybrids appear to be reshuffled resulting in a transgressive phenotype. They are related to diet and swim capability, might play a key role for the isolation allowing it to occupy unique niche. Correspondingly, genomic and selection analyses, suggest that it based more likely on a combination of standing variation of the two parental species, rather than de novo mutation within the hybrid lineage, that led to the emergence of this new hybrid species.

Symposium 15: Transposable elements in genome evolution and biotechnology development

TBA

Josefa Gonzalez

Institut Botànic de Barcelona, CSIC

Adaptive transposable elements leads to phenotypic variation via introducing cis-elements in Arabidopsis

Yalong Guo

Institute of Botany, CAS

TEs are repetitive DNA fragments that can transpose across the genome and constitute a large portion of the genome in many organisms. Transposition of TEs can affect gene function and genome stability, promote phenotypic divergence, and contribute to speciation and adaptation. High throughput genome sequencing speeds up the studies of transposable elements. Nevertheless, it is still largely unknown about the effect of TEs on the phenotypic variation and adaptation, which is a fundamental biological question. Here we focused on one TEs (STES) that is under positive selection in natural populations. STES is a 732-bp retrotransposon located at the first intron of the AtSL gene. The presence of STES enhances the expression level of AtSL by introducing two cis-regulatory elements, thereby increasing both silique length. Ecological niche modeling suggests that STES likely plays a significant role in the adaptation of Arabidopsis. Overall, this study elucidates the manner in which an adaptive TE can affect its host genes and phenotypic variation, highlighting the significance and potential impact of TEs on adaptation in general.

Giant genomes of lungfishes grew by massiv TE expansion

Axel Meyer

University of Konstanz

The genomes of living lungfishes can inform on the molecular-developmental basis of the Devonian sarcopterygian fish-tetrapod transition. We de-novo sequenced the genomes of the African (Protopterus annectens) and South American lungfishes (Lepidosiren paradoxa). The Lepidosiren genome (~91 Gb, ~30x the human genome) is the largest animal genome sequenced so far and more than twice the size of the Australian (Neoceratodus forsteri) 1 and African2 lungfishes due to enlarged intergenic regions and introns with high repeat content (≈90%). All lungfish genomes continue to expand since some transposable elements (TEs) are active still today. Particularly, Lepidosiren's genome grew extremely during the last 100 million years (MY), adding the equivalent of one human genome every 10 MY. Massive genome expansion appears to be related to a reduction of piRNAs and C2H2-zinc-finger and KRAB-domain-protein genes that suppress TEs expansion. Although TE abundance facilitates chromosomal rearrangements, lungfish chromosomes still conservatively reflect the Ur-Tetrapod karyotype. Neoceratodus' limb-like fins still resemble their extinct relatives and remained phenotypically static for ~100MYA. We show that the secondary loss of limb-like appendages in the Lepidosiren-Protopterus ancestor is likely due to loss of sonic hedgehog (Shh) limb-specific enhancers.

Exploring endogenous retroviruses in ruminant genomes: They might not be all dead after all

Emmanuelle Lerat

Universite Claude Bernard Lyon 1

Endogenous retroviruses (ERVs), traces of ancestral retroviral infections, constitute a significant portion of mammalian genomes and play crucial roles in host evolution. Some ERVs have retained their coding capacities, enabling them to act as active transposable elements, capable of duplicating within host genomes. While extensively studied in humans, ERVs remain underexplored in livestock, notably in ruminants. Using de novo repeat identification approaches, we characterized ERVs from 24 families across domestic and wild sheep and goats, and cattle reference genomes. Thirteen families were shared across the five species, eight were exclusive to small ruminants and three to cattle revealing multiple endogenization events over the last 40 million years. A high-resolution annotation allowed the identification of 100,534 ERV insertions, representing 0.5–1% of ruminant genomes, with varying copy numbers across species. Notably, the domestic goat exhibited a higher number of copies from two specific ERV families. Population-level analyses of over 100 goat genomes from 37 breeds uncovered abundant low-frequency copies from these two families, indicative of recent insertions. Whereas the youngest family, includes insertions with complete coding capacities and is closely related to oncogenic exogenous retroviruses; the second family displays incomplete elements; suggesting different duplication mechanisms and potential ongoing endogenization. These results demonstrate the recent transpositional activity of ERVs in goats and highlight unique evolutionary dynamics among ruminants. This study underscores the importance of ERVs as models for understanding species evolution and host-virus interactions and calls for further research into their impact on the livestock genomic landscapes.

The horizontal escape of a human retrotransposon into one contagious virus genome reveals the schizophrenic life of BC200

Cheng Sun

Capital Normal University

Horizontal transfer (HT) is the passage of genetic material between non-mating species. Transposable elements (TEs) have inherent ability to mobilize and proliferate within genomes, which might be especially prone to HT. Over 200 cases of HT of TEs have been documented so far, and HT of TEs was considered to be an essential step in their lifecycle. However, to date, the involvement of human TE in HT was poorly documented. Here we present the first documented HT of a human TE into an extremely contagious poxvirus. We found that one Alu element was transferred from human to the human poxvirus Molluscum contagiosum virus by retrotransposition. Further analysis indicated that BC200, a domesticated Alu element and functions in human neural cells, was the master gene for this retrotransposition event. In addition, our mobilization assay indicated that BC200 was still retrotranspositionally competent in human culture. Finally, we found that BC200 was the master gene for a group of Alu elements in every primate where it exists, and the generation of new insertions by BC200 was an ancient and most possibly, an on-going process. Taken together, our results uncovered a horizontal escape of human retrotransposon into one contagious poxvirus and revealed that BC200 should be a novel human mutagen causing somatic mutation. Our study emphasized a previously overlooked function of the domestication of TEs, which should be another method employed by TEs to attain their successful proliferation.

Transcriptional and translational signatures of ovarian aging in Drosophila melanogaster based on transposable elements

Qiheng Xu

Peking University

Ovarian aging means a decline in oocyte quality and quantity, which is closely related to genomic instability, epigenetic changes and mitochondrial dysfunction associated with transposable element (TE) jumping. piRNA acts as a regulatory factor to inhibit transposons and regulate gene expression, but its specific role in aging has not been widely explored. In this study, the regulation between piRNA and transposon and its molecular mechanism were discussed from the perspective of ovarian aging in Drosophila melanogaster. The results showed that with ovarian aging, DNA repair-related pathways were down-regulated, leading to the accumulation of genomic and mitochondrial damage, a decline in mitochondrial activity, an increase in chromatin openness, and ultimately, increased transcriptional activity of TEs and a higher number of insertions. This abnormal TEs activation further threatens genomic stability and drives the cellular aging process by disrupting gene expression and translation accuracy. In addition, as an important mechanism of transposon inhibition, studies have shown that increasing piRNA expression levels only partially limits the activity of TEs, and the regulatory ability of different transposon types is different. At the same time, the targeting effect of piRNA on protein-coding genes also causes body disorders and further promotes aging. These findings provide a new perspective for understanding the molecular mechanism of reproductive aging and may provide a theoretical basis for investigating interventions for diseases associated with ovarian aging.

Activation and evolution of endogenous retroviruses in gall bladder

Jie Cui

Fudan University

Gallbladder inflammation, especially gallbladder cancer, is a rapidly progressing disease. During our research on the evolution of endogenous retroviruses, a type of LTR-retrotransposon, we found that these elements become abnormally active in inflamed gallbladders. We classified these elements and studied the relationship between viral heterogeneity at the single-cell level and inflammation, developing a single-cell typing system based on endogenous retroviruses. We discovered that the expression of viruses such as HERVH is related to the occurrence and development of the disease. We also verified the possible regulatory functions of these elements by experimental studies. We hope that the observed phenomena can be used for immunotherapy of gallbladder related diseases.

Detection of plant LTR tetrotransposons through extrachromosomal circular DNA sequencing: Insights into transposition and novel non-canonical proteins

ILIA KIROV

All-Russia Institute of Agricultural Biotechnology

Extrachromosomal circular DNAs (eccDNAs) are enigmatic DNA molecules found in various organisms, including plants, where they exhibit diverse genomic origins and can be derived from transposable elements. However, the structures of individual eccDNA molecules and their dynamics in response to stress remain poorly understood. In this study, we performed nanopore sequencing of eccDNAs from different plant species, including Arabidopsis thaliana, Helianthus annuus, Nicotiana benthamiana, Brassica napus, and Solanum lycopersicum, grown under conditions that induce changes in DNA methylation or the RNA-directed DNA methylation (RdDM) pathways. We discovered that long terminal repeat (LTR) retrotransposons are among the most active producers of eccDNA. Using this methodology, we identified and experimentally validated new active LTR retrotransposons in tomato, sunflower, tobacco, and rapeseed plants. The active LTR retrotransposons found in sunflower, belonging to the Tekay clade, possess open reading frames (ORFs) for conserved non-canonical proteins present in both sunflower and other plant species. This was revealed through protein structure prediction, domain analysis, and conservation studies. The structural features of these non-canonical proteins resemble those of plant DRBM-containing proteins, suggesting their potential role in RNA-silencing modulation. Our findings demonstrate that active plant LTR retrotransposons encode additional non-canonical proteins, paving the way for their functional analysis and potential biotechnological applications in the future.

DNA transposons mining and engineering for efficient CAR-T modification

Chengyi Song

Yangzhou University

Cut-and-paste DNA transposons are widely recognized as natural and stable non-viral gene delivery vectors that offer several advantages, including low cost, low immunogenicity, low cell toxicity,and large cargo capacity. These transposons have found broad applications in gene therapy, transgenesis, and mutagenesis studies. After conducting large-scale genomic data mining and activity screening using a cell-based transposition assay, we identified over 4,000 cut-and-paste DNA transposons from more than 20 families, belonging to seven superfamilies (hAT, piggyBac, PHIS, Tc1/Mariner, DD82E/Sailor, pogo, Gambor/Hiker). Based on sequence analysis and K divergence of the selected transposons, we submitted 5 Spy/PHIS, 3 PiggyBac, 9 hAT, 3 pogo, one Mariner/Tc1, and 3 DD35E/Hiker for transposition activity evaluation. Among them, we identified 10 highly active transposons: PS from pogo, ZB from Tc1/Mariner, miCN and lrCN from Cleaner/hAT, anBuster from Buster/hAT, PB13, PB84, and PB108 from PiggyBac, and cgSpy and cvSpy from Spy/PHIS. We further characterized the cargo capacity, overproduction inhibition (OPI), and insertion preferences of PS, ZB, and cgSpy. Furthermore, we achieved significant improvements in transposition efficiency and cargo capacity by further engineering and optimizing the PS transposon systems. Additionally, the enhanced PS (renamed JL) transposon system, in combination with the Tini-plasmid technique (small vector backbone, approximately 500 bp), significantly improves the efficacy of CAR-T cells engineered even at low doses, resulting in impressive therapeutic outcomes in hematological and solid tumors. This evidence underscores the promise of the optimized JL transposase, paired with the condensed donor vector, as an efficacious and secure strategy for gene therapy in clinical practice.

Roles of TEs in modulating secondary metabolite biosynthesis in plants

Li Wang

Chinese Academy of Agricultural sciences

Plant secondary metabolites play a crucial role as biochemical signals in response to abiotic and biotic stress. However, the involvement of transposable elements (TEs) in modulating secondary metabolite biosynthesis remains poorly understood. Here we present two empirical examples illustrating how TEs can either hinder or promote secondary metabolite biosynthesis.

In the first example, pervasive inversions in Curcuma genomes were found to be enriched with TEs compared to the genomic background, as they were associated with suppressed recombination and relaxed purifying selection. The proliferation of TEs within these inversions led to the pseudogenization of TPS genes, ultimately impeding terpenoid biosynthesis in Curcuma.

In contrast, TEs appear to play a beneficial role in facilitating complex coumarin biosynthesis in Apiaceous medicinal plants. Specifically, long terminal repeat-retrotransposons (LTR-RTs) may act as molecular vehicles, guiding the assembly of a complex coumarin biosynthetic gene cluster (BGC). This cluster includes key enzymes such as C2'H, the gatekeeping enzyme C-PT, and modifying enzymes like O-PT, OMT, and CYP. The coordinated activity of these enzymes results in the high production of complex coumarins in the Apioideae subfamily. Additionally, TE exonization likely contributes to the emergence of novel catalytic enzymes, such as C-PTs, which are essential to produce complex coumarins.

In summary, TEs play a pivotal role in shaping biochemical traits and enabling plants to adapt to challenging environments.

Transposable elements as drivers of genome evolution and functional divergence following whole-genome duplication in orychophragmus

Jing Wang

Sichuan University

Whole-genome duplication (WGD) is a major evolutionary event that reshapes genomes and creates opportunities for genetic innovation. Following WGD, transposable elements (TEs) play a critical role in genome evolution by influencing the retention, structural diversification, and functional divergence of duplicated genes. In this study, we investigate the role of TEs in driving genome evolution and the neo- or sub-functionalization of duplicated genes in Orychophragmus violaceus. Compared to related species, our comprehensive genome, transcriptome, and epigenetic analyses reveal that TEs have proliferated extensively in Orychophragmus violaceus following its unique WGD, leading to significant genome restructuring. Our results demonstrate that TE insertions near duplicated genes are enriched in regulatory regions identified by ATAC-seq, where they contribute to transcriptional rewiring and the emergence of novel regulatory elements. Through gene expression and methylation profiling, we show that TEs influence the divergent expression patterns of duplicated gene pairs, facilitating their sub- or neo-functionalization by mediating chromatin modifications. Many of these genes are associated with stress responses, metabolic pathways, and developmental processes, and we experimentally validated the functions of several of them. Taken together, by combining genomic, transcriptomic, and epigenomic approaches, this study provides novel insights into how TEs shape the evolutionary trajectories of WGD-derived genomes and promote the functional diversification of duplicated genes.

History of transposable element invasions in drosophila melanogaster

Robert Kofler

VetMedUni

Largely based on phenotypic effects generated by transposable elements (TEs), previous work revealed that three TEs invaded Drosophila melanagaster in the last century.

Based on the genomes of historical specimens, more than 600 publicly available short read data sets and novel bioinformatics approaches, we substantially extend this invasion history. We show that in total 11 TEs spread in D. melanogaster during the last 200 years. These 11 invasions increased the genome by 1 Mbp, thereby accelerating the rate of genome evolution about 3800-fold as compared to base substitutions.

Leveraging data from over 1400 arthropod genomes, we suggest that the TE invasions were triggered by horizontal transfers, with D. simulans and species of the D. willistoni group acting as donors.

We further show that TEs can penetrate populations extremely fast. For example, Transib1 spread from three isolated epicentres in 2014 to all investigated populations within just two years.

Our findings suggest that anthropogenic activities, facilitating habitat and population expansions of D. melanogaster, might have accelerated the rate of horizontal transposon transfer as well the spread of the TEs into the worldwide population.

Given the significant impact of TEs and the potential involvement of humans in their dispersal, our research has important implications for both evolution and ecology.

Transposable elements reconcile micro and macroevolution for adaptive genome changes

Ying Liu

Hebei university

Transposable elements (TEs) are generally major component of a genome and contribute in large extent to evolution of genome.However, it remains a great challenge to reconcile micro and macroevolution given the huge pool and complex dynamics of TE types or families in genomes. We here attempted to address this issue using a notorious worldwide migratory locust, Locusta migratoria. The locust has a large genome of up to 6.5G bp, and TEs constitute approximately 65% of the genome. Based on a high-quality genome of the locust at the chromosome level, we first analyzed the evolutionary trajectories of TEs in eight orthopteran species to explore their impact on genomic evolution. These results indicate that genome rearrangement breakpoints occur in regions with high TE content. Several TE superfamilies have propagated in the locust genome and are highly correlated with genome size variation. TE insertion polymorphisms (TIPs) are a prevalent type of genomic structural variant. We then conducted whole-genome resequencing of six geographic populations from different altitudes, and analyzed genomic landscape of TIP variation. TIPs can differentiate these populations into two distinct lineages, i.e., plateau and plain groups. Thousands of adaptive TIPs associated with latitude selection were identified. Several propagative TE subfamilies contributed to the adaptive TIPs. Finally, we functionally validated a PiggyBac-mediated selective sweep that resulted in alternative splicing of the local genes. These results demonstrate that TEs are crucial players in macro and microevolution of genomic changes that shape genome landscape and drive adaptive changes.

Heterologous survey of DNA transposons in human cells highlights their functional divergence and expands the genome engineering toolbox

Shengjun Tan

Institute of Zoology, CAS

Experimental studies on DNA transposable elements (TEs) have been limited in scale, leading to a lack of understanding of the factors influencing transposition activity, evolutionary dynamics, and application potential as genome engineering tools. We predicted 130 active DNA TEs from 102 metazoan genomes and evaluated their activity in human cells. We identified 40 active (integration-competent) TEs, surpassing the cumulative number (20) of TEs found previously. With this unified comparative data, we found that the Tc1/mariner superfamily exhibits elevated activity, potentially explaining their pervasive horizontal transfers. Further functional characterization of TEs revealed additional divergence in features such as insertion bias. Remarkably, in CAR-T therapy for hematological and solid tumors, Mariner2_AG (MAG), the most active DNA TE identified, largely outperformed two widely used vectors, the lentiviral vector and TE-based vector SB100X. Overall, this study highlights varied transposition activity and evolutionary dynamics of DNA TEs, and increases the TE toolbox diversity.

Symposium 16: Thematic symposium: evo-devo

How to build a placenta from scratch: novel organ evolution in fish

Sophie Kraunsoe

The Francis Crick Institute

Differences in development underly the incredible diversity of vertebrate organs. However, understanding how developmental programs have evolved to create species differences is challenging due to the ancient origins of most organs. One exception is the recent, independent evolution of placentas in the fish family Poeciliidae. In contrast to egg-laying teleosts, Poeciliids fertilise their eggs internally and give birth to live offspring. Fertilisation takes place without ovulation with the ovarian follicle continuing to surround the embryo until birth. Remarkably, in some species, this novel interaction between the ovarian follicle and the embryonic yolk sac has evolved into a placenta, with the ovarian follicle transitioning from an ancestral function of oocyte provisioning to transferring nutrition directly to the embryo. Combining single-cell RNA sequencing, spatial transcriptomics and imaging, we have characterised the ovarian cell type diversity in Poeciliids. We identified that the ovarian follicle of oogonia is formed by classical granulosa and thecal cells, but these cells are later replaced by a novel Poeciliid-specific cell type. This cell type differentiates before fertilisation, likely from granulosa cells and forms the maternal structures of the placenta. In some species, it undergoes significant proliferation, creating intricate villi that expand the placental surface area for nutrient exchange. The presence of the Poeciliid-specific cell type in the ovaries of both placental and non-placental species indicates a new organ can be created by repurposing existing cell types into novel structures. Our findings reveal how a key cellular innovation enabled live-bearing and was subsequently co-opted multiple times for placenta development.

The relationship between molecular evolution and mechanisms underlying unique anatomical development and extreme radiotolerance in tardigrades

Dong Yang

National Center for Protein Sciences (Beijing), AMMS

Tardigrades, commonly known as water bears, are captivating organisms for their resilience in extreme environments. We investigated the role of genes with distinct evolutionary traits in its special development pattern and exceptional radiotolerance based on omics data.

We discovered that the evo-devo pattern of Hypsibius exemplaris (He) basically follows the 'hourglass model', with an exception that the late embryonic stage showed a transition from non-conserved to conserved state. He also showed a simplified pattern of some classical pathways during development. Some of the tardigrade-specific genes (TSGs) may participate in glutathione and lipid metabolism, which may be associated to the ecdysone synthesis and tardigrade-specific storage cell formation. (Evol Dev, 2024)

To explore their molecular basis of radiotolerance, multi-omics strategies were employed for Hypsibius henanensis (Hh), a novel species identified in our study. From the evolutionary viewpoint, the radiotolerance mechanisms were summarized into three types: Firstly, horizontally transferred genes endow tardigrades with special stress-resistance abilities. A DOPA dioxygenase gene DODA1, confers radiotolerance by synthesizing betalains with free radical scavenging properties. Secondly, proteins encoded by TSGs tend to be highly disordered. A tardigrade-specific radiation-induced disordered protein, TRID1, facilitates DNA damage repair through a mechanism involving phase separation. Thirdly, some ancient proteins exhibit a unique radiation response pattern in tardigrades. The mitochondrial respiratory chain complex assembly protein BCS1 underwent widespread gene expansion in the genomes of various tardigrades. BCS1 was significantly upregulated after irradiation, promoting mitochondrial NAD+ regeneration and accelerating NAD+-dependent damage repair. These three observations expand our understanding of mechanisms of tardigrade radiotolerance. (Science, 2024)

Origin of cell differentiation in animals

Victoria Shabardina

Institute of Evolutionary Biology

Over the past two decades, studies of the unicellular relatives of animals have greatly advanced our understanding of the origins of animal complexity, including cell differentiation and development. For instance, it was shown that some transcription factors involved in embryonic development predate the emergence of animals. In this study, we use the ichthyosporean Abeoforma whisleri as a model to explore the origins of cell differentiation in the animal lineage. Abeoforma exhibits a morphologically diverse non-linear life cycle. Using microscopy, culture manipulation, bulk RNA sequencing, and single-nuclei RNA sequencing, we identify and characterize its distinct life stages. Notably, we describe, for the first time, a highly polarized cell morphology with a complex cytoskeleton and vesicular infrastructure. These cells form at a specific time point and environmental cues, and carry the function of motility and dispersion. After performing their task, the cells die. This is reminiscent of the terminal cell differentiation in animals, the phenomenon previously undocumented in unicellular organisms. Our findings suggest that the unicellular ancestor of animals had a complex life cycle regulation involving animal-like molecular factors and with prerequisites to cell differentiation. This highlights the evolutionary principle of co-opting existing genes and pathways into new functions and physiological processes, highlighting this important mechanism in the evolution of animals.

Shared and Distinct Molecular Programs in Developing Human and Big Animal Brains

Mingfeng Li

Huazhong University of Science and Technology

The mouse is widely regarded as an ideal model for studying molecular and cellular principles analogous to those in humans. However, the central nervous system (CNS) exhibits significant differences between mice and humans, largely due to disparities in gestation periods and brain structure. To address these limitations, our lab has undertaken a decades-long genomic project utilizing large animals such as monkeys and pigs to compare developmental stages and anatomical regions in these species with those of humans, i.e., various cortical and subcortical regions from early embryonic stages to adulthood. Through the development of integrative and comparative analysis methodologies, we uncovered a key finding characterized by the hourglass model, which demonstrates the progression of regional diversification shared between humans and monkeys during development, alongside a distinct W-shaped divergence between the two species (Zhu et al. 2018; Li et al. 2018). Additionally, we generated single-cell transcriptome data for monkeys to explore species conservation and divergence, with a focus on identifying the spatiotemporal profiles of cellular and molecular programs (Micali et al. 2023). Our findings revealed minimal transcriptomic alterations during interneuron migration within the ganglionic eminence subdivisions, whereas cortical tangential migration exhibits dramatic transcriptomic changes, highlighting destination-specific molecular specification and species divergence (Feng et al. 2025). Overall, our research highlights the potential of large animal models as valuable proxies for understanding the relationship between cellular diversity and functional progression, providing new insights into the study of human brain development and evolution.

Evo-Devo of flowering transition, floral color and symmetry in neotropical orchids

Yesenia Madrigal Bedoya

New York University

With ca. 29,000 species, orchids exhibit a huge diversity in flowering responses and adaptations to local microclimates. In addition, they have highly modified showy flowers, with early bilateral symmetry, organ elaboration and fusion, and a plethora of colors resulting in a floral ground plan fit for ornamental purposes. Although the genetic bases of reproductive transitions, floral organ identity and symmetry, as well as floral color have been well documented in model species, less is known about them in Neotropical orchids. During the last decade our lab has developed transcriptomic resources from ca. 15 species of non-model neotropical orchids. We have also performed exhaustive gene evolution analysis, evaluated comparative expression of candidate transcription factors for all three processes and studied their protein interactions in selected species. Our results suggest orchid-specific adjustments in the genetic pathways controlling flowering and floral symmetry because of lineage specific duplication events. When comparing reproductive transitions, we identified the retention of key flowering integrators and their interactions in three species with distinct habits, very different flowering seasons, and unique pre-flowering requirements. We also noted that orchids, both temperate and tropical seem to have lost canonical vernalization genes. Conversely, floral symmetry genes deviate extensively from those previously identified in core eudicots. As per floral color we identified the key paralogs that form the MBW canonical anthocyanin promoting complex in leaves versus flowers. Finally, we propose Epidendrum fimbriatum, a miniature neotropical terrestrial orchid as an emerging model system to assess the function of candidate genes in Orchidaceae.

Taurine-Driven Chemotaxis and Metamorphosis in Proto-Vertebrate Ascidian

Likun Yang

Ocean University of China

Settlement of marine invertebrate larvae at suitable sites for metamorphosis, growth, and reproduction is crucial for population expansion, also causing ecological problems such as bioinvasion and biofouling. Chemosensation plays an essential role in marine invertebrate larval settlement preference. However, the mechanisms by which these larvae perceive various chemical cues to elicit distinct behaviors that shape their communities remain unknown. Using urochordate ascidian larvae, the prominent marine fouling organisms affecting coastal ecosystems, we explored the mechanism of larval chemosensation and its role in settlement preference. Here, we identified taurine, a specific sulfur-containing amino acid secreted from marine adult animals, as a nutritional sensor for ascidian larvae to locate salubrious environments for metamorphosis. Taurine stimulates palp sensory neurons, and this neuronal excitation can be integrated into an evolutionarily conserved chemosensory neuronal circuit in ascidian larval ‘head’ to elicit chemoattraction and attachment of swimming larvae. This work offers a significant contribution to the field of marine Eco-Evo-Devo research by establishing a model to study marine larval behaviors and developmental regulation from both ecological and evolutionary perspectives. It also provides novel insights into developing anti-fouling techniques by targeting the process of taurine-driven chemotaxis and metamorphosis to prevent biofouling and ecological collapse.

Stepwise Co-option of Gene Regulatory Networks Drives Episodic Radiations in Flies

Yaoming Yang

Westlake University

Understanding how organisms acquire morphological novelty and complexity remains a central challenge in evolutionary biology. The role of gene regulatory networks (GRNs) in facilitating the emergence of novel structures is still poorly understood. The pupal prothoracic respiratory organ in dipteran insects, hypothesized to be a serial homolog of the insect wing, offers a unique model for exploring how changes in GRNs contribute to morphological innovation. Here, we investigate the evolutionary and developmental origins of this organ. Transcriptomic and genetic analyses in Drosophila melanogaster and Aedes albopictus reveal that this organ shares core GRNs and a serial homologous developmental origin with the wing, confirming it as a modified prothoracic wing homolog. We demonstrate that its development is shaped by the GRNs governing appendage development in basal lineages, followed by the incorporation of tracheal GRNs in more derived fly lineages. Using ancestral state reconstruction based on genomic and morphological data from over 200 species, we show that this organ first emerged approximately 232 million years ago (MYA) in the common ancestor of Diptera and underwent a major modification with increased complexity around 66 MYA in the common ancestor of Schizophora. These key morphological innovations align with the Carnian Pluvial Episode and the K-Pg mass extinction, respectively, suggesting a connection between these evolutionary events and the diversification of dipterans. These findings highlight the stepwise co-option of GRNs as a fundamental mechanism driving morphological innovation and complexity during major geological events, contributing to the evolution of novel adaptive structures over macroevolutionary timescales.

Less, but More: New Insights from Appendicularians on Gene Loss and Evolutionary Innovations

Gaspar Sanchez Serna

Okinawa Institute of Science and Technology (OIST)

Gene loss is widespread across the tree of life, yet its impact on the diversification of taxa and the emergence of evolutionary innovations remains poorly understood. Here, we use the Fibroblast-Growth-Factor (FGF) signaling pathway as a case study to investigate how gene loss has shaped the evolution of tunicates—the sister group of vertebrates—whose evolution within the chordate phylum is marked by extensive gene losses. Through phylogenetic analyses, examination of gene and protein structures, and a comprehensive study of developmental FGF expression in the appendicularian tunicate Oikopleura dioica, we reconstruct the evolutionary history of appendicularian FGF genes. Appendicularians have lost all chordate FGF subfamilies except FGF9/16/20 and FGF11/12/13/14, which may constitute the minimal chordate FGF toolkit; moreover, these surviving subfamilies have undergone an extraordinary expansion and diversification. This genetic rearrangement offers insights into key evolutionary innovations that distinguish appendicularians from other tunicates, such as their filter-feeding house and the absence of drastic metamorphosis and tail absorption—traits underpinning their distinctive fully free-living lifestyle. Consequently, we propose a new conceptual framework, “less, but more”, as an extension of the classical “less is more” model to cases where gene loss coincides with gene duplications. This scenario is characterized by four outcomes: (I) conservation or (II) loss of ancestral functions, (III) function shuffling between lost and duplicated genes, and (IV) emergence of novel functions after the duplications. Our results provide fresh insights into the evolutionary mechanisms underlying gene loss and establish a novel model for understanding its role in shaping species evolution.

Symposium 17: Thematic symposium: regulatory genomics and non-coding elements

Mistranslation reduces error load in evolving proteins through negative epistasis with mistranscription

Xiaoyi Zhang

Sun Yat-sen University

Phenotypic mutations have increasingly been recognized for their critical role in evolution at various levels, ranging from improving protein stability to facilitating the emergence of novel traits in cancer. However, there is little knowledge about the relationship between the two major types of phenotypic mutation, namely mistranscription and mistranslation. In this study, genome-wide analyses of mistranslations and mistranscriptions was conducted in five model organisms using mass spectrometry and circ-seq data, respectively. Unexpectedly, genes with frequent mistranslations are consistently found to have a lower mistranscription rate, a pattern we hypothesized as caused by potential negative epistasis between the two types of errors. We tested our hypothesis by systematic experimental measurements of within-gene epistasis between genetic mutations. Such epistasis is predominately negative, which makes proteins affected simultaneously by both types of errors significantly more deleterious than expected. More importantly, an in silico simulation of molecular evolution suggests that the extra deleterious effects caused by the negative epistasis, when scaled by the error rates, can facilitate selection against mistranscription for genes with frequent mistranslation. Collectively, our results reveal a previously unknown interaction between mistranscription and mistranslation due to their negative epistasis.

Decoding the evolutionary origins of open chromatin using deep learning in Drosophila

Li Zhao

The Rockefeller University

The evolution of genomic regulatory regions is critical to shaping phenotypic diversity across species. While chromatin accessibility and conformation are largely sequence-dependent, the complexity of biological systems obscures how novel regulatory elements emerge over evolutionary timescales. A key unresolved question is how open chromatin states—and their associated regulatory sequences—originate and evolve. To address this, we developed a deep learning framework to dissect the sequence determinants of chromatin accessibility across Drosophila species and tissues. We trained hybrid convolutional-attention neural networks to predict ATAC-seq peaks using only local DNA sequences as input. Strikingly, these models generalize across evolutionarily divergent insect species, revealing that sequence rules governing accessibility are highly conserved. Through interpretable machine learning, we identified both novel and known DNA motifs predictive of accessibility, linking sequence features to regulatory function. In this talk, I will expand on these results to investigate the long-term evolutionary dynamics of open chromatin origination, providing insights into how enhancers, promoters, and other regulatory elements emerge genome-wide. Our findings demonstrate deep conservation in the sequence basis of chromatin architecture and highlight the robustness of regulatory logic over evolutionary time. Our work also underscores the power of deep neural networks to resolve fundamental questions in regulatory evolution.

Accelerated evolution in the human lineage shaped regulatory networks controlling the expression of key neurodevelopmental genes

Lucia Franchini

INGEBI-CONICET

An enduring objective of evolutionary biology is to decode how changes in gene regulatory networks contribute to human-specific traits. Comparisons of the human with other vertebrate genomes allowed to identify Human Accelerated Regions (HARs), highly conserved noncoding sequences that have a significant excess of human-specific nucleotide changes and that are prime candidates to encode novel human molecular features. We have found that neurodevelopmental genes such as the transcription factor NPAS3 and the splicing regulator RBFOX1 are hot spots of human-specific evolution since both genes accumulate large numbers of HARs, suggesting that they play a role in human brain evolution. We employed transgenic reporter assays to determine RBFOX1-HARs and NPAS3-HARs are neurodevelopmental transcriptional enhancers. Moreover, 10 RBFOX1-HARs and 4 NPAS3-HARs exhibit loss or gain of function during embryonic development when compared to their chimpanzee ortholog sequences. We selected two RBFOX1-HAR (2xHAR.321/2xHAR.505) and two NPAS3-HARs (HAR202/2xHAR.142) that showed loss or gain of function to generate genetically engineered mouse models. We found changes in the expression of Rbfox1 and Npas3 at different brain developmental stages in the mutant mice. RNA-seq analysis revealed a high number of differentially expressed genes and changes in the alternative splicing of several key genes involved in the synaptic machinery in the engineered Rbfox1 mice.

Our work shows that the accelerated evolution process that RBFOX1 and NPAS3 underwent in the human lineage shaped their gene regulatory networks and represents one step further to fill the gap between human-specific genetic changes and their phenotypic consequences.

Transcriptome and epigenome atlas of the rhesus macaque brain reveal molecular architecture changes during evolution

Xuerui Zeng

KUNMING INSTITUTE OF ZOOLOGY .CAS

The rhesus macaque serves as a vital model species for investigating phenotypic innovations during primate evolution and human origin. However, research on its epigenome significantly trails behind that of humans and mice. A comprehensive transcriptomic and epigenetic landscape of the rhesus macaque brain is crucial for gaining in-depth insights into brain evolution and function. In this study, we describe the transcriptomes and epigenomes of fetal monkey brains across nine developmental stages, ranging from embryonic day (E40) to birth. We also generated transcriptomic data for 21 brain regions of adult and aged monkeys, together with the epigenome profiles of ten brain regions. The epigenomic data encompass whole-genome methylation, chromatin accessible regions, histone modifications, and CTCF binding. We conducted an in-depth analysis of the transcriptomic and epigenomic characteristics of the rhesus macaque brain, and we dissected the patterns associated with development and brain regions. We identified primate-specific and human-specific epigenomic patterns. We found that the evolution of gene regulatory networks plays an important role in shaping the primate brains. We also discovered that epigenomic research on rhesus macaques aids in understanding the risk loci and pathogenic mechanisms of human brain diseases. Our results enhance the understanding of the baseline transcriptome and epigenome of the rhesus macaque brain and contribute to expanding its application in studying human brain evolution and diseases.

Epigenomic responses to long-term snowpack decline in a hibernating marmot

Stavi Tennenbaum

Princeton University

In temperate climates, gene regulatory mechanisms allow hibernating species to enter and emerge from hibernation on the optimal date to align with snowpack depth. Plasticity in length and emergence time from hibernation is a key molecular adaptation in species within seasonal ecosystems, yet how these mechanisms will endure perturbations under climate change remains unexplored. A better understanding of how regulatory mechanisms may respond to environmental variability is urgent, as hibernating species are particularly vulnerable to global snowpack declines. Previous research indicates that hibernating species exhibit extreme variation between summer and winter in DNA methylation, a well-studied regulatory mechanism, suggesting the potential for flexible epigenomic responses to facilitate species survival under new climate conditions. To assess historical trends in DNA methylation, we leveraged 165 longitudinal blood samples from 37 yellow-bellied marmots (Marmota flaviventris), a high altitude specialist in western Colorado, USA over 17 years. We evaluated individual level genome-wide DNA methylation in response to annual snowpack and temperature using remote sensing data downscaled to marmot habitats. Individuals exposed to earlier snowmelt date in spring emerged earlier from hibernation and showed increased DNA methylation compared to those emerging later in years with delayed snowmelt, opposing age-related DNA methylation trends. We cultured primary dermal fibroblasts from marmots to validate gene expression modulation within differentially methylated regions, finding candidate genes regulating metabolism and heat shock proteins. Epigenomic responses with regulatory effect to earlier snowmelt underscore the need to integrate organismal, population and ecosystem-level data to understand the current and future impacts of environmental change.

Thermal acclimation and parental effects: The role of microRNAs in a coral reef fish

Lucrezia Celeste Bonzi

The University of Hong Kong

Besides the key role of small non-coding RNAs in various fundamental biological processes such as animal development, cell division, and stress response, their potential involvement in epigenetic inheritance is yet to be fully elucidated. Indeed, some classes of these epigenetic regulatory mechanisms, for example microRNAs (miRNAs), might play a role in non-genetic trait transfer and in acclimation strategies that are passed on from parents to the next generations. Here, we benefited from the ability of a coral reef fish to acclimate to elevated water temperature over a generation to investigate the possible involvement of miRNAs in such thermal transgenerational plasticity. By exposing mothers, fathers, and offspring of the spiny damselfish Acanthochromis polyacanthus to increased water temperature simulating future climate change conditions, we were able to analyze offspring livers for miRNA expression changes due to the offspring's own thermal exposure experience, as well as the ones of their two parents. Our findings therefore not only characterize the miRNA repertoire involved in thermal acclimation in this species, but also investigate the shared and unique contributions of each generation and parent to this epigenetic response. Overall, this study sheds light on the significance of miRNAs in a coral reef fish's thermal acclimation and offers new insight into the role of non-coding RNA in transgenerational plasticity, a potentially vital mechanism for species resilience in the face of environmental challenges.

Vitellogenin receptor mediates heat adaptability of oocyte development in mud crabs and zebrafish

Muhua Wang

Sun Yat-sen University

Climate-driven warming affects the reproduction of oviparous ectotherms. However, whether oviparous ectotherms possess a protection mechanism against heat stress for oocyte development, which is essential for maintaining the continuity of animal populations, is largely unknown. Under high temperatures, female mud crabs (Scylla paramamosain) typically have well-formed ovaries, while a few crabs were found to experience oocyte development failure. To investigate the heat stress protection mechanism of oocyte development in mud crabs, we constructed a chromosome-level genome of this species and identified an enhancer of the vitellogenin receptor (VtgR) that stimulates its expression under high temperature. Lacking this enhancer due to an intronic deletion leads to low VtgR expression in abnormal crabs, resulting in abnormal vitellogenic oocyte formation in these individuals when exposed to high temperatures. Furthermore, a similar heat stress protection mechanism was identified in zebrafish. Disruption of Lrp13, a VtgR-like protein in zebrafish, resulted in impaired vitellogenin absorption and ovarian degeneration in zebrafish exposed to high temperatures. Our results reveal a VtgR-mediated mechanism that protects vitellogenic oocyte formation against heat stress in mud crabs and zebrafish, contributing to their heat adaptability during oocyte development.

Canalization and de-canalization in organisms -when the transcriptome is constantly but weakly perturbed

Guangan Lu

Sun Yat-sen University

Recent studies have increasingly pointed to microRNAs (miRNAs) as the agent of gene regulatory network (GRN) stabilization as well as developmental canalization against constant but small environmental perturbations. To analyze mild perturbations, we construct a Dicer-1 knockdown line (dcr-1 KD) in Drosophila that modestly reduces all miRNAs by, on average, ~20%. The defining characteristic of stabilizers is that, when their capacity is compromised, GRNs do not change their short-term behaviors. Indeed, even with such broad reductions across all miRNAs, the changes in the transcriptome are very modest during development in stable environment. By comparison, broad knockdowns of other regulatory genes (esp. transcription factors) by the same method should lead to drastic changes in the GRNs. The consequence of destabilization may thus be in long-term development as postulated by the theory of canalization. Flies with modest miRNA reductions may gradually deviate from the developmental norm, resulting in late-stage failures such as shortened longevity. In the optimal culture condition, the survival to adulthood is indeed normal in the dcr-1 KD line but, importantly, adult longevity is reduced by ~90%. When flies are stressed by high temperature, dcr-1 KD induces lethality earlier in late pupation and, as the perturbations are shifted earlier, the affected stages are shifted correspondingly. Hence, in late stages of development with deviations piling up, GRN would be increasingly in need of stabilization. In conclusion, miRNAs appear to be a solution to weak but constant environmental perturbations.

Pervasive latent selection in the human noncoding genome

Swetha Ramesh

UCLA

Comparative genomics approaches that detected constrained sequences across distantly related species estimate that approximately 5% of the human genome is under purifying selection. Most of this constrained sequence lies in the noncoding portion of the genome. It is often assumed that mutations in these sequences are deleterious across the phylogeny and mutations outside these regions are neutrally evolving. However, sites that have biological function in certain lineages but not in others, i.e. functional turnover, may be invisible to comparative genomics approaches. This latent, or hidden, selection from functional turnover obscures the true amount of selection in the genome and its effects are still unclear. Here, we use a series of simulations to show that there are statistically significant differences when comparing the site frequency spectra for simulations with and without latent selection. The magnitude of this difference depends on the amount of total sequence, the amount of sequence under latent selection, and the sample size. We employ the same statistical comparison technique on human genetic variation data from the 1000 Genomes Project to estimate the amount of latent selection in noncoding, non-conserved sequence by comparing it to quiescent sequence. We find that 3-5% of unconstrained bases in the human genome are experiencing latent selection, which is on par with previous estimates of functional turnover. Our findings demonstrate that overlooked latent selection in noncoding sequence can lead to erroneous inferences in downstream applications.

Quantifying and characterizing the rate and strength of adaptation in human genomic regulatory sequences in response to viruses

Addison Lander

University of Arizona

It is well-established that pathogens, especially viruses, are dominant drivers of genomic adaptation in the hosts they infect. There has been much research investigating this dynamic, particularly in humans, and yet the majority of this research focuses on quantifying adaptation in genomic sequences that are translated into proteins. It has been shown, however, that mammalian genomes experience more adaptation at loci that house regulatory features. We hypothesize that there is a significant amount of virus-driven adaptation occurring across the regulatory landscape. In this study, we investigate the rate and strength of virus-driven adaptation in human genomic regulatory sequences using an Approximate Bayesian Computation method applied to the McDonald-Kreitman test (ABC M-K), and we compare its contribution to the overall genomic adaptation to viruses relative to coding sequences. Given that regulatory sequences do not possess synonymous mutations, we modify the M-K test to circumvent these restrictions. In doing so, we investigate the utility of a comparative metric, "relative alpha", as a means of assessing relative contribution to overall genomic adaptation of regulatory sequences compared to coding sequences. We further investigate whether alignment quality affects estimates of positive selection through the implementation of a consensus alignment pipeline. Initial findings suggest that there is a substantial amount of regulatory adaptation that has occurred in humans in response to epidemics in deep time.

The error hypothesis of RNA variations: evidence and applications

Chuan Xu

Shanghai Jiao Tong University

RNA variations refer to differences in RNA sequences that arise from various mechanisms, such as alternative transcription, RNA editing, RNA modifications, etc., contributing to transcriptome diversity and functional complexity. Although RNA variations are ubiquitously observed in eukaryotes, and some have been shown to possess important functions, it remains unclear whether most of them are functional or not. Based on the principles of genetics and evolution, we proposed the error hypothesis, which posits that RNA variations primarily result from molecular error and are therefore mostly nonadaptive. Through a series of analyses, we previously tested and validated this hypothesis. In our further in-depth studies, we then found additional evidence supporting the error hypothesis from a population genetics perspective and introduced new case studies. Moreover, although most RNA variations are non-adaptive and functionally neutral, we leveraged them to develop a novel method for identifying functional RNA variations. While our findings reinforce the notion that the majority of RNA variations are non-adaptive, they also present a new approach for uncovering rare, functionally important RNA variations, thereby advancing research on gene regulation.

The landscape of fitness effects of putatively functional noncoding mutations in humans

Chenlu Di

University of California, Los Angeles

While annotations of functional non-coding regions in the human genome are becoming more precise, the fitness effects of mutations in these regions remain unclear. We developed a pipeline to infer the distribution of fitness effects (DFE) of non-coding mutations using polymorphism data from the 1000 Genomes Project. Our analysis of enhancers, promoters, and conserved non-coding regions revealed distinct selection patterns. While mutations in enhancers are often neutral, approximately 30% of mutations in promoters are deleterious. The most conserved non-coding regions, showing reduced divergence across mammals and primates, have the highest proportion of deleterious mutations. However, primate-specific conserved sites have a higher proportion of deleterious mutations than regions that are conserved across mammals, but less conserved in primates specifically. Notably, we infer the most conserved sites across mammals, frequently targeted in biomedical studies, account for only a minority of deleterious mutations in noncoding regions. For example, the top 5% of conserved non-coding sites encompass fewer than 20% of deleterious mutations, indicating that functional non-coding regions are widely distributed across evolutionary constraints. Our findings highlight the dynamic evolution of gene regulation and shifting selection pressures over deep evolutionary timescales. These insights have broad implications for using comparative genomics to identify non-neutrally evolving sequences in the human genome.

Massively parallel characterization of common and rare SNP pairs on gene regulation

Rita Kreevan

University of Tartu

Interactions between single nucleotide polymorphism (SNP) regulate gene expression and contribute to variation in complex traits. Models of complex trait architecture have assumed that nearby SNP effect are independent. However, there are indirect evidence that under stabilizing selection, linkage masking can occur, whereby haplotypes containing linked SNPs with opposite effects on a trait have reduced effects on fitness and escape negative selection.

We explored this by directly examining the effects of proximal SNP-pairs on gene expression level. By employing an episomal massively parallel reporter assay (MPRA), we investigated the functional interaction effects within 14,570 rare (AF <0.05) and common (AF > 0.1 SNP pairs identified from ~2,400 Estonian whole-genome sequencing (WGS) samples from the Estonian Biobank. In the selection criteria we focused on SNP pairs within tissue-specific blood enhancers (12,686 pairs) and broad enhancers active in over 85 cell/tissue types (1,884 pairs). The SNP pairs, separated by distances of up to 75 bp, included at least one cis-eQTL (-log10(p) ≥ 15) identified in eQTLGen. We designed a library of 30,000 oligos representing these SNP pairs and transfected it into two biologically relevant cell lines: the GM12878 lymphoblastoid cell line and the K562 immortalized myelogenous leukemia cell line. We expect this MPRA will help to decipher the functional impact of noncoding variants and provide new insights into the regulatory mechanisms of SNP-SNP interactions.

Accelerated cell-type-specific regulatory evolution of the human brain

Dennis Manjaly Joshy

Yi lab, UC Santa Barbara

The molecular basis of human brain evolution is a key piece in understanding the evolution of human-specific cognitive and behavioral traits. Comparative studies have suggested that human brain evolution was accompanied by accelerated changes of gene expression (referred to as regulatory evolution), especially those leading to an increase of gene products involved in energy production and metabolism. However, the signals of accelerated regulatory evolution were not always consistent across studies potentially due to the diversity of distinctive cell types in the human brain. Here, we leveraged single-cell human and nonhuman primate transcriptomic data to investigate regulatory evolution at cell-type resolution. We relied on six well-established major cell types: excitatory and inhibitory neurons, astrocytes, microglia, oligodendrocytes, and oligodendrocyte precursor cells. We discovered pervasive signatures of accelerated regulatory evolution in human brains compared to chimpanzees. These signatures also demonstrated high cell-type specificity, as are the evolutionarily differentially expressed genes (DEGs) that are influenced by them. Our results support the accelerated cell-type specific functional programs as an important feature of human brain evolution. Furthermore, we investigate the effectiveness of canonical cell type markers and discuss results that suggest the involvement of human-specific evolution in age-dependent expression trends of these canonical cell type markers.

A-to-I RNA editing frequently impedes ribosome elongation to regulate protein synthesis in Drosophila

Xinning Dong

Peking University

Adenosine-to-inosine (A-to-I) RNA editing is an evolutionarily conserved mechanism. Inosine is typically recognized as guanosine by cellular machinery such as ribosomes, while the impact of inosine on decoding efficiency during translation remains unclear. Using Drosophila, which exhibits extensive RNA editing in coding regions, we observed frequent ribosome stalling and reduced protein production at these sites. The reduction in protein production was confirmed by mass spectrometry analysis and in vitro translation experiments. Profiles of Adar5G1 mutant flies further support translation inhibition by editing events, independent of RNA structures, codon usage bias and amino acid identities. Notably, ribosome stalling at the Adar S>G nonsynonymous site adds another layer to the auto-feedback loop of ADAR activity, indicating multi-level regulation of the global editome. Phenotypic experiments on AdarG and AdarS mutants with different adenosine deaminase activities confirm the adaptive significance of RNA editing. Our work uncovers the stalling at the S>G recoding site of Adar adds an additional layer of regulatory complexity, affecting both the editomes and transcriptomes, and thereby influencing phenotypes. We propose that A-to-I RNA editing serving as an elongation brake at functionally important sites, thereby regulating ribosome elongation. This work deepens our understanding of the biological significance and functional impact of A-to-I RNA editing in Drosophila, where editing events are prevalent in coding regions.

Symposium 18: Thematic symposium: frontiers in single-cell level analysis

Advanced bioinformatics tools for single-cell somatic mosaicism analysis and visualization

Yanmei Dou

Westlake University

Somatic mosaicism, the presence of genetically distinct cell populations within an individual, influences human development, aging, and diseases. To address the complexities of somatic mosaicism, we have developed bioinformatics tools for high-resolution analysis and visualization, enabling precise detection, lineage tracing, and advanced visualization of mosaic variants.

BayesMonSTR: Single-Cell Somatic Mosaicism of Short Tandem Repeats: BayesMonSTR leverages Bayesian inference to detect somatic mosaicism in short tandem repeat (STR) loci at single-cell resolution. It enables accurate identification of haplotype-resolved mosaic mutations, providing insights into STR dynamics across various biological contexts.

Precise Detection of Mosaic Mutations in Single-Cell Omics Data: This tool accurately detects mosaic mutations in single-cell omics data, integrating lineage information with multi-modal data for a comprehensive understanding of genetic diversity at the single-cell level.

PhyloSOLID: Error-Tolerant Phylogenetic Tree Construction: PhyloSOLID is an error-tolerant method for constructing phylogenetic trees from single-cell data, facilitating robust lineage tracing despite sequencing errors.

Advanced Visualization of Mosaic Variants: Our visualization tool offers detailed views of mosaic variants in single-cell sequencing data, revealing the context and nature of mosaic mutations.

Together, these tools provide a comprehensive pipeline for somatic mosaicism analysis, advancing our understanding of genetic diversity and cell population dynamics, with potential applications in human development, aging, and disease research.

A lineage-informed deep neural network for cellular dynamics inference using single-cell transcritpome and lineage tracing data

Yuhong Wen

University of Chinese Academy of Sciences

The advent of single-cell transcriptomics and lineage tracing technologies has opened new avenues for understanding complex dynamic processes in multicellular systems, such as cell differentiation and tumorigenesis. However, integrating and modeling gene expression and lineage tracing data over time remains a significant computational challenge, particularly for large-scale datasets. To address this, we introduce **DeepTracing**, a lineage-informed deep neural network designed to infer cellular developmental dynamics from time-series single-cell RNA sequencing and lineage tracing data. By coherently integrating gene expression, lineage tracing, and temporal information, DeepTracing accurately reconstructs cell differentiation trajectories and captures cross-clonal interactions. This framework provides a robust and scalable solution for elucidating the mechanisms driving cell fate decisions and development, offering new insights into the dynamics of multicellular systems.

Coevolution of transcriptional rate parameters across vertebrates

Catherine Felce

California Institute of Technology

Whilst principled biophysical modeling has been recognized as crucial for analzying RNA-seq data, phylogenetic comparative methods are currently performed at the level of steady-state transcript abundances. Focusing on steady-state abundances rather than dynamic parameters obscures information about the relative contributions of different regulatory mechanisms in shaping evolutionary divergence in gene expression.

We analyze single-cell transcriptional data from spleen immune cells across seven different species, using an inference framework called Monod. We input spliced and unspliced RNA count matrices, and fit a bursty transcription process with splicing, along with technical noise. The outputs are per-gene transcriptional parameters (burst size, splicing rate and decay rate). We then use multivariate phylogenetic models to analyze how the different parameters of the transcription process have (co)evolved across history. Using hierarchical models for evolutionary rates across genes, we can exploit the large number of measured genes and compensate for the limited number of species.

Following the approach of Cope et al. (2024) who jointly model RNA and protein evolution, we investigate the joint adaptation of transcriptional burst size, mRNA splicing and decay rate. With coupled OU models we answer unexplored questions, such as: which transcriptional rates are under the strongest selection pressure? How closely do the various mechanistic parameters coevolve? We then test whether additional regressors for the number of transcription factors and other gene-specific variables give an even more complete picture of tran-scriptional rate evolution.

SciPhy: A Bayesian phylogenetic framework using sequential genetic lineage tracing data

Sophie Seidel

University of Washington

CRISPR-based lineage tracing provides a powerful approach for reconstructing single-cell phylogenies, particularly in organisms where microscopy-based tracking is challenging. A recent advancement in this field is sequential genome editing, which not only records genetic modifications but also preserves the order in which they occur. To fully leverage this enriched data, we introduce SciPhy, a simulation and inference tool integrated within the BEAST 2 framework. SciPhy employs a Bayesian phylogenetic approach to estimate time-scaled single-cell phylogenies and cell population dynamics.

After validating SciPhy through simulations, we compare it to the widely used clustering method UPGMA and find that SciPhy-derived phylogenies differ in key structural aspects and show later cell division times. In addition, SciPhy reports uncertainty as well as cell proliferation rates, neither of which are available within a UPGMA analysis.

We apply SciPhy to two lineage tracing datasets for which we estimate time-scaled trees together with cell proliferation rates. For a proof-of-concept dataset from a monoclonal cell culture we obtain realistic cell proliferation rates. Second, we apply SciPhy to murine gastruloids, where we use a time-varying birth-death model to capture realistic aspects of this in-vitro model of early mammalian development. Notably, our analysis reveals a slowdown in cell proliferation rates following CHIR stimulation, a trend that was recently observed independently using microscopy.

Together, these examples showcase the application of advanced phylogenetic and phylodynamic tools to reconstruct and analyse single-cell phylogenies, laying the groundwork for studying developmental processes in multicellular organisms.

Seasonal Profile and Microevolution in Freshwater Bacterioplankton: Insights from Single-Cell Genomics

Yanru Jiang

Beijing Normal University

The consequences of temporally varying selection for the maintenance of genetic diversity have been elusive. Prokaryotic microbes typically have short generation times, for which seasonal changes in the environment become a prevalent selection force relevant to genetic diversity maintenance. Here we studied a natural bacterioplankton community using the single-cell genomics technique; and tested a hypothesis that species occupying broader seasonal niches (and thus under seasonally varying selection) would have greater genetic diversity. The results were contrasting to our prediction. Among species that differed in seasonal niche ranges but had similar abundances, broader seasonal niche breadth was associated with lower genetic diversity at both non-synonymous and synonymous loci. The strength of purifying selection estimated based on the ratio of nonsynonymous to synonymous nucleotide diversity ($\pi N/\pi S$) was greater in generalists. For the temporal niche generalists, the genetic diversity indices calculated for the pool of samples of two seasons, or all three ice-free seasons was not greater than those for single-season samples, suggesting limited turnover in population composition among seasons. Our results suggest that, in a temporally generalist species, only genotypes that may cope well with every season can persist, while temporally niche specialists can escape selection during unfavorable seasons by e.g. entering a dormancy state. Our results imply that the mechanisms for temporal environmental fluctuations to maintain diversity (relative nonlinearity of competition and storage effect) may require substantial ecological difference among organisms and be rarely possible for genetic variants within species.

LatinCells: A human cell map of Latin American diversity

Jose Antonio Corona Gomez

Unidad de Genómica Avanzada, Langebio, Cinvestav

Latin America's vast genetic diversity, shaped by Indigenous and continental admixture, remains underrepresented in the Human Cell Atlas, limiting ancestry-informed medical insights. The LatinCells initiative addresses this gap by generating single-cell expression atlases from Indigenous and admixed populations across Brazil, Chile, Colombia, Mexico, Peru, Uruguay, and U.S. Latino communities. Our approach integrates community engagement, cutting-edge technology, and capacity building. We collaborate with Indigenous communities in remote regions, conduct single-cell RNA sequencing (scRNA-seq) on immune and gallbladder cells, and develop tailored computational pipelines for analysis. Additionally, we promote regional research capacity by providing workshops and training materials on single-cell analysis, designed for researchers in Latin America.

To date, we have created single-cell atlases of immune cells from 17 Latin American ancestries. Using these datasets, we aim to identify cell-type-specific expression quantitative trait loci (eQTLs), revealing ancestry-driven differences in immune-regulatory molecules such as receptors and activators. Through LatinCells, we bridge gaps in global genomic datasets and empower local researchers, fostering a collaborative network to advance research capacity across Latin America.

Cell-type specific evolution of the Drosophila germline transcriptome

Imtiyaz Enayatali Hariyani

University of California Irvine

Gene expression changes in the germline are key drivers of adaptation and phenotypic evolution, yet much of our understanding comes from whole-body or bulk tissue studies, often involving distantly related species with notable allometric differences. Using single-nucleus RNA-seq, we characterized transcriptional divergence at the cell-type level in the testis and ovary of Drosophila melanogaster and D. simulans, two species that diverged 1.4 mya. Our analysis identified 10 testis and 18 ovary cell types, revealing limited variation in cell type composition and marker gene expression, both within and between species. While the transcriptional program of early germ cells was highly conserved in the testis, interspecific divergence was most pronounced in the meiotic and post-meiotic cell types. In contrast, the ovary exhibited a more consistently conserved gene expression program across cell types, although substantial differential expression was detected in the germline cells of the germarium. Overall, 15% of the genes were found to be differentially expressed between species in at least one cell type and tissue. Further, a systems-level analysis revealed divergent coexpression modules, including those associated with mitochondrial and microtubule-related processes in the late spermatocytes, and others related to germ cell development and meiotic cell cycle regulation in the ovary–suggesting lineage-specific regulatory patterns of evolution in early oogenesis. By integrating network- and gene-level analyses, our findings support a more dynamic model of transcriptome evolution in the gonads than previously recognized. Lastly, our snRNA-seq dataset offers a valuable resource for further investigating the functional implications of transcriptional divergence during gametogenesis.

Single-Cell Resolution of Rare Adaptive Mutations from Evolving Populations of Microbes

Parker Crossland

Arizona State University

Rare genotypes play a fundamental role in adaptation, ecological dynamics, and disease evolution, yet their detection and characterization remain major challenges in population genetics and evolutionary biology. Traditional bulk sequencing obscures rare variants and loses haplotype structure, while single-cell isolation is labor-intensive and inefficient for studying low-frequency genotypes. We introduce Fluorescent Labeling for Intracellular Genomics (FLInG), a novel approach that integrates in situ amplification of genomic DNA with fluorescence-activated cell sorting to allow whole-genome sequencing of rare genotypes. By enriching for cells harboring these rare genetic markers of interest, FLInG enables whole-genome recovery from adaptive genotypes while preserving haplotype information and cellular context.

We apply FLInG to a barcoded S. cerevisiae evolution to uncover adaptive mutations conferring drug resistance that traditional sequencing failed to detect. Previous studies have used statistical techniques, such as Hidden Markov Models (HMMs), to predict haplotype frequencies. FLInG directly associates genetic mutations with specific cell lineages, providing empirical haplotype-frequency data that can improve inference-based approaches. Our method facilitates the integration of sequencing data with population genetics models of evolutionary dynamics. FLInG's ability to isolate and sequence rare genotypes allows us to identify additional adaptive mutations and haplotypes within these lineages. This has broad applications for inferring lineage-specific adaptation, reconstructing evolutionary trajectories, and characterizing genetic heterogeneity in microbial communities, cancer evolution, and host-pathogen interactions at the single-cell level.

Symposium 19: Thematic symposium: conservation genomics

Population genomic landscapes and insights for conservation of the critically endangered island-endemic Chinese pangolin in Taiwan

Jingyang Hu

Yunnan University

Island-species/populations are characterized by evolutionary uniqueness and a high degree of endemism and extinction. The conservation and restoration of island populations have become the most challenging and urgent issues in biodiversity conservation. Taiwanese pangolin (Manis pentadactyla pentadactyla) is a well-known, critically endangered, island-endemic Chinese pangolin subspecies, which has been the focus of conservation concern. However, the subspecies' demographic history and the genomic consequences of population declines, as well as survival potential are unknown, which severely hinders effective development of conservation strategies. Here, we first generated large-scale population genomics data for Taiwanese pangolin covering northern, central, southern, and eastern parts of Taiwan island. We revealed that Taiwanese pangolin originated in southeast China and is differentiated into Northern and non-Northern populations due to the isolation by the Hsuehshan and Central Mountain Ranges, proposing treat them as separate conservation units. We found that the Northern population had experienced a more severe bottleneck and isolation than the non-Northern population, which corresponded to the estimated current lower genetic diversity, higher inbreeding and genetic load of the Northern population. We revealed that Northern populations is affected by climate change more serious. Evolutionary potential simulation indicated that the Northern population need a higher population growth rate to achieve evolutionary potential equal to the non-Northern population over the next 100 years, deserving conservation prioritization. This study enhances understanding of the genetic background, conservation status and future prospects for Taiwanese pangolins, and provides knowledge concerning the genetic consequences of small and isolated island population.

Genetic rescue in Florida panthers maintained local diversity and provided temporary fitness benefits

Diana Aguilar Gomez

UCLA

Pumas (Puma concolor) have an extensive range from Canada to Argentina. However, due to habitat loss and unregulated hunting, the Florida panther represents the only remaining breeding population east of the Mississippi River. By the 1990s, fewer than 30 individuals remained, suffering from severe inbreeding depression. To counteract these effects, eight pumas from Texas were translocated into southern Florida in 1995. This genetic rescue effort successfully increased population size and reduced inbreeding depression. While genetic rescue is frequently proposed for mitigating small population effects, its underlying mechanisms and long-term outcomes remain insufficiently explored. To assess the genomic consequences of this intervention, we sequenced the genomes of post-translocation Florida panthers (PTFPs). Our analysis of local ancestry revealed that no genomic regions were entirely replaced by Texas ancestry, ruling out genetic swamping. The fitness benefits of the translocation were primarily driven by increased heterozygosity, which mitigated the effects of recessive deleterious alleles, rather than by a direct reduction in the overall number of deleterious mutations. We found no evidence that selection systematically favored Texas over Florida genetic material. Simulations indicated that while heterozygosity remained elevated over time compared to a no-translocation scenario, the fitness benefits were temporary. These findings provide critical insights for managing the Florida panther population and offer guidance for applying genetic rescue in other inbred wildlife populations facing conservation challenges.

Inbreeding load in a small and managed population: two decades of Hihi / Stitchbird (Notiomystis cincta) genomics

Hui Zhen Tan

University of Auckland

Population declines result in increasingly small and isolated populations in the wild. These small populations often experience genetic diversity loss and increased inbreeding, which can reduce fitness and adaptive potential. While genomics is now routinely used to characterise threatened populations, more studies on the dynamics of genetic load and purging are needed to assess management strategies. Our study system is the Hihi / Stitchbird, Notiomystis cincta, a forest-dwelling songbird endemic to Aotearoa, New Zealand. The hihi went extinct from the North Island of Aotearoa in the late 1800s and underwent a prolonged bottleneck on a single offshore island. Translocations to restore hihi to its former range began in the late 1900s, and seven reintroduced populations currently exist. In this study, we used whole genome resequencing data across two decades from the reintroduced island population of Tiritiri Matangi. Low-coverage sequences were first imputed to quantify inbreeding levels using runs of homozygosity. We then tested for evidence of inbreeding depression by modelling lifetime reproductive success against genome-wide inbreeding. We calculated genetic load in lethal equivalents using a logistic regression model and quantified the opportunity for purging based on our long-term pedigree. We found high inbreeding levels in the hihi and some evidence for inbreeding depression. The genetic load carried by the population reflects past bottlenecks and subsequent management. Purging of mildly deleterious mutations could be less effective in the face of genetic drift and reduced selection. Our study contributes to understanding small population dynamics and the evolutionary mechanisms that affect adaptive potential.

Convergent evolutionary deadend and breakdown of hard chorion in egg-guarding fish reproductive strategies

Tatsuki Nagasawa

Institute of Science Tokyo

Fish exhibit a diverse array of reproductive strategies adapted to various ecological niches. Parental egg-guarding, including live-bearing, mouth-brooding, and male egg protection by brood pouches, represents an effective strategy for ensuring larval survival and has emerged independently in multiple lineages. Despite the recognized evolutionary bias that favors a strategy transition from non-guarding to egg-guarding, the genetic mechanisms underlying this bias and the commonalities among egg-guarding species remain elusive. This study explores the relationship between egg-guarding and the chorion hardening system crucial for protecting eggs in non-guarding species. By analyzing whole genome sequences of 212 species of Acanthopterygii fish across 25 orders, we discovered that multiple genes associated with chorion hardening have become pseudogenes in various egg-guarding species, indicating a collapse of the chorion hardening system in these fish. These findings suggest that the evolutionary bias in fish reproductive strategies not only aims to enhance survival efficiency but also imposes a constraint on egg-guarding species, preventing them from reverting to a reproductive strategy for relying on a hardened chorion. In particular, alveolin, previously characterized as a single mutant resulting in significantly fragile chorion in medaka, suggests a strong correlation between egg-guarding strategy and gene loss. Our results suggest that evolutionary deadends because of gene loss may impose evolutionary constraints at the behavioral level in reproduction. The observed association between gene loss and reproductive strategies provides insights into suitable reproductive environments for each species and may facilitate non-invasive estimation of reproductive strategies in species with unknown breeding behaviors.

1000 arctic genomes for diversity, adaptations and longevity: Polar Bear, Muksun, Beluga Whale, Bowhead Whale and others

Dmitry Pustoshilov

Biotech Campus LLC, Russia

The "1000 Arctic Genomes" project is a Russian-led genomic initiative aimed at uncovering the genetic basis of adaptation, resilience, and longevity in Arctic species facing extreme environmental conditions and climate change pressures. Understanding their genomic diversity is crucial for evolutionary biology, conservation, and physiological research, particularly in cold tolerance, oxygen utilization, and longevity. To achieve this, the project integrates high-quality de novo genome assemblies, population genomics, phylogenetics, gene annotation, and comparative genomics, utilizing MGI short-read sequencing, MGI st-LFR, Hi-C, flow cytometry, ONT long and ultra-long reads, and RNA-Seq.

We present the project's progress, including its roadmap, species selection, and initial findings. Pilot species include the polar bear (Ursus maritimus), muksun, beluga whale, and others, selected for their unique Arctic adaptations. Among them, the polar bear serves as a key model species, yet faces significant threats from climate change-driven sea ice loss, underscoring the need for genomic research in conservation.

A major milestone is the high-quality de novo genome assembly of the polar bear, surpassing the best available reference in NCBI Assembly. This assembly has enabled the discovery of novel genes linked to Arctic adaptations and improved phylogenetic reconstructions of the Ursidae family. These findings provide critical insights into Arctic resilience mechanisms and contribute to broader efforts in biodiversity conservation, evolutionary research, and climate change adaptation studies.

Comparative genomics supports ecologically induced selection as a putative driver of banded penguin diversification

Fabiola León

Pontificia Universidad Católica de Chile

The relative importance of genetic drift and local adaptation in facilitating speciation remains unclear. This is particu-larly true for seabirds, which can disperse over large geographic distances, providing opportunities for intermittentgene flow among distant colonies that span the temperature and salinity gradients of the oceans. Here, we delve intothe genomic basis of adaptation and speciation of banded penguins, Galápagos (Spheniscus mendiculus), Humboldt(Spheniscus humboldti), Magellanic (Spheniscus magellanicus), and African penguins (Spheniscus demersus), by analyz-ing 114 genomes from the main 16 breeding colonies. We aim to identify the molecular mechanism and genomic adap-tive traits that have facilitated their diversifications. Through positive selection and gene family expansion analyses,we identified candidate genes that may be related to reproductive isolation processes mediated by ecological thermalniche divergence. We recover signals of positive selection on key loci associated with spermatogenesis, especially dur-ing the recent peripatric divergence of the Galápagos penguin from the Humboldt penguin. High temperatures intropical habitats may have favored selection on loci associated with spermatogenesis to maintain sperm viability,leading to reproductive isolation among young species. Our results suggest that genome-wide selection on loci asso-ciated with molecular pathways that underpin thermoregulation, osmoregulation, hypoxia, and social behavior ap-pears to have been crucial in local adaptation of banded penguins. Overall, these results contribute to ourunderstanding of how the complexity of biotic, but especially abiotic, factors, along with the high dispersal capabilitiesof these marine species, may promote both neutral and adaptive lineage divergence even in the presence of gene flow.

Integrating genomic tools to safeguard siamese crocodile genetic purity

Thitipong Panthum

Kasetsart University

Hybrids between the critically endangered Siamese crocodile (Crocodylus siamensis) and the IUCN-listed least concern saltwater crocodile (C. porosus) in captive populations present a significant challenge to conservation and reintroduction efforts for C. siamensis. This study emphasizes the importance of utilizing alternative DNA markers and analyzing diverse populations to accurately identify and exclude hybrids from reintroduction programs. To achieve this, DArT sequencing was employed to detect genome-wide single nucleotide polymorphisms (SNPs) in both species, confirming the genetic profiles of parental species and their hybrids. A population of Australian saltwater crocodiles was used to compare the distribution of species-specific diagnostic SNPs. Multiple analytical methods were applied to assess hybridization levels, particularly in three individuals exhibiting evidence of potential backcrossing. To validate species-diagnostic SNP loci, a PCR-based approach was developed, with 20 SNP loci selected for primer design. Of these, three loci successfully differentiated between the parental species and various hybridization levels. By integrating mitochondrial and nuclear genetic data—including microsatellite genotyping and species-diagnostic DNA markers—this study introduces a novel approach to overcome the limitations of individual methods. This comprehensive strategy supports conservation prioritization by ensuring genetic integrity before reintroducing individuals into the wild, contributing to the long-term survival of the species through effective management and reintroduction programs.

Conservation genomics of an important mountain ungulate, Capra sibirica

Shamshidin Abduriyim

Shihezi University

As a relic and the largest and heaviest species of the genus Capra, the Asiatic ibex (Capra sibirica) is a key animal in the mountain ecosystems of southern and central Asia. It is of critical conservation concern both globally and locally. However, its population structure and evolutionary history remain poorly understood. Here, we preliminarily obtained 32 high-quality whole-genome resequencing datasets of C. sibirica from most of its natural range and 23 skulls from five locations. Consequently, we identified >3.18 million SNPs for the nuclear genome and 32 mitogenome sequences for downstream analyses. Structure, PCA, and phylogenetic analyses based on nuclear SNPs revealed two main clusters: samples from the East Tianshan Mountains and Altai Mountains formed Cluster I, those from the Pamir Plateau North Kunlun Mountains and Hindukush Mountains were grouped into Cluster II. However, mitogenome clusters were inconsistent with nuclear clusters, suggesting mitochondrial introgression. PCA of skull indices formed two clusters as well. The genomic heterozygosity for geographic populations ranged from 0.00089 for the Koktokay population to 0.00146 for the Ulugqat population. Demographic history analyses showed that different geographic populations underwent differential population fluctuations from the end of the Naynayxungla Glaciation to the Last Glacial Maximum. Mitochondrial introgression may have played an important role in restoring effective population sizes in some populations. Our results provide an important baseline for planning protection strategies and ensuring the effective conservation of the Asiatic ibex.

Genetic diversity loss in the Anthropocene will continue long after habitat destruction ends

Kristy Mualim

Stanford University

Genetic diversity within species is the basis for evolutionary adaptive capacity and has recently been included as a target for protection in the United Nations' Global Biodiversity Framework (GBF). However, there exists a lack of reliable large-scale predictive frameworks to quantify current genetic diversity loss, let alone how to quantitatively predict future losses under different habitat destruction scenarios in the 21st century. Combining spatio-temporal population-scale genetic theory with population-scale genomic data of 29 plant and animal species, we studied the dynamics of genetic diversity across varying habitat area loss scenarios. We show that genetic diversity reacts slowly to initial population and habitat area declines, and continues to decline for many decades even after habitats are fully protected. To understand the magnitude of this problem, we combined our predictive method with species' habitat area and population monitoring reported in the Living Planet Index, the IUCN Red List, and new GBF indicators. We approximate existing genetic diversity loss in 13,808 species to account to a short-term genetic diversity loss of 13–22% and a long-term loss of 42–48% with substantial deviations depending on the level of habitat fragmentation. These results highlight that safeguarding existing habitats alone is insufficient to maintain the genetic health of species, and relying solely on continuous genetic monitoring likely underestimates long term impacts. We provide an area-based spatio-temporal predictive framework to develop quantitative predictions of global genetic diversity.

Molecular adaptation and conservation genomics in carnivorous bats

Huabin Zhao

Wuhan University

Dietary specialization stands as a major factor in the study of adaptive evolution and the field of conservation biology among animals. Although bats show unparalleled dietary diversification among mammals, specialized carnivory remains relatively rare within this group. Consequently, our comprehension of the genetic and conservation aspects associated with this specific dietary niche in bats has largely remained unclear. To investigate molecular adaptations and conservation genetics in carnivorous bats, we produced a new draft genome assembly for the carnivorous bat Vampyrum spectrum. Furthermore, we utilized this genome alongside another distantly related carnivorous bat Megaderma lyra, to conduct genome-wide comparative analyses with other bat species. Our findings unveil that genes linked to lipid metabolism exhibit signatures of positive selection and convergent molecular adaptation in the two divergent lineages of carnivorous bats. Additionally, our genomic analyses also reveal low genetic diversity in both carnivorous bat species. This pattern is attributed to their continuously declining population sizes and low levels of heterozygosity, signaling their vulnerability and emphasizing the pressing need for conservation efforts. The genuine carnivorous behavior is a rare occurrence, observed in fewer than one percent of all bat species (n=9). We plan to assess the population trend, genetic diversity and molecular adaptation for all carnivorous bats in the future.

Population genomics studies of critically endangered hawksbill turtles in Southeast Asia

Regine Tiong

Nanyang Technological University

Hawksbill turtles are critically endangered and aiding in their recovery would help to protect the coral reef ecosystems and their biodiversity. However, genetic studies on marine turtles have relied on limited genetic markers, including the mitochondrial control region, providing incomplete and biased estimations of genetic diversity and population structure. Here, we generated a de novo genome assembly and high-quality whole-genome population datasets from hawksbill turtles nesting and foraging in Singapore. We identified approximately 12 million single nucleotide polymorphisms from the genome dataset, significantly enhancing the genetic markers to examine population structure and genetic characteristics of this species. Our analyses revealed an extent of genetic diversity and demographic history of the hawksbill turtle of Singapore. There is a pronounced degree of inbreeding within the population: most individuals share at least a first cousin relationship. Furthermore, we identified a multiple-paternity nest, while the parents are related. The estimated demographic history suggests the impact of past climate change on the declining hawksbill turtle population. Positive selection was detected in genes critical for environmental sensing and response, significant to natal behavior, emphasizing the strong correlation of turtle survival and climate and environmental changes. The study shows the significance of whole genome sequencing data for studying this species and the vulnerability of Singapore hawksbill turtle population facing climate change and human activities from land reclamation plans at their nesting sites. This project aims to direct conservation efforts for the critically endangered hawksbill turtle population in Singapore and in the Indo-Pacific region.

Museomics and indigenous knowledge confirm existence of the marozi - spotted lions in East Africa

Xiaodong Liu

University of Copenhagen

Big cats rank among the most iconic species worldwide. Kenya in East Africa is home to two big cat species: the lion and the leopard. Another highly cryptic and elusive big cat, called the 'marozi', also inhabits Kenya. Similar in size to a lion but having shorter front legs, it bears spots reminiscent of a leopard. Since its first documentation, the marozi has long puzzled researchers and conservationists regarding its true nature. Here, we investigate the evolutionary affiliation of the marozi by analyzing its nuclear genome extracted from the only known specimen and incorporating indigenous knowledge from the Kikuyu and Maasai communities in Kenya. Our genetic analysis reveals that the marozi is a lion color morph evolutionarily close to local lion populations in Kenya and Tanzania. Genome screening of the marozi and comparison with previously published lion and leopard genomes indicate that its distinct coloration is unlikely due to genetic mutations alone. Furthermore, knowledge keepers from the Kikuyu and Maasai report that spotted lions like the marozi occurred on mountains like the Aberdares of Kenya. In conclusion, we demonstrate that the marozi is a unique morph of lion that likely continues to inhabit high elevations in Kenya.

Remnant island populations of Australian mammals purge genetic load

Emily Roycroft

Monash University

Australian mammals have the world's highest contemporary extinction rate. With recent collapse of mainland populations, and natural replicates of long-term bottlenecks on islands, they are a model for understanding the genomic impact of decline. Genome-scale data from extinct and living populations across six flagship species shows that now-extinct mainland populations had high genetic diversity, and a high burden of segregating genetic load. While theory predicts small populations will accumulate harmful mutations, we find that genetic load is purged from isolated island populations by selection. Mixed-source reintroductions show that genetic rescue decreases fixed genetic load but increases segregating load in the near-term, allowing additional purging under population growth. These results highlight the importance of 'pre-purged' populations in species recovery and inform genetic management of remnant populations globally.

Symposium 20: Thematic symposium: evolution of immunity

Where is the innate immune system of fungi? Molecular evolution of a fungal receptor gene family

S. Lorena Ament-Velásquez

Stockholm University, Sweden / PHIM, France

A universal feature of cellular organisms is the capacity to recognize self from nonself. Fungi have large and diverse repertoires of genes that resemble NOD-like receptors (NLRs), intracellular proteins that detect signs of infection in plants and animals. A few fungal NLRs are known to have a role in allorecognition, which leads to the rejection of conspecific nonself tissues. However, the function of all the other NLRs remains unknown. Using high-quality genome assemblies of the Podospora anserina species complex, we show that fungal NLRs have a rapid turnover, with more gains and losses and longer phylogenetic distances than observed in random genes. A subgroup of NLRs is characterized by having superstructure-forming repeats with very high sequence identity (known as high internal conservation or HIC). We show that HIC NLRs are more associated with transposable elements and have higher synonymous nucleotide diversity, suggesting higher mutation rates. Moreover, we find evidence that some HIC NLR genes exchange repeats by unequal crossing-overs or other recombination events that likely maintain the high sequence similarity. At the population level, regions introgressed between sibling species are significantly enriched with NLRs, which confounds balancing selection signals, and several NLR loci show evidence of selective sweeps. We conclude that fungal NLRs evolve under dynamics similar to that of plants and animals, consistent with an equivalent immune function, and we further suggest that association with transposable elements and the mechanisms maintaining high similarity between repeats boost diversity beyond the genomic mutational background.

Evolution of Erm-associated Antibiotic Resistance

Yongjun Tan

Saint Louis University

Modification of ribosomal RNAs by the erythromycin resistance RNA methyltransferase (erm) is the major cross-resistance mechanism for clinically important macrolides, lincosamides, and streptogramins (MLS phenotype). However, this resistance depends on regulated erm expression, which is typically induced by macrolide-mediated ribosome stalling in the upstream leader sequence. Such stalling triggers a conformational switch in intergenic RNA hairpins that permits erm translation. In other words, merely harboring erm does not ensure its expression, underscoring the need to examine non-coding allelic variation in its regulatory region and its effect on the MLS phenotype. We investigated the evolutionary emergence of upstream erm regulatory elements and the impact of allelic variation on erm induction. Systematic profiling of upstream sequences across known erm operons revealed that certain subfamilies (e.g., ermB and ermC) independently evolved distinct combinations of small upstream ORFs and palindromic repeats. A population-wide genomic analysis of ermB further uncovered substantial non-random allelic variation at multiple positions within its upstream regulatory region. Integrating machine learning classification with RNA secondary structure modeling, we found that many alleles act cooperatively to influence the stability of alternative RNA hairpin structures, ultimately affecting ermB inducibility and MLS phenotypes. Experimental validation of 11 randomly selected variants confirmed 91% accuracy in predicting resistance. Moreover, the phylogenetic distribution of MLS-sensitive and MLS-resistant ermB loci indicated repeated, independent evolution of MLS resistance. Together, this study elucidates the evolutionary processes driving MLS resistance and highlights the potential of non-coding genomic allele data to predict antibiotic resistance phenotypes.

Reference-quality bat genomes illuminate clade-specific immunity to viruses

Aaron Irving

Zhejiang University - University of Edinburgh Institute

Bats are suggested to harbour more zoonotic viruses than other mammalian orders, with infections being largely asymptomatic, exhibiting limited tissue-damaging inflammation and immunopathology. To investigate a genomic basis for enhanced immunity, the Bat1K project generated reference-quality genomes of ten bat species, including potential viral reservoirs. A systematic analysis covering 17000 orthologs from 115 mammals identified immune selection signatures more prevalent in bats. Excess immune gene adaptations were identified in the ancestral chiropteran branch and also many descending bat lineages, including viral entry and detection factors, and regulators of antiviral and inflammatory responses. ISG15, an antiviral gene contributing to hyperinflammation during COVID-19, exhibits key residue changes in rhinolophid and hipposiderid bats. Experimental validation revealed species-specific antiviral differences, ISGylation potential and highlighted an essential role of protein conjugation in antiviral function for bat ISG15, separate to its role in secretion and inflammation in humans. In contrast to humans, ISG15 from most rhinolophid and hipposiderid bats shows strong anti-SARS-CoV-2 activity. Our work reveals molecular mechanisms that contribute to viral tolerance and disease resistance in bats, highlighting clade-specific and species-specific differences in antiviral immunity.

Evolutionary repurposing of TLR21 as a broad-spectrum recogniser of CpG-DNA

Steven Fiddaman

The Pirbright Institute / University of Oxford

The ability to detect non-self nucleic acid as an indicator of pathogen presence is a major component of the pattern-recognition receptor repertoire. Specifically, the abundance of unmethylated CpG dinucleotides in microbial genomes, compared to their relative scarcity in vertebrate nuclear genomes, provides a point of difference that can be used to differentiate between self and non-self. This study focuses on the functional evolution of TLR21 – one of two vertebrate Toll-like receptors (TLRs) that respond to unmethylated CpG motifs (the other being TLR9). CpG recognition by TLRs is dependent on the surrounding sequence context, with different species favouring particular motifs. Using functional assays, we find that TLR21 of fishes and some amphibians are fastidious in their CpG recognition – generally favouring only a single CpG motif. Birds and reptiles, however, have generalist recognition profiles, responding indiscriminately to a wide range of different motifs. We argue that this transition from specialist to generalist is a result of the loss of TLR9 in the avian/reptilian lineage, which is a known generalist in fishes and mammals. The evolution of generalism in avian/reptilian TLR21 is supported by a burst of positive selection in this lineage. For the first time, we also identify the DNA-binding motif of TLR21, and identify polymorphisms that likely alter the avidity of DNA binding by the receptor between the different species. Our results highlight one of the most dramatic functional repurposings of a vertebrate TLR, and underscore the importance of generalist CpG recognition ability in vertebrate pattern recognition.

A genomic deep dive into the evolution of structural and genetic diversity of Major Histocompatibility Complex (MHC) regions of African, Asian, and South American primates

Joanna Malukiewicz

University of Hamburg

Adaptively, the major histocompatibility complex (MHC) region 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 structurally unique features like frequent gene duplications, high gene density, high polymorphism levels, and high repetitiveness define this region. However, due to these unique evolutionary features, it has historically been difficult to deduce genome-level structural variation of the MHC regions in organisms beyond humans, including non-human primates. To address this gap, we harnessed long read sequencing technologies to comparatively resolve "problematic genomic" regions such as the MHC in several primate genera. Our dataset includes diploid, phased telomere-to-telomere genomes from Pan, Gorilla, Pongo, Symphalangus, long-read genomes from Callithrix jacchus, and Callithrix target enriched long-read MHC sequences. Here we describe and compare the gene organization, structure, and copy number variation of classical and non-classical MHC genes within the full MHC region of the primate taxa mentioned above. Our characterisations of immunologically important MHC genes include previously annotated and as well as novel loci in apes, and we provide novel characterisation of the siamang and marmoset MHC region. Our comparisons reveal structural variation at both the intra-individual and the interspecific level- variation which likely resulted from both neutral and adaptive processes, reflecting different evolutionary trajectories of primate species. This study highlights recent sequencing advances that facilitate understanding inter-specific structural variation in challenging genomic regions, particularly in laying down the groundwork for pangenomic studies of primate MHC variation.

Domesticated Plants Show Elevated Immune Receptor Gene Loss Across Diverse Lineages

Noah Bourne

University of Sheffield

Plant domestication is sometimes associated with a reduction in the diversity of immune receptor genes (IRGs), critical for pathogen recognition and defence. Yet, the extent to which this occurs and the mechanism driving this pattern remain unclear. Using a comparative genomics framework, we analysed the IRG repertoires of 15 domesticated crop species and their wild relatives, representing nine plant families. We examined both cell-surface pattern recognition receptors (PRRs) and intracellular nucleotide-binding leucine-rich repeat receptors (NLRs). Our results show that five crops—grapes, mandarins, rice, barley and yellow sarson—exhibited significantly reduced IRG repertoires compared to their wild counterparts, driven by increased IRG loss. Moreover, the longer a species has been domesticated, the greater the extent of IRG loss, indicating that this process is ongoing and likely impacts most crops. This study provides insights into how domestication impacts plant immunity, with implications for breeding strategies enhancing crop disease resistance.

Not the same: genomic analysis explains variation in strength of positive selection across avian immune genes

Michal Vinkler

Charles University, Faculty of Science

Host-pathogen co-evolution is assumed to diversify immune genes through arms races. Nonetheless, the strength of the selection can differ between individual immune genes. Although evidence available for the Major Histocompatibility Complex (MHC) and innate immune receptors (e.g. Toll-like receptors, TLRs) suggests that these essential activators of vertebrate immunity experience strong positive selection it remains unclear if this is the general rule. We propose that the strength of positive selection is dependent on combination of gene molecular function and cellular localization. Using 50 publicly available avian genomes, we performed the largest comparative analysis of positive selection in immune genes up to date. We defined five different semi-manually classified immune catalogues to annotate all immune genes with resolved bird-mammal orthology. Applying dN/dS approaches, we analysed positive selection in 453 protein-coding immune genes. There were no differences in intensity of selection between innate and adaptive immunity. However, core immunome (genes with primary functions in immunity) was subject to stronger selection than peripheral immunome (genes having also non-immune functions). Effector and membrane-bound / extracellular molecules were under stronger positive selection than others. Combination of functional classification and localisation appears more suitable for characterisation of major trends in immune gene evolution than the GO / KEGG classifications. Our results highlight that immune genes are diversified in their evolutionary patterns, with those having primary roles in immunity and those whose products physically interact with pathogen structures being subjects to stronger positive selection than others.

The prokaryotic roots of eukaryotic immune systems

Max Burroughs

National Library of Medicine, National Institutes of Health

A sweeping consequence of the genome era has been the discovery of deep evolutionary links between prokaryotes and systems once thought to be quintessentially eukaryotic. Perhaps nowhere is this more evident than the parallels observed between eukaryotic and prokaryotic immunity and apoptosis systems. Across the tree of life, these systems share an operational framework, shaping their biochemical logic and evolutionary trajectories. The diversification of immune genes in the prokaryotic superkingdoms, followed by lateral transfer across the eukaryotic tree of life, was central to the emergence of innate immunity in the latter. These include protein domains related to nucleotide second messenger-dependent systems like cGAS, OAS, and STING, NAD+/nucleotide degradation, and P-loop NTPase domains of the STAND and GTPase clades playing pivotal roles in eukaryotic immunity and inflammation. Moreover, several domains orchestrating programmed cell death, including the recently discovered bacterial progenitors of the Death and TRADD-N domains, suggest an intimate link between immunity and the emergence of multicellularity in eukaryotic lineages such as animals. We contextualized the totality of these linkages, dividing them into two categories: 1) acquired eukaryotic immune components functioning ‘as is’ in prokaryotes and 2) domains that are ‘reused’ in eukaryotes for distinct functional roles from prokaryotic conflict systems, often in convergent, immunity-based roles.

From the dawn of molecular biology, immune proteins have furnished foundational biotechnological reagents, including restriction enzymes and CRISPR-Cas genome-editing endonucleases from prokaryotes and antibodies and the Piwi/Ago-dependent small interfering RNA systems from eukaryotes. Several newly discovered immune components hold substantial biotechnological potential.

Symposium 21: Thematic symposium: virus evolution and pathogen evolution

Tracking and quantifying reassortment divergence of panzootic H5Nx highly pathogenic avian influenza viruses and the key drivers of its spatial heterogeneity

Lu Lu

University of Edinburgh

Avian influenza viruses (AIVs) undergo frequent genetic reassortment, which can be accompanied by large phenotypic 'shifts' that play a critical role in the emergence and cross-species transmission of novel viral lineages. We analysed over 9000 2.3.4.4b clade H5 highly pathogenic (H5Nx) genomes to better understand the evolutionary and ecological dynamics of reassortant lineages emerging since the 2019/2020 wave. Specifically, we identified 209 distinct reassortants, of which the vast majority (83%) contributed little to overall circulation, but six (3%) 'dominant' reassortants had high spatial diffusion coefficients, infected a diverse host range, and circulated across multiple seasons. Applying phylodynamic and statistical modelling to these data, we estimated that the largest number of new reassortants observed since the 2019/2020 wave originated from Asia, but only 40% showed signs of onwards spread. Of those reassortants successfully spreading, those originating from North America spread the most, followed by reassortants from Asia. Across all regions, further spread was positively associated with persistence in wild migratory birds. Additionally, our analysis revealed that poultry trading is significantly associated with the transmission of reassortants, particularly in North America. Altogether, our findings highlight the geographically and ecologically segregated dynamics of H5Nx reassortants. We also found poultry trading may strongly linked to reassortant transmission, especially in North America. While the long-distance transmission of HPAIV is primarily driven by wild bird migration and not easily controllable, human activities—such as enhanced surveillance of wild populations, stricter trade controls, and supervised vaccination remain a crucial backbone for effective control strategies to contain avian influenza.

An expanding universe of mutational signatures and its rapid evolution in single-stranded RNA viruses

Yue Yang

Institute of Zoology，Chinese Academy of Sciences

The study of mutational processes in somatic genomes has uncovered a wide array of endogenous and exogenous factors associated with somatic changes. However, the overall landscape of mutational processes in germline mutations across the tree of life and associated evolutionary driving forces are rather unclear. In this study, we analyzed mutational processes in single-stranded RNA (ssRNA) viruses which are known to switch between different hosts with divergent exogenous environments. We found that mutational spectra in different ssRNA viruses differ significantly and are mainly associated with their genetic divergence. Surprisingly, host environments contribute much less significantly to the mutational spectrum, challenging the prevailing view that exogenous cellular environment is a major determinant of the mutational spectrum in viruses. To dissect the evolutionary forces shaping viral spectra, we selected two important scenarios, namely the inter-host evolution between different viral strains as well as the intra-host evolution. In both scenarios, we found mutational spectra change significantly through space and time, strongly correlating with levels of natural selection. Combining the mutations across all ssRNA viruses, we identified a suite of mutational signatures with varying degrees of similarity to somatic signatures found in humans, indicating universal and divergent mutational processes across the tree of life. Taken together, we unraveled an unprecedented dynamic landscape of mutational processes in ssRNA viruses, pinpointing important evolutionary forces shaping fast evolution of mutational spectra in different species.

Unraveling epistatic interactions between sites under drug-dependent selection in the Mycobacterium tuberculosis genome

Anfisa Popova

HSE University

Understanding genetic mutations linked to drug resistance, as well as epistatic interactions between them that might promote or hinder resistance acquisition, is essential for improving resistance detection and informing treatment strategies for Mycobacterium tuberculosis (MTB). However, such analysis is challenging due to entangled selection pressures. Simultaneous administration of multiple antibiotics leads to strong episodic selection, with resistance mutations to different drugs often being selected in parallel. This coordinated selection makes it difficult to identify genuine associations and epistatic interactions.

Using publicly available genome data from more than 11,000 MTB clinical isolates with known resistance phenotypes for at least one of 13 antitubercular drugs, we reconstructed MTB phylogeny and identified mutations that are associated with drug resistance or occur preferentially in either resistant or sensitive lineages, carefully accounting for the interdependence among resistance phenotypes.

Among sites carrying these mutations, we searched for concordantly (CSP) and discordantly evolving site pairs (DSP), disentangling epistatic interactions from simultaneous drug-associated selection and other coordinated selective forces. We identified some established as well as novel interactions. Specifically, we identified two CSPs and thirteen DSPs between sites that are known to be strongly associated with resistance to anti-TB drugs. The two CSPs and four out of 13 DSPs were between sites whose mutations were associated with resistance to different drugs, while the other nine DSPs were between sites located either in the same genes or in different genes involved in alternative adaptive pathways to the same drugs.

Genetically distinct within-host subpopulations of hepatitis C virus persist after Direct-Acting Antiviral treatment failure

Lele Zhao

University of Oxford

Analysis of viral genetic data has previously revealed distinct within-host population structures in both untreated and interferon-treated chronic hepatitis C virus (HCV) infections. While multiple subpopulations persisted during the infection, each subpopulation was observed only intermittently. However, it was unknown whether similar patterns were also present after Direct Acting Antiviral (DAA) treatment, where viral populations were often assumed to go through narrow bottlenecks. Here we tested for the maintenance of population structure after DAA treatment failure, and whether there were different evolutionary rates along distinct lineages. We analysed whole-genome next-generation sequencing data generated from a randomised study with DAAs. We focused on longitudinal samples collected from patients (N=84) who failed treatment. Given the short-read nature of the data, we used a number of methods to identify distinct within-host lineages including tracking concordance in intra-host nucleotide variant (iSNV) frequencies, applying sequenced-based and tree-based clustering algorithms to sliding windows along the genome, and haplotype reconstruction. Distinct viral subpopulations were maintained among a high proportion of individuals post DAA treatment failure. Using maximum likelihood modelling and model comparison, we found an overdispersion of viral evolutionary rates among individuals, and significant differences in evolutionary rates between lineages within individuals. These results suggest the virus is compartmentalised within individuals, with the varying evolutionary rates due to different viral replication rates and/or different selection pressures. We endorse lineage awareness in future analyses of HCV evolution and infections to avoid conflating patterns from distinct lineages, and to recognise the likely existence of unsampled subpopulations.

Intrahost dynamics, together with genetic and phenotypic effects predict the success of viral mutations

Cedric Tan

University College London

Predicting the fitness of mutations in the evolution of pathogens is a long-standing and important, yet largely unsolved problem. In this study, we used SARS-CoV-2 as a model system to explore whether the intrahost diversity of viral infections could provide clues on the relative fitness of single amino acid variants (SAVs). To do so, we analysed ~15 million complete genomes and nearly ~8000 sequencing libraries generated from SARS-CoV-2 infections, which were collected at various timepoints during the COVID-19 pandemic. Across timepoints, we found that many successful SAVs were detected in the intrahost diversity of samples collected prior, with a median of 6-40 months between the initial collection dates of samples and the highest frequency seen for these SAVs. Additionally, we found that the co-occurrence of intrahost SAVs significantly captures genetic linkage patterns observed at the interhost level (Pearson's r=0.28-0.45, all $p<0.0001$). Further, we show that machine learning models can learn highly generalisable intrahost, physiochemical and phenotypic patterns to forecast the future fitness of intrahost SAVs (r2=0.48-0.63). Most of these models performed significantly better when considering genetic linkage (r2=0.53-0.68). Overall, our results document the evolutionary forces shaping the fitness of mutations, which may offer potential to forecast the emergence of future variants and ultimately inform the design of vaccine targets.

Antagonistic selection drives SARS-CoV-2 evolution within and between host

Mei Hou

Sun Yat-sen University

In viral evolution, a new mutation has to proliferate within the host (Stage I) in order to be transmitted and then compete in the host population (Stage II). We analyzed the intra-host single nucleotide variants (iSNVs) in a cohort of 79 SARS-CoV-2 infected patients with most transmissions tracked. Here, every mutation has two measures: i) iSNV frequency within each individual host in Stage I; ii) occurrence among individuals ranging from 1 (private), 2-78 (public) to 79 (global) occurrences in Stage II. In Stage I, a small fraction of non-synonymous iSNVs are sufficiently advantageous to rise to a high frequency, often 100%. However, such iSNVs usually fail to become public mutations. Thus, the selective forces in the two stages of evolution are antagonistic. For that reason, successful mutants, including many VOCs (variants of concern), have to avoid being eliminated in Stage I when they first emerge. As a result, they may not have the transmission advantage to outcompete the dominant strains and, hence, are rare in the host population. Few of them could manage to slowly accumulate advantageous mutations to compete in Stage II. When they do, they would appear suddenly as in each of the 6 successive waves of SARS-CoV-2 strains. In conclusion, Stage I evolution, the gate-keeper, may contravene the long-term viral evolution and should be heeded in viral studies.

A tale of two cancer-associated viruses

Huihui Li

SYSUCC

Human papillomavirus (HPV) and Epstein-Barr virus (EBV) are both double-stranded DNA viruses that bear strong association with various types of human cancers. Despite intensive studies in the past, it remains elusive how their pathogenicity got evolved at the genomic level. Here we performed pangenome studies on these two representative human cancer associated viruses based on comprehensive global samplings of thousands of isolates. Our population genomic and phylogenetic analyses revealed intrahost codivergence and niche adaptation shaped the evolution of both HPV and EBV, which explained why these two viruses are so widely distributed across the globe. By further combined with pangenome-based association analyses, we identified key mutational events associated with the pathogenicity evolution of HPV and EBV respectively and explored potential molecular mechanisms of these key mutations in conferring higher risks of HPV- and EBV-associated cancers (e.g., cervical cancer for HPV and nasopharyngeal cancer for EBV). Finally, we further validated such strong variant-risk association with hundreds of independently acquired clinical samples for both HPV and EBV, highlighting the biomarker value of these risk-associated HPV and EBV variants. Taken together, our study illuminated the pathogenicity evolution of HPV and EBV from a pangenomic perspective and provides valuable insights towards the development of better prevention and treatment strategies against these cancer-associated viruses in the era of precision medicine.

Discovery of sickle haemoglobin-associated mutations in the Plasmodium falciparum genome using 6,000 infections from three populations

Annie Forster

University of Oxford

The sickle haemoglobin polymorphism (HbS) is a striking example of balancing selection in the human genome, maintained due to its protection against severe malaria. Recently, protein-altering mutations in regions of the malaria parasite genome—termed Plasmodium falciparum sickle-associated (Pfsa)—were found to strongly associate with HbS. These 'Pfsa+' mutations appear to have coevolved to enable parasites to infect, grow, and cause disease in HbS carriers. With only a limited number of human-parasite genetic datasets available, the extent of this coevolutionary relationship, including whether additional loci are involved, remains unknown.

Here, we develop a novel resource of parasite whole-genome sequencing data from Gambian cases collected during high malaria transmission in the late 1980s, with linked human genotyping. We combine this with previously published data from Ghana, Gambia and Kenya to conduct a meta-analysis of association with HbS across 6257 infections. Our analysis confirms associations at previously reported loci (Pfsa1-4) but also implicates multiple novel putative loci of HbS interaction, some of which are supported across populations.

Several Pfsa loci, including newly identified ones, lie within complex genomic regions. Using whole-genome assemblies, we resolve copy number and structural variation and identify paralogous sequence variants. While Pfsa+ haplotypes are presumed to be positively selected for by HbS, they have not reached fixation in any population, raising questions as to the evolutionary forces at play. Our analysis establishes a foundation for future work to unravel the biological basis and evolutionary dynamics of this fascinating host-parasite interaction.

Malaria parasite genotype dictates asymptomatic infection in sickle cell carriers

Helena Hopson

University of Utah

The sickle cell allele (HbS) is protective against malaria symptoms caused by Plasmodium falciparum parasites but does not prevent infection. In fact, longer infections, lower parasite densities and higher transmissibility have all been reported in heterozygous HbS carriers. Intriguingly, parasites from symptomatic HbS carriers were recently reported to be almost entirely infected by parasites with specific nonsynonymous mutations in four parasite genes, named P. falciparum sickle associated alleles (Pfsa), suggesting these changes allow parasites to overcome the protective effect of HbS. However, it is unclear why overcoming HbS protection would benefit the parasite, since HbS carriers appear to be favorable reservoirs for malaria transmission. To test if the HbS-Pfsa association extends to asymptomatic infection as well, we screened 2,700 school-age children from Cameroon for asymptomatic blood stage parasitemia. By genotyping both host and parasite alleles, we find that HbS and Pfsa are already strongly associated in asymptomatic individuals and Pfsa already show strong inter-chromosomal linkage disequilibrium, supporting potential epistasis. Surprisingly, we also find that the majority of infections in non-HbS carriers are with non-Pfsa carrying parasites. We propose that HbAS and HbAA provide host environments that select for different parasite alleles, with the frequency of the host allele maintaining the parasite alleles at intermediate frequency. In support of this, Pfsa loci show evidence of long-term balancing selection, as is also observed for HbS. Furthermore, this work suggests the protective mechanism of HbS may involve resistance rather than tolerance, potentially driving parasite adaptation to survive in HbS red blood cells.

Varroa destructor antiviral responses across different life stages reveal the transmission dynamics of viruses

James Damayo

University of Sydney

One of the most devastating threats to global honey bee health is the infestation of the ectoparasitic mite and viral vector Varroa destructor. However, the transmission dynamics of the pathogenic viruses it vectors and Varroa-specific viruses are poorly understood. Most invertebrates (including V. destructor) utilise antiviral RNA interference (RNAi) pathways as their main antiviral mechanism. This results in the production of virus-derived small interfering RNA (vsiRNA) fragments, typically 21–23 nt with bidirectional polarity, depending on the host. Interestingly, an established viral infection in V. destructor produces antisense 24-nt vsiRNA fragments that resemble the secondary vsiRNAs observed in nematodes. To better understand viral transmission dynamics in V. destructor, we performed small RNA sequencing to identify vsiRNA signals indicative of viral replication in 31 V. destructor mites from seven life stages. We identified five viruses that consistently produce antisense 24-nt vsiRNA across all life stages, suggesting these viruses are vertically transmitted. Additionally, the pathogenic bee virus, deformed wing virus A (DWV-A), produces bidirectional 23-nt vsiRNA fragments at earlier life stages that transition to antisense 24-nt vsiRNA at later life stages. These data suggest that once a virus is acquired, the primary vsiRNA fragments produced are 23-nt and bidirectional, as the infection progresses secondary 24-nt antisense vsiRNA is generated. Unravelling the transmission dynamics of viruses associated with V. destructor, including the pathogenic bee virus DWV-A, enhances our understanding of the persistence and dissemination of viruses within honey bee-mite populations, which is crucial for developing novel control methods to improve honey bee health.

Exploring Pathogen diversity in Patagonian Hunter-Gathers in a 6000-100-year time transect through Paleogenomics

Florencia Alvarez Gallego

Universidad Autónoma de México-UNAM

The genomic study of ancient pathogens provides insights into human migrations and the adaptation of pathogens over time. Here, we present the first metagenomic study of ancient pathogens from Patagonia, targeting bacterial and viral pathogens in human remains from the early Holocene. We analyzed 35 teeth and 19 bones from 32 individuals dated from 6000 yBP to 100 yBP, both with and without palaeopathological evidence. Unmapped reads to the human genome were taxonomically classified using Kraken2, MALT, MetaPhlan4, alignment against virulence factors, and competitive mapping to exclude false positives. We detected three ancient pathogen DNA of Tannerella forsythia (periodontal disease) from an individual dated to 950 years before present (yBP), hepatitis B virus (HBV) genotype H (liver disease) from an individual dated to 1500 yBP, and Mycobacterium lepromatosis (the second causative agent of lepromatous leprosy) from 3 individuals dated to 4700-2130 yBP. We reconstructed their genomes using capture-enrichment strategies, such as Whole Genome In-Solution Capture (WISC) and MyBaits. The phylogenetic analysis of T. forsythia, using a maximum likelihood method, revealed an association with pre-contact lineages previously reported from Mexico. Additionally, we recovered an HBV genotype H genome at 6.8X coverage, M. lepromatosis at 1.78X coverage. Further phylogenetic analysis will allow us to determine the epidemiological differences between these ancient pathogen strains and those present in current populations across the Americas.

Bronze Age Yersinia pestis genome from sheep sheds light on hosts and evolution of a prehistoric plague lineage

Felix M. Key

Max Planck Institute for Infection Biology

Most human pathogens are of zoonotic origin. Many emerged during prehistory, coinciding with the domestication of animals providing more opportunities for spillover into human populations. However, we lack direct DNA evidence linking animal and human infections during prehistory because unlike humans most animal remains entered the archeological record as undiseased and heat-exposed food waste. Here we present the first prehistoric Yersinia pestis genome recovered from a 3rd millennium BCE domesticated sheep from the Eurasian Steppe belonging to the Late Neolithic Bronze Age (LNBA) lineage - a lineage until now exclusively identified in ancient humans from across Eurasia. Interestingly, the closest genetic relative to this genome is a human infection from the same cultural complex, implicating sheep in the pathogen transmission landscape among this exclusively pastoralist society. We show that this ancient lineage underwent independent ancestral gene decay paralleling extant lineages, but evolved under distinct selective pressures contributing to its lack of geographic differentiation. We collect evidence supporting a scenario where the LNBA lineage, unable to efficiently transmit via fleas, spread from an unidentified reservoir to humans, sheep, and likely other domesticates - increasing its known host range, and thus the infection risk for humans. Collectively, our results connect prehistoric livestock with infectious disease in humans and showcase the power of moving paleomicrobiology into the zooarchaeological record.

Tracing the evolutionary history of Epstein-Barr Virus using modern and ancient genomes

Yuejiao Huang

University of Copenhagen

Epstein–Barr Virus (EBV) is a ubiquitous human pathogen with two main genotypes (type 1 and type 2), which exhibit distinct geographic distributions and disease associations. Although EBV infection is often asymptomatic, it is also implicated in various malignancies and autoimmune diseases, prompting extensive research into its genetic diversity and pathogenic mechanisms. However, due to the lack of ancient EBV genomes, its long-term evolutionary dynamics remain understudied. In this study, we report the oldest ancient EBV genomes sequenced to date. The five genomes were recovered from pieces of chewed birch tar from Mesolithic and Neolithic contexts in Europe. We demonstrate that one of the Neolithic genomes belongs to type 2 EBV, while the remaining four are classified as type 1. Phylogenetic analyses reveal that these type 1 genomes cluster closely with contemporary Asian EBV lineages, suggesting that these ancient European strains and modern Asian lineages share a deeper common ancestor, and highlighting the complexity of EBV's prehistoric dispersal. One ancient strain carried two out of three variants within BALF2 associated with high risk of nasopharyngeal carcinoma, indicating that oncogenic variants of EBV already circulated thousand years ago in Europe. Besides, we identified a missense mutation in the same genome in glycoprotein 350, associated with EBV's ability to invade host B cells. These mutations may reflect early adaptations influencing viral infectivity and oncogenic potential. Our findings provide valuable insights into the genetic diversity and evolutionary dynamics of EBV over millennia, shedding light on its long-term adaptation and disease associations in human populations.

Tracing the evolution of Marek's disease virus virulence using ancient and historical collections

Martin Tesicky

LMU Munich, Germany

Marek's disease virus (MDV) is a highly contagious poultry pathogen that induces lymphoid tumors, causing mortality rate of up to 90% in unvaccinated flocks. In our previous work (Fiddaman et al., Science 2023), using ancient DNA we demonstrated that MDV has been circulating in chickens for at least 1,000 years and, importantly, that these ancient strains were non-tumorigenic, unlike their modern counterparts. These findings suggest that MDV virulence has evolved considerably over the last millennium, possibly driven by factors such as imperfect vaccination practices, which can favor the proliferation of more virulent strains, and the dramatic increase in global chicken populations in the 20th century. To investigate this hypothesis and to quantify the selective pressures acting on the MDV genome during the 20th century, we screened 200 historical chicken samples (1800s–1960s) from museum collections for the presence of MDV. We observed a high prevalence of MDV (~20%) in these historical samples. Employing a targeted capture approach, we generated a dense time series of over 50 high-coverage MDV genomes over the last 200 years. This genomic time series enables us to determine the timing, intensity, and targets of selection within the MDV genome during this transformative era. These results underscore the value of historical collections for tracking the molecular evolution of virulence in economically important pathogens.

Symposium 22: Thematic symposium: evolution theory and recent methodological advances

The Generalized Haldane (GH) model tracking population size changes and resolving paradoxes of genetic drift

Yongsen Ruan

Sun Yat-sen University

Genetic drift, the random fluctuation of allele frequencies, is a fundamental evolutionary force. The Wright-Fisher (WF) model, which describes drift as inversely proportional to population size (1/N or 1/Ne) through binomial sampling, has been a cornerstone in population genetics. However, it fails to account for complexities observed in real populations, where drift can intensify with increasing population size due to heightened competition. In contrast, J.B.S. Haldane's branching process model offers an alternative framework, where alleles independently produce offspring, and genetic drift is determined by the variance in offspring number (V(K)/N). This model, however, allows population size (N) to drift unboundedly, prompting the development of the Generalized Haldane (GH) model, which constrains N while preserving Haldane's core principles. The GH model resolves paradoxes in genetic drift, explaining how drift can intensify in larger populations under specific ecological conditions, contrary to conventional expectations. It shows that small populations, especially those undergoing exponential growth, experience minimal drift, while populations near carrying capacity display paradoxical patterns observed in both field studies and experiments. Thus, the GH model reconciles the apparent contradiction where drift increases with population size in certain settings. A key application of the GH model is its explanation of sex-biased drift, where males, due to exceptional reproductive output (ERO), experience stronger drift than females. Evidence from humans and apes reveals a disproportionate contribution of a few males to the gene pool, emphasizing the role of ERO in shaping evolutionary dynamics.

On the probability of reaching high peaks in fitness landscapes by adaptive walks

Yang Li

University of Michigan

Adaptive evolution can be described by an uphill walk in a fitness landscape. However, climbing the global peak in a multi-peak landscape is improbable because of the high chance of being trapped at a local peak. Nonetheless, over three-quarters of simulated adaptive walks in the fitness landscape of the Escherichia coli dihydrofolate reductase (DHFR) gene were reported to end at the highest 14% of peaks, suggesting that biological systems may be substantially more evolvable than commonly thought. To investigate the cause and generality of this observation, we estimate in empirical and theoretical fitness landscapes the probability of reaching high peaks by adaptive walks (PHP), where high peaks refer to the highest 1, 5, 14, or 25% of all peaks. We find that (1) PHP varies substantially among landscapes, (2) PHP in empirical landscapes is generally comparable to or smaller than that in same-size Rough Mount Fuji landscapes of similar ruggedness, and (3) lowering landscape ruggedness boosts PHP. As observed in DHFR, we find in every examined landscape a positive correlation between the fitness of a peak and its basin size, which is the number of genotypes that can reach the peak through adaptive walks. Yet, this correlation does not guarantee a large PHP because of the influences of other factors. We conclude that evolvability depends on the specific fitness landscape and that the large PHP in the DHFR landscape results from its relatively low ruggedness and is not a general property of empirical or theoretical fitness landscapes.

Fluctuating selection among individuals (FSI): a new theory of molecular evolution

Xun Gu

Iowa State University

Fitness of a mutation could vary among individuals who carry this mutation, due to different genetic or epigenetic backgrounds. Fluctuating Selection among Individuals (FSI) refers to the phenomenon that a mutation tends to have a broader fitness variation among individuals than the wildtype. I developed a novel Wright-Fisher diffusion model under FSI, challenging some current wisdom of molecular evolution. First, FSI plays a major role in genetic drift rather than the effective population size (Ne), except for a very small population. This result explains why dN/dS (the ratio of nonsynonymous to synonymous rates of protein sequences) is inversely related to log-Ne in a broad range from mammals, amniotes to metazoans. Second, the MacDonald-Kreitman (MK) test should be reinterpreted by the means of FSI. Let pN/pS be the ratio of nonsynonymous to synonymous diversity within population. FSI theory showed that the null of MK test 0 corresponds to , where is the coefficient of selection averaged over individuals, the FSI-coefficient measures the strength of FSI, and the constant >0. Because dN/dS=1/(1+2Ne) when , an intriguing feature called selection-duality emerges: around the null of MK test (, the mutation is expected to be, on average, beneficial in fitness (, but subject to a negative selection during the across-species evolution (dN/dS<1). Further analysis based on metazoan genomics demonstrated that the overwhelming excesses (α>0) is likely the consequence of FSI without adaptation. The current study may shed lights on resolving the long-term selection-neutrality controversy.

A narrow range of transcript-error rates across the Tree of Life

Weiyi Li

Stanford 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 across the Tree of Life using a modified rolling-circle sequencing method, revealing that the range in error rates is remarkably narrow across diverse species. Transcript errors tend to be randomly distributed, with little evidence supporting local control of error rates associated with gene-expression levels. A majority of transcript errors result in missense errors if translated, and as with a fraction of nonsense transcript errors, these are underrepresented relative to random expectations, suggesting the existence of mechanisms for purging some such errors. To quantitatively understand how natural selection and random genetic drift might shape transcript-error rates across species, we present a model based on cell biology and population genetics, incorporating information on cell volume, proteome size, average degree of exposure of individual errors, and effective population size. However, while this model provides a framework for understanding the evolution of this highly conserved trait, as currently structured it explains only 20% of the variation in the data, suggesting a need for further theoretical work in this area.

Emergence of oscillatory balanced polymorphism in age-structured populations

Yuseob Kim

Ewha Womans University

10 years ago, a surprisingly large number of loci in the fruit fly genome were discovered to harbor polymorphism in which allele frequencies oscillated in seasonal cycles. This suggested that fluctuating selection under variable environments is prevalent. However, it was questioned which evolutionary process can produce such multi-locus oscillatory polymorphism because, according to the classical theory, changing directions in selection rather makes it difficult to maintain genetic variation. Recent progress in theoretical investigations now provides an answer. If a subset of a population – a particular group of individuals or particular stage in the life cycle – is protected from selection, negative frequency-dependent selection on alleles with fluctuating fitness arises. This diversity-promoting mechanism – discovered first in community ecology – is called the 'storage effect'. While the storage effect is expected to arise in various biological settings, this study focuses on the age structure with overlapping generations, a condition satisfied by most animals, including fruit flies. Using the analysis of simple models and computer simulations, I demonstrate that environment-driven population size changes generate an even stronger force of balancing selection, called the 'noncovarying storage effect (NSE)'. With random demographic and fitness fluctuations, mutant alleles that would otherwise be eliminated are positively selected by NSE and are either maintained in polymorphism with oscillating frequencies or reach fixation in the population. This NSE in age-structured populations is quite robust to perturbations in model parameters and can potentially generate oscillatory polymorphism at a large number of loci as observed in the fly data.

Assessing phylogenetic posteriors for single-cell DNA sequencing tumour data: A simulation study of variant calling parameters

Eva Li

University of Auckland

Single-cell DNA sequencing (SCS) provides valuable insights into tumour heterogeneity and enables the reconstruction of tumour cell phylogenies, but its high error rates present significant challenges. We aim to analyse the posterior distributions generated by Bayesian phylodynamic inference for SCS tumour data. Our study involves three experimental setups using the same pipeline but different simulation models: (1) a simple tree with strict clock model, (2) a calibrated tree with local clock model, and (3) a more realistic tree with local clock model. Key steps in our pipeline include: (i) simulating trees and sequence alignments with implemented LinguaPhylo (LPhy) models, (ii) evaluating how different variant calling parameters mitigate sequencing noise, and (iii) estimating phylogenetic features through Bayesian analysis while assessing the performance of error models. This simulation pipeline is designed to generate realistic sequences, identify optimal variant calling parameters to minimize noise, and improve the accuracy of tumour cell phylogenetic inference. Future work will extend this pipeline by integrating machine learning models to simulate more realistic sequences and applying deep learning to enhance variant detection sensitivity on SCS genome-wide data, with potential applications in tumour patient treatment.

The impact of non-neutral synonymous mutations when inferring selection on non-synonymous mutations

Aina Martinez i Zurita

UCLA

The distribution of fitness effects (DFE) describes the proportions of new mutations that have different effects on reproductive fitness. Accurate measurements of the DFE are important because the DFE is a fundamental parameter in evolutionary genetics and has implications for our understanding of phenomena like complex disease or inbreeding depression. Current methods to infer the DFE for nonsynonymous mutations from natural variation first estimate demographic parameters from synonymous variants to control for the effects of demography and background selection. Then, conditional on these parameters, the DFE is inferred for nonsynonymous mutations. This approach relies on assuming that synonymous variants are neutrally evolving. However, some evidence points toward synonymous mutations having measurable effects on fitness. To test whether selection on synonymous mutations affects inference of the DFE we simulated several models of selection on synonymous mutations and attempted to recover the DFE of nonsynonymous mutations using Fit∂a∂i, a common method for DFE inference. We conducted simulations with a constant population size and with a human-like demographic history. Our results show that the presence of selection on synonymous variants leads to incorrect inferences of recent population growth. Furthermore, under certain parameter combinations featuring pervasive selection on synonymous mutations, inferences of the DFE may overestimate the proportion of weakly deleterious nonsynonymous mutations. This bias can be eliminated if the correct demographic parameters are used for DFE inference instead of the biased ones inferred from synonymous variants. Our work demonstrates how unmodeled selection on synonymous mutations may affect downstream inferences of the DFE.

Deep learning: Balancing, linkage and effects of selection (DEEPBLUES)

Carolin Kosiol

University of St Andrews

The phenomenon of Balancing Selection (BS) observed across a multitude of organisms refers to the processes that sustain genetic variation over large numbers of generations in populations. After decades of debate about the role of BS, new life has been breathed into the topic of BS with advanced genomic techniques. The subtle nature of BS, and its complex interactions with factors such as LD have significant effects even over timescales longer than speciation events.The implementation of polymorphism-aware phylogenetic models (PoMos) allows the inference of phylogenies and directional selection and more importantly BS with the recently developed PoMoBalance. We adopt a methodology that involves the training of neural network (NN) models, using convolutional NNs and feed-forward NNs. Features extracted from our data sets are fed into a machine learning framework. The NNs will thus be applied to data simulated with SLiM in different scenarios combining LD and BS with corresponding ancestral recombination graphs (ARGs). Training and testing data are simulated for a great ape population scenario, incorporating the effect of dominance at intermediate allele frequencies that represent BS.

In the first instance, we rely on a large number of population genetics summary statistics (SS), including Tajima's D and site frequency spectra, among others. Additionally, we also rely on SS that summarise features of ancestral recombination graphs (ARGs) calculated from branch lengths, tree topology and lineage-through-timeplot. Here we present a comparison of the performance of the different sets of SS in the context of balancing selection and linkage disequilibrium.

Distinguishing false positives from real gene-by-gene interactions between nearby variants

Anastasia Ignatieva

University of Oxford

A common phenomenon which significantly complicates statistical testing for gene-by-gene interactions is phantom epistasis, where the omission of a causal variant (with a purely additive effect on the phenotype) causes the spurious inference of a significant interaction between two SNPs. Phantom epistasis is more likely to arise between SNPs in close proximity, making it very difficult to tell apart true epistatic signals in cis from false positives, a pervasive issue in previous studies claiming to have identified such signals. We characterise the causes of phantom epistasis from a genealogy-based perspective, and use reconstructed genealogies (in the form of ancestral recombination graphs) to develop a novel method for determining whether a given pairwise interaction is likely to be the result of phantom epistasis. Our method can be applied to pairs of SNPs regardless of the genetic distance between them, uses only publicly available data (not requiring access to the phenotypes and genotypes used for detecting interactions), and can be especially useful when imputation using large reference panels is not possible (for instance, when studying species other than humans). We use simulations to show that the method achieves good sensitivity at even low genetic distances (around 0.5cM), and applying the method to real data from previous studies demonstrates high power to detect phantom epistasis. This presents a novel methodological advance towards the exciting goal of distinguishing spurious interactions in cis from those reflective of real biological effects.

PLVFormer: A transformer-based tree rearrangement for phylogenetic tree search

Trong Nhan Ly

Australian National University

Widely used phylogenetic inference tools, such as IQ-TREE, RAxML, and PHYML, rely on heuristic tree search algorithms to construct phylogenetic trees that maximize the likelihood of observed genetic sequence alignments. However, these tree search methods are computationally intensive and often prone to local optima. To address these challenges, we present PLVFormer, a new method based on Transformer - the neural network architecture behind ChatGPT and other AI tools - to propose tree rearrangements while searching in the tree space. PLVFormer applies the Phyloformer architecture in order to reconstruct a connected region of the tree based on partial likelihood vectors (PLVs) at the boundary of the connected region. Experimental results show that PLVFormer obtains reasonable accuracy while being significantly faster than IQ-TREE. For example, IQ-TREE required a total of 1.6 days to reconstruct 1,000 phylogenetic trees from 20-taxon alignments while PLVFormer took only 5 minutes, a speedup of over 450 times.

A comprehensive representation of selection at loci with multiple alleles that allows complex forms of genotypic fitness

Nikolas Vellnow

TU Dortmund University

Genetic diversity is central to the process of evolution. Both natural selection and random genetic drift are influenced by the level of genetic diversity of a population; selection acts on diversity while drift samples from it. At a given locus in a diploid population, each individual carries only two alleles, but the population as a whole can possess a much larger number of alleles, with the upper limit constrained by twice the population size. This allows for many possible types of homozygotes and heterozygotes. Moreover, there are biologically important loci, for example those related to the MHC complex, the ABO blood types, and cystic fibrosis, that exhibit a large number of alleles. Despite this, much of population genetic theory, and data analysis, are limited to considering biallelic loci. However, to the present, what is lacking is a flexible expression for the force of selection that allows an arbitrary number of alleles (and hence an arbitrary number of heterozygotes), along with a variety of forms of fitness. In this work, we remedy this absence by giving an analytical representation of the force of selection that emphasises the very different roles played by the diversity of the population, and the fitnesses of different genotypes. The result presented facilitates our understanding and applies in a variety of different situations involving multiple alleles. This includes situations where fitnesses are: additive, multiplicative, randomly fluctuating, frequency-dependent, and it allows fitnesses which involve explicit gene interactions, such as heterozygote advantage.

PhyloVAE: Unsupervised learning of phylogenetic trees via variational autoencoders

Tianyu Xie

Peking University

Learning informative representations of phylogenetic tree structures is essential for analyzing evolutionary relationships. Classical distance-based methods have been widely used to project phylogenetic trees into Euclidean space, but they are often sensitive to the choice of distance metric and may lack sufficient resolution. In this paper, we introduce phylogenetic variational autoencoders (PhyloVAEs), an unsupervised learning framework designed for representation learning and generative modeling of tree topologies. Leveraging an efficient encoding mechanism inspired by autoregressive tree topology generation, we develop a deep latent-variable generative model that facilitates fast, parallelized topology generation. PhyloVAE combines this generative model with a collaborative inference model based on learnable topological features, allowing for high-resolution representations of phylogenetic tree samples. Extensive experiments demonstrate PhyloVAE's robust representation learning capabilities and fast generation of phylogenetic tree topologies.

Evolutionary dynamics of polygenic adaptation under diverse ecological scenarios

Yuna Zhang

University of Cologne

An important goal of ecological genetics is to understand how adaptation occurs in new environments. Quantitative traits, which encompass a wide range of functionally important characteristics, are controlled by numerous genetic loci. These traits are likely subjected to stabilizing selection in static environments, but changing environments lead to shifts in trait optima and directional selection. Hence, it is essential to understand how quantitative traits behave during environmental change and what factors influence adaptability. To address this, we used forward-in-time simulations to study the phenotypic and genetic changes within populations adapting to new trait optima under three ecological scenarios: sudden, gradual and fluctuating shifts. We varied trait architecture and the distance to the new phenotypic optimum to model different magnitudes of environmental change. Our results show that sudden shifts are more likely to result in extinction compared to gradual and fluctuating scenarios. Adaptation is primarily driven by segregating sites rather than fixed mutations, with standing genetic variations playing a key role. Genetic variance, commonly used as an indicator of adaptability, is strongly influenced by standing variations during sudden shifts while new mutations contribute more to genetic variance during gradual and fluctuating shifts. Notably, traits governed by a large number of small-effect alleles benefit more from standing variations, enabling faster adaptation to distant optima. In summary, our findings reveal the complex interplay between environmental dynamics and trait architectures, emphasizing the importance of standing variation for adaptation and the need to mitigate rapid environmental changes for species conservation.

Symposium 23: Thematic symposium: Symbiosis

Spatial transcriptomics and comparative genomics reveal key evolutionary innovations in a symbiotic bioluminescent Flashlight fish

Chang Liu

Sun Yat-sen University

We still lack understanding of the genomic changes underlying the coevolution between endosymbionts and hosts. The Flashlight fish Anomalops katoptron hosts an obligate endosymbiotic bacterium Candidatus Photodesmus katoptron in their specialized subocular light organ, providing bioluminescence for complex social behaviors. We first employ single-cell spatial transcriptomics of the light organ to identify key genes and novel cell types involved in this symbiosis. Next, by combining de novo genome assembly, comparative genomics and population genetics, we elucidate the selection pressures and genetic changes that shaped the evolution of the light organ, with a focus on understanding the molecular evolutionary processes that led to the emergence of bioluminescence in A. katoptron from non-luminescent ancestors. Our study sheds light on how new symbiotic relations can evolve from existing gene networks in non symbiotic ancestors.

Microbiome responses to climate-induced stressors and implications for host and ecosystem health

Jingdi Li

University of British Columbia

Climate change is driving unprecedented global temperature increases, which could be linked to rising infectious disease risks and pose significant threats to species and ecosystems. All animals and plants harbor microbiomes that play crucial roles in host health and fitness. Due to their vast population sizes and shorter generation times, host-associated microbiomes often exhibit heightened sensitivity and adaptability to environmental changes and biotic stressors, making them key indicators of altered host fitness. Understanding how microbiomes respond to abiotic factors, such as temperature fluctuations, and biotic factors, such as pathogen invasion, is critical for disentangling host fitness responses and associated risks. Here we explore general microbiome responses to warming and cooling, as well as their responses to the combined pressures of temperature stress and pathogen invasion. Our ultimate goal is to investigate how manipulating microbiomes can mitigate these stressors, enhance host resilience, and ultimately contribute to ecosystem and public health.

Interactions between the honeybee immune system and gut bacteria

Shiqi Luo

China Agricultural University

Honeybee workers maintain a stable gut microbiota that interacts closely with honeybee immunity. Honeybees regulate their immune system precisely to maintain the colonization of host-specific gut symbiotic bacteria and create an inhospitable environment for foreign bacteria, ensuring symbiotic specificity. Inoculating honeybees with the gut bacteria Gilliamella strain from bumblebees increases prostaglandin levels in the honeybee gut, activating the honeybee immune response via the immune deficiency and Toll pathways. This leads to elevated Duox expression, reactive oxygen species production, and suppression of the bacteria strain from bumblebees. Additionally, gut bacteria contribute to honeybee immune defense by producing butyrate, enhancing honeybee resistance to pathogens. This process involves butyrate-mediated gene expression regulation, which stimulates the production of immune-related prostaglandins, thereby increasing systemic cellular and humoral immune responses in honeybees. Overall, the interaction between honeybees and gut bacteria represents a dynamic equilibrium that maintains symbiotic stability and enhances honeybee resilience against pathogens, underscoring the intricate nature of host-microbe interactions.

Taxonomically different symbiotic communities of sympatric Arctic sponge species show functional similarity with specialization at species level

Anastasiia Rusanova

Koltzov Institute of Developmental Biology

Marine sponges harbor diverse communities of associated organisms, including eukaryotes, viruses, and bacteria. Associated microbiomes contribute to the health of the host organisms by defending them against invading bacteria and providing them with essential metabolites. Here we describe bacterial communities of three sympatric species of cold-water marine sponges – Halichondria panicea, Halichondria sitiens, and Isodictya palmata – sampled over the period of six years at the White Sea. Using FISH microscopy, 16S and shotgun metagenomics, we identified sponge species as low microbial abundance (LMA) and detected stably associated bacteria that represent new taxa of unculturable sponge symbionts within Alpha- and Gammaproteobacteria. Studied sponges carried unique sets of unrelated species of symbiotic bacteria illustrating varying complexity of microbiomes: I. palmata had a minimalistic microbiome with a single dominant symbiont, H. panicea carried one dominant and one minor symbiont, while H. sitiens had one dominant and three minor symbionts specialized in sulfur metabolism. On a community level, microbiomes shared common symbiotic features; they encoded multiple eukaryotic-like proteins (ELPs), biosynthetic pathways, and transporters of amino acids and vitamins essential for sponges. On species level, however, different classes of ELPs and pathways were distributed between dominant and minor symbionts indicating specialization within microbiomes. Particularly, taurine and sulfoacetate metabolism pathways were associated exclusively with dominant symbionts in all three sponge species. Our study demonstrates strong functional convergence and co-evolution of microbiomes of sympatric cold-water sponge species with distribution of functions between community members.

Relaxed selection on Blattabacterium genes led to convergent gene losses

Zhuli Cheng

Okinawa Institute of Science and Technology

Genomes of insect endosymbionts are well-known for losing most of their genes, but the processes and causes of gene losses remain unclear. Here we present a study that compared the selection pressures acting on the 596 protein-coding genes of Blattabacterium, the obligate endosymbiont of cockroaches. We used dN/dS ratio as the measure for selection pressure and found that this ratio varied drastically across the genes. We further found that genes that went through independent losses in multiple cockroach families exhibited higher dN/dS ratios, indicating relaxed selection. These findings suggest that relaxed selection is a key driver of gene loss in Blattabacterium, shedding lights on the role of natural selection in genome erosion in insect endosymbionts.

Molecular adaptation of two hot-vent chemosymbiotic Alviniconcha snails revealed by multiple-omics analyses

Hui Wang

Ocean University of China

Deep-sea hydrothermal vents represent an ‘extreme’ environment with highly fluctuating and stressful physicochemical conditions for animals, where specially adapted animals form dense aggregations, mostly those relying on chemosymbionts. The molecular cross-talking between the host animal and their symbionts at vents has evolutionary significance that favor the survival of the holobiont to take advantage of such environment. A major example of such holobiont is the hairy snail genus Alviniconcha which enlarged gills which house Gammaproteobacterial or Campylobacterota symbionts at the intermediate phase between intracellular and extracelluar, but the molecular mechanisms allowing the host animal to carry out this symbiosis remained unclear. Here, we assembled two high-quality Alviniconcha genomes, one harboring Gammaproteobacteria (A. adamantis) and the other harboring Campylobacterota (A. marisindica), to explore the adaptive strategies shared by these two species. From the gene family analyses, we show that the solute carrier gene families which can transport various nutrients are significantly expanded. In the gill of Alviniconcha we detected high protein and gene expression of myoglobin as opposed to previously suggested hemoglobin, suggesting a strategy of storing oxygen in myoglobin as an adaptation to highly fluctuating oxygen availability in the environment. Comparing the gene expression level in the bacteriocytes versus other parts in the gill filaments highlighted that the host likely obtain nutrients from the symbionts predominantly by digestion via phagocytosis, but the milking without digestion was also present. Overall, our results provide a holistic view of host-symbiont interactions, revealing key adaptations of Alviniconcha snails to the deep-sea hydrothermal vent.

Coevolution of host and gut bacteria via metabolism synergy during honeybee acclimation

Min Tang

Xi'an Jiaotong-Liverpool University

Our mechanistic understanding of how gut symbionts coevolve with hosts during host adaptation is limited. Honeybees, with long-term co-diversified gut bacteria, adapted to temperate climates and provide an ideal model. We reveal brown adipose-like fat synthesis via host-symbiont synergy using metagenomics and genomics, in vitro and in vivo experiments, saccharide metabolome, and transcriptome. Compared to tropical honeybees, cold-adapted honeybees possess enhanced genetic capacity for glucose, pyruvate, lipid and glucuronate production. Significantly, enriched gut bacteria Gilliamella and host-specific Lactobacillus can promote host lipogenesis and thermogenesis by providing more glucose and pyruvate. Two gut symbionts both excel in hydrolyzing β-glucan to glucose but coevolve with hosts to different extents. Gilliamella formed an intimate metabolic interaction with hosts that promotes lipogenesis with pyruvate and D-xylulose-5P, while selected Lactobacillus strains reduced specific functions and saved pyruvate to hosts. By different pathways, both bacteria restrained their population size and thus reduced competition with hosts. The highly coordinated host-symbiont metabolisms give insights into honeybee physiological adaptations to temperate climates, highlighting the diversity of host-symbiont interactions in long-term coevolutionary history.

Genomic Underpinnings of Cytoplasmic Incompatibility

Yongjun Tan

Saint Louis University

Cytoplasmic incompatibility (CI), a non-Mendelian genetic phenomenon, involves the manipulation of host reproduction by Wolbachia, a maternally transmitted alphaproteobacterium. The underlying mechanism is centered around the CI Factor (CIF) system governed by two genes, cifA and cifB, where cifB induces embryonic lethality, and cifA counteracts it. Recent investigations have unveiled intriguing facets of this system, including diverse cifB variants, prophage association in specific strains, copy number variation, and rapid component divergence, hinting at a complex evolutionary history. We utilized comparative genomics to systematically classify CIF systems, analyze their locus structure and domain architectures, and reconstruct their diversification and evolutionary trajectories. Our new classification identifies ten distinct CIF types, featuring not just versions present in Wolbachia, but also other intracellular bacteria, and eukaryotic hosts. Significantly, our analysis of CIF loci reveals remarkable variability in gene composition and organization, encompassing an array of diverse endonucleases, variable toxin domains, deubiquitinating peptidases, prophages, and transposons. We present that the components within the loci have been diversifying their sequences and domain architectures through extensive, independent lateral transfers and interlocus recombination involving gene conversion. The association with diverse transposons and prophages, coupled with selective pressures from host immunity, likely underpins the emergence of CIF loci as recombination hotspots. Our investigation also posits the origin of CifB-REase domains from mobile elements akin to CR (Crinkler-RHS-type) effectors and Tribolium Medea1 factor, which is linked to another non-Mendelian genetic phenomenon. This comprehensive genomic analysis offers novel insights into the molecular evolution and genomic foundations of Wolbachia-mediated host reproductive control.

Symposium 24: Thematic symposium: colonizing new environments and invasiveness

Impacts of genome architecture on rapid parallel adaptation during habitat invasions

Carol Eunmi Lee

University of Wisconsin-Madison

An extremely high proportion of destructive aquatic invaders are invading from saline to freshwater habitats. Understanding genetic mechanisms of parallel evolution during invasions could help uncover mechanisms of invasive success. Populations of the copepod Eurytemora affinis species complex have recently invaded many freshwater habitats worldwide. During these invasions, distinct clades (sibling species) with different chromosome number show remarkable parallel evolution, with the same loci under selection during invasions. To determine mechanisms of parallel adaptation, we imposed salinity reduction for 10-20 generations using clades with 15 chromosomes (E. affinis proper, ancestral state) and 4 chromosomes (E. carolleeae, fused state). In both clades, ion transport-related genes were the predominant functional category under selection. However, selection acted very differently on selection lines from the two clades, with strong signals of positive epistasis among beneficial alleles in the 15-chromosome clade but selection on standing variation in the 4-chromosome clade. In the 15-chromosome genome, beneficial alleles are concentrated near the telomeres, where recombination and positive epistasis could bring them together within genomes. Simulations indicated that under high epistasis, higher chromosome number enhances selection response and parallelism between replicate selection lines. In contrast, in the 4-chromosome genome, beneficial alleles are already linked near fusion sites (especially centromeres) where recombination is low, such that selection acts on balanced polymorphisms of large linkage groups. Our results reveal that genome architecture has profound impacts on adaptation of populations during habitat invasions.

Evolution of Darwin's Finches in the Galápagos Islands: Genomics perspectives on adaptation, speciation, and biological invasion.

Sangeet Lamichhaney

Kent State University

Island ecosystems serve as natural laboratories for understanding the molecular mechanisms underlying key evolutionary processes such as adaptive evolution, speciation, and biological invasion. Darwin's finches, inhabiting the Galápagos Islands, represent one of the most iconic systems for studying these evolutionary processes. As a part of our long-term research in the Galápagos Islands, we have conducted population-scale whole-genome re-sequencing for > 400 individuals of Darwin's finches and have identified two major genetic mechanisms key to their evolution (1) Two transcription factors, ALX1 and HMGA2, affecting craniofacial development, have driven beak diversification, enabling the finches to utilize a broader range of food resources (2) Extensive interspecific gene flow has been critical in maintaining genetic diversity both within and between species, including a documented case of rapid hybrid speciation. Furthermore, recently, the invasive avian vampire fly (Philornis downsi) has become one of the greatest threats to the Darwin's finches in Galápagos, that parasitizes their chicks and cause high mortality. We have generated a high-quality reference genome and conducted population-scale whole-genome re-sequencing of P. downsi populations from the Galápagos Islands and mainland Ecuador. Our findings have revealed reduced genetic diversity in Galápagos (indicative of a recent bottleneck), ongoing gene flow among island populations, and positive selection in genes involved in neural signaling, muscle development, and metabolism, traits that have likely facilitated the fly's successful invasion in Galápagos. Our study offers a framework for investigating the genetic mechanisms underlying species colonization, local adaptation, and dynamics of invasive species within island ecosystems.

From sub-Saharan Africa to China: evolutionary history and adaption of Drosophila melanogaster revealed by population genomics

Junhao Chen

Peking University

Drosophila melanogaster is a widely used model organism for studying environmental adaptation. However, the genetic diversity of populations in Asia is poorly understood, leaving a notable gap in our knowledge of the global evolution and adaptation of this species. We sequenced genomes of 292 D. melanogaster strains from various ecological settings in China and analyzed them along with previously published genome sequences. We have identified six global genetic ancestry groups, despite the presence of widespread genetic admixture. The strains from China represent a unique ancestry group, although detectable differentiation exists among populations within China. We deciphered the global migration and demography of D. melanogaster, and identified widespread signals of adaptation, including genetic changes in response to insecticides. We validated the effects of insecticide-resistance variants using population cage trials and deep sequencing. This work highlights the importance of population genomics in understanding the genetic underpinnings of adaptation, an effort that is particularly relevant given the deterioration of ecosystems.

Cooperation shapes bacterial niche breadth evolution and patterns of diversification

Chunhui Hao

East China Normal University

Bacteria exhibit varying niche breadths, with generalists thriving in diverse environments and specialists confined to specific habitats. This variability reflects the adaptability of bacteria to their environment and may influence their speciation and extinction rates. We used phylogenetic causal inference and diversification analysis techniques to investigate the influence of cooperation on bacterial niche breadth evolution and patterns of diversification across 25,785 species. We found: (1) a positive correlation between the proportion of genes for cooperation and niche breadth; (2) a decreased proportion of genes for cooperation promotes niche contraction; and (3) species with a higher proportion of genes for cooperation show increased speciation and extinction rates when their niche breadths are narrower. Our study highlights the role of genes for cooperation in shaping both niche breadth and patterns of diversification of bacteria, underscoring their critical function in maintaining the ecological versatility and diversity of bacteria.

Nutritional Adaptation of Drosophila suzukii Larvae

Yan Hou

Westlake Laboratory

Understanding the genetic foundations of niche adaptation is crucial for elucidating the molecular mechanisms that drive the emergence of novel traits and strategies to cope with varied ecological environments. Nutritional profiles, particularly the protein-to-carbohydrate (P:C) ratio, vary across habitats and significantly influence physiological and metabolic responses. Drosophila suzukii, a major invasive pest of fruits, has larvae that thrive on fresh fruits—a low-protein, high-sugar habitat—differ from most congeners that inhabit decaying food, rich in protein and carbohydrates. To explore the genetic basis of D. suzukii's adaptation to this specialized dietary niche, larval performance was compared across D. suzukii and its relatives, D. biarmipes and D. melanogaster, on diets with varying P:C ratios. Drosophila suzukii larvae exhibited superior performance on low P:C diets, even in the absence of microbial contributions. Transcriptional profiling revealed that, under low P:C conditions, D. suzukii exhibited the fewest significantly differentially expressed genes among the three species, suggesting minimal transcriptional reprogramming under nutritional stress. Candidate genes involved in insulin signaling, feeding regulation, and sensory perception were identified as potential contributors to interspecies differences in performance. Functional validation using the GAL4/UAS system in D. melanogaster confirmed that regulating these genes improved larval development on low-protein, high-sugar diets, supporting their role in D. suzukii's adaptation. These findings offer critical insights into the genetic mechanisms underlying D. suzukii's dietary specialization, which likely drive its ecological success and invasive potential. Furthermore, this study provides a broader perspective for understanding adaptive evolution in Drosophila species.

The Genetic Blueprint of an Invasive Pest: Insights from the Genomes of Historical and Modern German Cockroaches

Sarah Saadain

University of Veterinary Medicine, Vienna

The German cockroach (Blattella germanica) is the world's most widespread cockroach pest, having colonized all inhabited continents over the past centuries. Despite its name, it originates from Asia and is exclusively found in human habitats, relying on heated environments to survive in temperate regions. It has spread globally through international trade, driven by its exceptional ability to adapt to human environments.

A major contributor to this adaptability is its strong resistance to insecticides, having evolved defenses against 42 active ingredients from synthetic chemicals, which has made it a persistent challenge for pest management over the past century. Studies on other insects, such as Drosophila melanogaster, suggest that mobile genetic elements, such as transposable elements, play a role in driving resistance development. Comparing B. germanica with its closest relative, the Asian cockroach (Blattella asahinai), reveals key ecological differences, such as B. germanica's avoidance of light and loss of flight capacity—adaptations that contribute to its unique ecological niche.

In our study, we analyze DNA from 150-year-old specimens of B. germanica collected in Vienna, Austria, and preserved at the Natural History Museum Vienna, alongside modern populations from the same region. Our goal is to uncover the genetic basis of the German cockroach's rapid and repeated adaptations to insecticides and other environmental pressures, shedding light on the mechanisms underlying its invasiveness.

Molecular mechanisms underlying genetic adaptation of Atlantic herring to the brackish Baltic Sea

Fahime Mohamadnejad Sangdehi

Uppsala University

The Atlantic herring (Clupea harengus), an ecologically and commercially significant marine fish, successfully colonized the Baltic Sea about 10,000 years ago despite its stark marine-to-brackish salinity gradient. While Atlantic herring typically spawn in marine waters (~35 ppt salinity), Baltic populations have adapted to much lower salinities (as low as 3 ppt), where their eggs and larvae must survive and develop. However, the mechanisms underlying this adaptation remain largely unknown. Our genome-wide scans between Atlantic and Baltic populations revealed strong signatures of selection at a genomic region containing hatching enzyme genes. These enzymes belong to the C6astacin family of metalloproteases, which play a key role in embryonic hatching by hydrolyzing the egg envelope. Short-read sequencing analyses suggested a striking copy number amplification at this locus in Baltic herring, with a strong negative correlation between gene copy number and water salinity. To resolve the structural complexity of this region, we leveraged long-read sequencing data from 15 individuals, allowing us to study 30 haplotypes. Our findings confirm extensive copy number variation, with some Baltic haplotypes carrying over 30 copies of hatching enzyme genes, while most Atlantic haplotypes lack them entirely. Additional analyses revealed these genes are distributed across multiple genomic regions, suggesting complex evolutionary dynamics. We identify candidate polymorphisms under selection, and plan experimental assays to assess gene expression and enzyme activity in varying salinity conditions. This study provides new insights into the genetic basis of salinity adaptation, enhancing our understanding of molecular mechanisms that enabled the species to colonize the Baltic Sea.

Pathways of introduction and expansion dynamics of invasive alien species in an island ecosystem

Jorge Gutierrez

Instituto de Biología, Universidad Nacional Autónoma de México (UNAM)

The establishment and spread of non-native species into new regions, known as biological invasions, can have significant economic and environmental consequences. Islands, which harbor a disproportionately high share of global terrestrial biodiversity, are particularly vulnerable to invasive species due to predation, competition, habitat disturbance, and disease transmission. Antaxius spinibrachius is an omnivorous bush-cricket primarily found in shrubland and forests of the Iberian Peninsula, with introduced populations in Mayenne (France) and the Macaronesian archipelago of Madeira (Portugal). Within its native range, the species is distributed across the northwestern Iberian Peninsula, with populations surrounding the northern Iberian Plateau. In this study, we used ddRADseq data and sequences of the Cox1 mitochondrial gene to investigate the invasion history of A. spinibrachius on Madeira Island. Our analyses aimed to identify source populations, map dispersal routes, determine the number of independent introductions, and explore key aspects of the invasion's evolutionary and demographic dynamics. Reconstructions of the genetic structure across the species' entire native range revealed a strong phylogeographic pattern. Preliminary results indicate that the source population is located in northwestern Portugal, near Porto. Additionally, a single mitochondrial haplotype was detected across all sampled individuals on Madeira Island, suggesting a single colonization event.

Symposium 25: Advancing Inclusion, Diversity, Equity, and Accessibility in Molecular Biology and Evolution

IDEA Taskforce Presentation

Ravinder Kanda

Oxford Brookes University

Most genomes consist of a substantial amount of sequences derived from transposable elements, including viruses (termed Endogenous Viral Elements, EVEs). While their diversity and impact on genome evolution has been well studied in humans and larger primates, considerably less is known about these insertions in the genomes of other species, especially outside of mammals. With a similar number of extant species and comparable genome sizes, Anurans represent an ideal group for comparison to mammals. We conducted a search for novel Endogenous Retroviruses (ERVs) through publicly available Anuran genomes. We identified 57 novel ERV families, represented by multiple loci, with several families shared across multiple species. The genomic structure of these insertions differs greatly to those currently recognised with retroviral classification, illustrating the mammal-centric bias in our current understanding of these genomic elements, and the previously unknown rich diversity of these genomic elements outside of mammals.

Skeletons in the closet: ethics, law, and politics in palaeontology

Nussaïbah B. Raja Nussaïbah B.

University of Erlangen-Nuremberg

Palaeontology is unique among scientific disciplines in that it thrives on the exchange of information across diverse communities, both academic and non-academic. However, palaeontological research does not always best serve these communities. A practice that is prevalent in palaeontology is "parachute science", referring to the practice whereby researchers drop into, collect data and leave without the involvement or interaction with the local community. Put simply, these parachuting palaeontologists are benefitting from the resources of a country and often, the efforts of local people, without giving anything back. Lack of paleontological and scientific involvement with communities in the Global South may be a remnant of colonial-era palaeontology. The history of natural science is inseparable from the history of European colonialism when local specimens were brought to the homeland of the colonisers to be reposited and studied in museums for the sake of the "greater scientific good". The culture of theft and plunder, a legacy of colonialism perpetuates in palaeontology, even now. The drive for discovering the next new extraordinary fossil can be linked to several ethical and legal issues, where fossils are excavated without record and smuggled across borders to finally end up in collections across the world.

Diversifying the curriculum in a research-intensive university with staff/student co-creation

Matteo Fumagalli

Queen Mary University of London

The diversification of the curriculum is a common approach in inclusive education, showcasing the contribution of a diverse community. It is vital for all students to being able to see themselves in the curriculum, and to recognise the great diversity that has shaped our disciplines, particularly in STEM. A diverse curriculum is widely regarded as an important factor in strengthening access, sense of belonging, and ultimately, success of students from all backgrounds.

Queen Mary University of London has the most diverse and first-gen cohort of students among the top research-intensive universities in England, representing the best case study to implement and evaluate practises to remove barriers. Initiatives like the Centre for Academic Inclusion in Science and Engineering (CAISE) support staff to champions inclusive education and fulfil the university's strategy to become the most inclusive university of our kind.

Here I will present a project co-led with PhD students to create a portfolio of scientists with historically underrepresented identities who have significantly contributed to our taught material in biological and behavioural sciences. We focused on several identities including ethnicity, gender, sexual orientation, religion, neurodivergence, and their intersectionality. The portfolio links profiles with keywords and taught modules making it seamlessly for lecturers to enrich their teaching material. I will showcase examples from the field of molecular evolution. I will also contextualise results from surveys from participating PhD students and teaching staff.

We argue that this transferable framework will enable a diverse research-oriented education, while fostering a meaningful co-creation culture within our universities.

HAAM Community: Fostering collaboration, inclusivity, and sustainable research in Human Ancient DNA, Ancestry, and Mobility

Miren Iraeta-Orbegozo

University of Lausanne, University of Copenhagen

The Human Ancient DNA, Ancestry, and Mobility (HAAM) Community, is an international grassroots network of early-career researchers. Recognising the historical and systemic inequalities within biomolecular archaeology, we are committed to promoting a more equitable future for the field. Through building an open and inclusive community, we aim to foster sustainable, ethical collaboration and networking, bridging divides across different regions and disciplines. Our free virtual initiatives consistently attract large, diverse participation across career levels and global regions. We maintain an online communication space with >280 members, including a channel on which to discuss conceptual, methodological or technical questions. In 2024, we launched the first online Summer School, "Connecting the Past, Present, and Future", which covered both technical and ethical aspects of human ancient DNA research. With volunteer instructors from various career stages, the program received 100+ applicants, 20% of whom were affiliated with groups based in the Global South. Selection prioritized researchers from smaller or underfunded labs, neighbouring disciplines, underrepresented backgrounds and lower-income countries. Our "HAAM-Radio" Webinars give early career researchers the opportunity to present their research to a wider audience outside the usual conference environments. The webinars ensure equitable representation by exploring regional topics from Eurasia, the Americas, Southeast Asia, and Oceania. Collaborating with local researchers, we emphasize cultural sensitivity and inclusivity. Additionally, we are developing an interactive world map of ancient biomolecular research groups, supporting newly established labs, particularly those in the Global South, attempting to mitigate inequalities stemming from limited visibility, underfunding, and unequal access to scientific resources.

Brain drain: recognising the elephant in the room

Mehmet Somel

Middle East Technical University

Low- and middle-income countries supply those in the richer world with not only cheap products but also cheap educated labour. This includes early-stage researchers. Here I argue that researcher brain drain has net negative effects on international scientific production, with examples from molecular evolution. 1) The loss of most well-educated researcher candidates compromises the quality and quantity of local research in the Global South. This leads to: a) a total reduction in global research, b) neglection of local but vital problems (e.g. regional infectious diseases, threatened endemic species), c) local problems being tackled only from the perspectives or interests of academic institutions in the Global North, potentially jeopardizing the interests of local communities or alienating them from science (e.g. population genetics studies that serve racist agendas). 2) The lack of qualified researchers undermines education and public outreach in the Global South, which a) limits the number of young scientists trained worldwide, and b) weakens scientific literacy globally. Overall, brain drain helps deepen socioeconomic inequalities. The structural bases for the drivers of brain drain in science (economic and social advantages of working in the Global North vis-a-vis difficulties of academic positions in the Global South) are formidable and may not be easily altered. Still, I argue that awareness of its global consequences may help mitigate this negative trend and help to improve science and education internationally.

Symposium 26: Associate editor's symposium

Mobility in East and West Eurasia through the Holocene: a comparison of temporal dynamics through paleogenomics

Mehmet Somel

Middle East Technical University

During the Holocene, both West and East Asia experienced similar cultural and economic changes, starting with sedentary life and food production, followed by the adoption of metallurgy and long-distance transport technologies, and the emergence of hierarchical social structures. In recent work, we described mobility and admixture patterns in Southwest Asia over the last 10,000 years as reflected in ancient genomic data. These included several temporal trends: increasing intraregional diversity, regional genetic homogenization during the Neolithic Transition followed by a widening range of mobility after the Bronze Ages, and an increase in relative male-bias in admixture signals. Here we argue that these observed changes in the Southwest Asian gene pool were shaped more by background (continuous, low-level) mobility rather than migration (unidirectional and episodic mass movement events). We further study mobility patterns during the Holocene in East Asia using 1000 published ancient genomes. Interestingly, we do not find the above-mentioned three trends in this region. We propose a number of possible explanations for the observed differences between the West and East Asia, including differences in the scale of mobility events and geographic constraints to regional connections, as well as historical contingency.

Ancient human extinction risk: Implications from high-precision computation and fossil evidence

Haipeng Li

Shanghai Institute of Nutrition and Health, CAS

Reconstructing ancient population size dynamics is essential for understanding human evolution. Using the fast infinitesimal time coalescent process (FitCoal), we identified a severe population bottleneck approximately 930,000 years ago, which pushed human ancestors to near extinction but left only subtle genomic traces in present-day humans. Detecting this signal requires precise computational methods. Compared with mushi and ten billion simulations performed using msprime, FitCoal provides the fastest and most accurate inference of expected branch lengths across site frequency spectrum categories. Independent fossil and paleoclimate records corroborate the timing and evolutionary impact of this bottleneck, highlighting the need to integrate genetic, paleontological, and environmental evidence. These findings refine our understanding of hominin evolution and the demographic events that shaped modern human ancestry.

Out-of-Africa migration and clonal expansion of a recombinant Epstein-Barr virus drives frequent nasopharyngeal carcinoma in southern China

Weiwei Zhai

Institute of Zoology, Chinese Academy of Sciences

Although Epstein-Barr virus (EBV) infection is ubiquitous worldwide, a high-risk subtype linked to the exceptionally high incidence of nasopharyngeal carcinoma (NPC) in southern China has been identified. However, the evolutionary history of EBV in humans and the origins of this oncogenic subtype remain poorly understood. By assembling one of the largest global datasets of EBV genomes, we demonstrate that EBV's evolutionary trajectory mirrors the out-of-Africa migration of modern humans. Within the high-risk subtype prevalent in southern China, we discovered a rapidly expanding clonal lineage originating from a recombination event between ancestral northern and southern Chinese EBV strains approximately 4,000 years ago. This recombination was followed by intense Darwinian selection, conferring a 4% fitness advantage that drove its dominance. Strikingly, this clonal lineage nearly doubles NPC risk compared to the broader high-risk subtype and accounts for ~66% of regional NPC cases, establishing it as the strongest genetic risk factor for NPC identified to date. Our findings reveal a profound co-evolutionary relationship between EBV and humans, wherein historical migration and population admixture facilitated the emergence of a highly virulent EBV strain through recombination. This evolutionary cascade has contributed to the enduring NPC endemic observed in southern China.

Gene tinkering and acid-resistance in Helicobacter pylori

Xuhua Xia

University of Ottawa

Helicobacter pylori can colonize the acidic environment of the mammalian stomach. It is acid-resistant but not acidophilic because it grows better under neutral pH. Its acid-resistance mechanism requires sensing the acidic environment and activating the acid-resistance response, both involving the tinkering of existing genes for new functions. Genomic comparisons between H. pylori and close relatives revealed different ways of tinkering that have evolved during the evolution of the H. pylori lineage, some by point mutations of existing genes and some by appending a sequence encoding metal-binding motifs. I will reconstruct this adaptation process to demonstrate the creativity of nature through mutation and selection.

Modulation of translation accuracy leads to phage resistance

Laasya Samhita

Ashoka University

Mistranslation (errors during protein synthesis) is a potent source of non-genetic variation that occurs in all living cells. In E. coli, about 10% of all proteins are mistranslated at any given time, contributing significantly to the total protein pool. Prior work has established that changes in the accuracy of protein synthesis can alter the rate and route of adaptation when cells are exposed to physical and chemical stressors. Till date, its role in a co-evolution framework remains unexplored. In this work, we test the impact of altering E. coli translation accuracy on adaptation in a host-pathogen system, using the virulent bacteriophage P1vir. We find that both increasing and decreasing the basal rate of mistranslation confers phage resistance, albeit in a mutation-specific manner. Preliminary results indicate that susceptible strains share similar paths to adaptation, while strains that gain resistance due to altered accuracy may be doing so in different ways

Virus-Host Coexistence in Phytoplankton Through the Genomic Lens.

Gwenael Piganeau

CNRS - Oceanological Observatory of Banyuls sur mer

Tiny photosynthetic microorganisms that form the picoplankton (with cells between 0.3 and 3 μm in diameter) are at the base of the food web in marine ecosystems. They are outnumbered by about ten to one by viruses, which can lyse host cells and thereby contribute to biogeochemical cycles by promoting the turnover of plankton populations.

Genomic and metagenomic approaches have revealed an unexpected genetic diversity of both microalga and viruses, at a rate outpacing our ability to identify host-virus specificity patterns and life history traits, which require isolation and cultivation in the lab.

In my lab in Banyuls sur mer, we have been isolating and sequencing many green eukaryotic microalga (Mamiellophyceae) and Prasinoviruses, large double stranded DNA viruses infecting them. Their compact haploid genomes (13 to 22 Mb) and ease of culture make them ideal models to understand the mode and tempo of phytoplankton-virus interactions and their evolution.During this talk, I will present our results obtained from both population genomics and experimental evolution approaches to understand the mechanisms involved in the Mamiellophyceae – Prasinovirus system dynamics, with a special focus on the “omics” bases of resistance and susceptibility in the microalga, combining genomics, transcriptomics and metabolomics data.

Mechanisms driving adaptive evolution in the cross-kingdom fungal pathogen Fusarium oxysporum

Li-Jun Ma

University of Massachusetts Amherst

Fungi constitute a vital component of our ecosystem, significantly impacting agricultural productivity and public health. Climate change has intensified the pressure caused by fungal diseases due to their adaptability in changing environments. Understanding mechanisms driving fungal adaptation, particularly in environments that are associated with disease and impacted by climate change, is vitally important. Fusarium oxysporum is a unique fungal pathogen with cross-kingdom virulence. Using multi-omics and short-term experimental evolution, we studied genome dynamic of a plant pathogenic strain Fol4287 and an eye infecting strain MRL8996. They share a core genome of about 46 Mb, but have distinct transposable element (TE) profiles and encode genes unique for host-specific virulence. Serial passaging of a clonal isolate resulted in populations displaying significantly increased fitness in the respective passaging condition. Sequencing of the evolved lines revealed changes both at the nucleotide and chromosome level, some are fixed in the population. The most detected variants in both evolving passages are TE insertions, as DNA transposons are highly active in Fol4287 and retrotransposons are active in MRL8996. Parallelism or the predictability was observed, as multiple independent lineages evolved under the same conditions carry events at the same locus. Epistasis was detected showing a stepwise optimization of mutations, one depending on the fixation of the other. Collectively, our results provide new insights into the genetic mechanisms underlying fungal adaptation and rectify the power of the experimental evolution in helping us to strategize novel means to control diverse fungal diseases under the rapidly changing environment conditions.

Implications of Sex-Specific Genomic Differentiation in the Great Tit (Parus major) Species Complex

Toni Gossmann

TU Dortmund University

The generation and exposure of genetic novelties to natural selection are central to evolutionary processes and sex-specific variation may function as a source of genomic adaptation. In birds—characterized by highly stable karyotypes and recombination landscapes—rapid evolution appears to occur under seemingly constrained genomic conditions. In this study, we performed comprehensive genomic analyses of more than 200 wild great tits (Parus major)—based on WGS and SNPchip data—from five Asian and seven European populations to uncover substantial cryptic and sex-specific genomic variation. We identified more than 50 autosomal genomic regions spanning over 70 Mb that harbor cryptic genetic elements. These gene-dense regions, which harbor genetic variation that is either female-specific or shared between sexes, show a significant overrepresentation of reproduction-associated genes. Using a long-range deep sequencing approach, we discovered distinct haplotypes enriched in repetitive sequences with high GC and CpG content, which may stem from the sex chromosomes. These haplotypes contain hundreds of novel and divergent homologous genes under purifying selection, indicating functional conservation—particularly in pathways linked to female reproduction and development. We also address the technical implications of genome-wide sex differentiation, examining how sampling composition bias and SNP-based analytical limitations can skew inferences related to local adaptation, gene flow, and population structure. Our findings underscore the importance of rigorous quality control and sex-aware filtering strategies to account for these regions in population genetic analyses. At the same time, they demonstrate how sex-specific genomics can serve as a valuable tool for pinpointing genomic origins of rapid adaptation.

Symposium 27: Population genetics through time session B

Utilizing Identity-by-Descent segments in time-transect ancient DNA data

Harald Ringbauer

MPI EVA Leipzig

The ability to detect pairwise shared long identical haplotypes in ancient human DNA (aDNA) unlocks powerful new ways to study recent evolutionary processes. These so-called identical-by-descent (IBD) segments are ideally suited for this purpose, particularly in aDNA research. A fast recombination clock determines the length of IBD segments, which is more robust to typical biases in error-prone aDNA data than the mutation clock widely used in other inference methods.

Here, I present a new generation of computational tools that use IBD-segment sharing in time-series aDNA data. We have developed IBD-segment-based tools for a broad range of downstream applications by extending analytical frameworks to deal with time-transect sampling. For instance, we can now fit effective population size trajectories using the increased power provided by time-series sampling or estimate the ages of ancient samples of previously unknown date attribution. I showcase the utility of these IBD-segment-based tools through various consequential applications.

ARGMix: Inferring ancient local ancestry with transformers in ancestral recombination graphs

Cole Shanks

University of California Santa Cruz

Local ancestry inference classifies DNA segments by identifying their ancestral origins. However, as the date of admixture becomes older, the task grows increasingly challenging, limiting many existing methods to relatively recent events. The expanding availability of ancient DNA offers a promising opportunity to use these samples as references for local ancestry inference. Previous work has integrated ancient samples into the ancestral recombination graph for local ancestry inference, but does not utilize the full structure of the tree. Here, we introduce ARGMix, a transformer-based model that infers local ancestry by leveraging the inherent structure of the inferred ARG. Our approach uses ancient samples as references in the marginal tree for predicting local ancestry. We train ARGMix on data reflecting well-understood European demographic history and demonstrate its improvement over existing methods, and evaluate it for robustness to misspecified demography. We apply ARGMix to an inferred ancestral recombination graph containing both ancient and present-day European samples, and use ancestry-specific methods to uncover previously unobserved population structure.

Joint parametric inference of complex demographic history using composite likelihood based on whole-genome genealogy

Xinzhu "April" Wei

Cornell University

We present mrpast, a tool for inferring complex demographic history from pairwise-sample coalescences in ancestral recombination graphs (ARGs). mrpast uses a composite likelihood formulation based on the coalescence probabilities from a continuous time Markov process (based on a user-specified model). The input model specifies per epoch the migration paths between populations and population splits, with symbolic parameters for migration rates, effective population sizes, population growth rates, and epoch transition times, all of which are jointly inferred. We evaluated the performance on a variety of models, including stepping-stone models with asymmetric migrations, and population-split models such as the 3-population out-of-Africa models (OOA3G09; previously inferred by ∂a∂i). On most tested models, mrpast is quite accurate when using simulated ARGs, and reasonably accurate using inferred ARGs. For example, the average relative error for the OOA3G09 model is 0.08 with simulated ARGs and 0.33 with inferred ARGs. Using data from the 1000 Genomes project, we explore revised OOA3 models; compared to OOA3G09, we observe different parameter estimates mainly in the recent 3000 generations, possibly explained by site frequency spectrum and ARG having different resolutions for recent history. Moreover, mrpast provides a user-friendly pipeline, which includes simulation, ARG inference, demographic inference, and model selection. The simulation component can help users determine the practical identifiability and accuracy that a model can achieve. As ARG inference techniques improve, mrpast's accuracy should also improve. We hope that mrpast provides a useful pipeline with the flexibility to answer many current and future demographic questions.

Deep-time speciation genomics of mammoth

Kelsey Moreland

Centre for Palaeogenetics

The unique characteristics of a species' genome result from evolutionary changes occurring across considerable timescales, often exceeding several hundreds of thousands of years. While comparative genomics allows for the identification of genome architecture changes across species, pinpointing the exact timing and rate of these evolutionary changes remains challenging when relying on modern snapshot data alone. The Pleistocene evolution of mammoths in Eurasia represents a framework for studying the evolutionary processes of environmental adaptation and speciation. Here we sequenced 29 late-Pleistocene woolly mammoth genomes as well as two mid-Pleistocene mammoths, dated to 1.1 and 0.7 million years old. Combining this with 64 additional woolly mammoth, African elephant, and Asian elephant genomes, we identified the protein-coding and regulatory changes that underlie the formation and speciation of the woolly mammoth across a million-year time transect. We find that most functional changes in woolly mammoths evolved in the time before the first woolly mammoths appear in the fossil record around 0.7 million years ago. Genes involved in fat metabolism, thermoregulation, and immune system functioning have been under continuous selection in woolly mammoths up to their extinction, whereas genes involved in fur and hair development, circadian rhythm, and taste and smell perception had already highly evolved in the mid-Pleistocene mammoths and remained stable since.

Efficient and flexible analysis of fine-scale population structure using rare variants

Lei Huang

Max Planck Institute for Evolutionary Anthropology

Various statistical methods have been developed to identify population structure from genetic data, including F-statistics, which measure the average correlation in allele frequency differences between two pairs of populations. However, the SNPs analyzed in F-statistics are often limited to relatively common allele frequencies as part of microarrays or, in the case of ancient DNA, SNP capture panels. Recent advances in sequencing technology allow us to increasingly obtain whole-genome sequencing data, both ancient and modern, which not only enable querying nearly every base of the genome, but also contain rare genetic variants. Rare variants, with their more population-specific distribution, allow detection of population structure with much finer resolution than common variants – an opportunity that has so far been under-exploited.

Here we develop a new statistical tool, RAS (Rare Allele Sharing), for summarizing rare allele frequency correlations, similar to F-statistics but with flexible ascertainment on allele frequencies. We test RAS-statistics on both published and simulated data and find that RAS estimates have better resolution in distinguishing populations with appropriate ascertainment. Taking advantage of that, we further find possible ways to improve RAS estimates under high levels of SNP missingness, and explore ways to use RAS to derive ancestry proportions, which is highly efficient at high migration rates. Some optimization methods are also adopted to achieve efficient calculation of RAS-statistics. In sum, our method can provide new perspectives to identify and model population structure, allowing us to understand more subtle relationships among populations in the recent human past.

Reconstructing population admixture graph using GTmix2

Yufeng Wu

University of Connecticut

Population admixture is an important genetic process that greatly affects the population demographic history of populations from many species. Admixture graph (AG) has been used extensively as a model for population admixture in the study of the evolution of populations using modern and ancient genomes. In the literature, AGs with ten or more populations and ten or more admixture events are being used for the understanding of complex population evolution. However, reconstructing AG of such complexity is challenging partly due to the astronomical size of possible AG topologies for this type of AGs. In 2020, I developed GTmix, a maximum likelihood AG inference method. GTmix uses inferred gene genealogies and aims at finding the optimal AG under the multispecies coalescent (MSC) model. One drawback of GTmix is that it cannot work with very complex AGs due to its computational inefficiency. In this talk, I will first describe how GTmix works. Then I will present GTmix2, an updated version of GTmix. I will show that GTmix2 significantly outperforms existing AG inference methods including TreeMix and AdmixtureTools2 in terms of inference accuracy. GTmix2 is much faster than the original GTmix. The current implementation of GTmix2 can reconstruct AGs with ten or more populations and up to six admixture events. Future releases of GTmix2 can potentially work for AGs with ten admixture events. Research partly supported by US National Science Foundation under grant IIS-1909425.

Ancient genomic insights into East Asian bovine and caprine domestication: Challenges, discoveries and future

Zehui Chen

Institute of Vertebrate Paleontology and Paleoanthropology

The domestication of Bovidae species–yaks, cattle, goats, and sheep–has significantly shaped East Asian agriculture. Recent ancient DNA (aDNA) studies have revealed complex regional variations and genetic contributions from Early populations to these later domesticated species. In Caprines, the discovery of mosaic origins from Anatolia has reshaped our understanding of the single or multiple origins of the caprine subfamily. Our further research suggests a subsequent domestication event in East Asia, highlighting the region as a secondary center for goat domestication, adding new perspectives on caprine evolutionary history. Additionally, sequencing 40,000-year-old ancient yak genomes has identified a distinct lineage of yak progenitors, contributing differently to modern yak populations on the Qinghai-Tibet Plateau. These findings underscore the resilience and adaptation of East Asian populations to diverse environments, further advancing our understanding of bovine and caprine domestication. Despite these advancements, challenges remain, including the limited availability of well-preserved aDNA and difficulties in quantifying genetic contributions from extinct wild progenitors. Future research should focus on expanding aDNA datasets from East Asia and integrating genomic, archaeological, and isotopic evidence to further explore domestication processes. By utilizing new genomic technologies and interdisciplinary approaches, we can gain deeper insights into the rise of global herding systems.

Population genomic insights from ancient sediment DNA in arctic plant species

Chenyu Jin

Swedish Museum of Natural History

Ancient DNA (aDNA) recovered from sediments has revolutionized our understanding of species presence and ecosystem composition in past environments. While sedimentary aDNA is widely used to detect species across temporal and geographic scales, leveraging this data for population genomic inferences remains a significant challenge. In this study, we analyzed both published and unpublished sedimentary aDNA datasets, ranging from 2 million to 40,000 years old, from multiple sites across the Arctic. We focused on plant species with high DNA abundance and applied an optimized filtering pipeline to identify genome-wide genetic variants. This approach enabled us to investigate population structure and reconstruct the demographic history of Arctic plant species across different periods. Our findings provide novel insights into the genetic diversity and historical dynamics of Arctic ecosystems, highlighting the potential of sedimentary aDNA for population genomic research.

A flexible and efficient new model to infer multiple admixture dates via joint analyses of modern and/or ancient individuals

Nancy Bird

University College London

Inferring admixture dates improves our knowledge of historical migrations, e.g. understanding whether there were multiple waves of intermixing. However, a complication is that individuals are typically combined into a single group to increase power. This assumes, perhaps incorrectly, that they all share the same admixture date(s). We present a new approach that instead starts from admixture dates inferred separately per individual, and uses this to infer the total number of admixture dates across all individuals, and cluster individuals by date. Our approach accommodates individuals with different radiocarbon dates, and hence ancient DNA sampled from multiple time points following the admixture. Furthermore, leveraging shared admixture signals, it can infer the best-fitting sample age for an aDNA sample if supplied with a range of ages. We show with simulations that our method vastly improves date estimates by borrowing strength across individuals, and can accurately identify the number of different admixture dates. We utilise the model to corroborate recent findings of more than one wave of Neanderthal gene flow into modern humans spanning roughly 5000 years. Among other applications, we report analyses of admixture between farmers and hunter-gatherers in Neolithic Europe. We provide new evidence for multiple waves of admixture in Germany and France during the early/middle Neolithic, which contradict conclusions from previous studies that grouped individuals into one population. We demonstrate the power of our new approach, which accommodates output from (e.g.) GLOBETROTTER and DATES, to elucidate previously unreported admixture dynamics using modern and/or ancient genomes.

Dual origin of Turkic speaking peoples

Daniel Tabin

Harvard University

Turkic languages are spoken by an estimated 200 million people across Eurasia from the coast of the Mediterranean to the coast of the Pacific. Although Turkic speakers cover much of Eurasian genetic diversity, their languages spread from Central Asia relatively recently, such that they remain highly mutually intelligible to this day. Despite the recent expansions of Turkic languages, the archaeological identity of the proto-Turkic-speaking population, as well as the expansions of the earliest Turkic-speaking groups, remain mysterious. Using ancient DNA, we find that the proto-Turkic language likely was spoken in the multi-ethnic Xiongnu empire. After the collapse of this empire in the first Century AD-Turkic-speaking populations spread outwards in two pulses. The first wave, which spoke proto-Oghuric, arrived in Europe around 350 AD and likely included the Huns. A second group settled in the Altai Mountains where they mixed with Scytho-Siberians from the Bulan Koby culture. This new group of dual Xiongnu-Bulan-Koby Ancestry spoke proto-Common-Turkic and spread both east and west from the Altai Mountains. This dual origin is corroborated by archeological, linguistic, and historical evidence, including the Chinese Book of Sui. Bulan Koby and Xiongnu ancestry are now ubiquitous among Turkic speaking peoples. Xiongnu ancestry, however, is found in other non-Turkic populations. Bulan Koby ancestry, however, can uniquely identify the genetic legacy left by the expansions of Common-Turkic-speaking populations.

Symposium 28: Human history and past social organisation in the light of palaeogenomes: new methods, new findings session B

The horse before and after our shared history

Ludovic ORLANDO

CAGT (CNRS, France)

The domestication of the horse marked a turning point in human history, accelerating travel, cultural exchange, and gene flow across continents. This long partnership endured until the 20th century, when combustion engines reshaped human mobility. Recent advances in ancient DNA research have refined the timeline of horse domestication, revealing that while early horse husbandry emerged in Central Asia ~5,500 years ago, the lineage of modern domestic horses originated over 1,000 years later in the Pontic-Caspian steppes. Expanding genome time-series data, our study uncovers how Late Pleistocene climate change contributed to shape the genetic diversity of wild horse populations near the time of domestication. We refine the timeline of lineage replacement as domestic horses spread beyond their homeland and identify key genetic adaptations that facilitated their success. Functional validation in model organisms highlights a specific genetic modification likely linked to the rise of horseback-riding and horse-driven mobility.

Ancient genomes from the Imdang-Joyeong burial complex reveal an extensive kinship network in a Three-Kingdoms period society in Korea

Hyoungmin Moon

Seoul National University

The burial complex of the Imdang-Joyeong site at Gyeongsan, Republic of Korea, is notable for the large number of tombs constructed for a short period of late 4th to early 6th centuries CE as well as frequent presence of burials with sacrificed individuals co-buried with the grave owner. To characterize this Three-Kingdoms period society from a genetic perspective, we tried ancient DNA analysis on 182 skeletal elements, resulting in a successful retrieval of genome-wide data from 78 individuals buried at the burial complex. Using the genome-wide data, we detected 11, 20, 23 pairs of the first, second, and three-or-more-distant degree relatives, respectively, providing rich information on the kinship network in the Imdang-Joyeong society. We found 5 individuals from consanguineous marriage, meaning that their parents were close relatives of each other, coinciding with historical records of consanguineous marriage in this region during the Three-Kingdoms period. Within the reconstructed pedigrees, we observed adult female descendants buried together with their kin, unlike several recent archaeogenetic studies in Europe reporting a strict pattern of female exogamy. We also observed no difference in the overall genetic profiles between the grave owners and the sacrificed individuals, confirming the presence of a homogeneous population ancestral to modern Koreans in the Korean peninsula by the Three-Kingdoms period. Our analysis sheds light on the burial customs and social structure at the Imdang-Joyeong site, providing a new perspective on the Three-Kingdoms period society in Korea.

Grave matters: Discerning ancient DNA profiles from grave dirt vs. skeletal remains

Gözde Atağ

Max Planck Institute for Evolutionary Anthropology

Sedimentary ancient DNA (sedaDNA) has become a valuable tool for studying past populations in the absence of skeletal human remains. Here, we present data from five archeological sites ranging from the Bronze Age to Medieval times, to assess whether sedaDNA can substitute for ancient DNA from skeletal elements in various settings and to explore its broader potential. Using target enrichment, we generated mitochondrial and nuclear genome data from both skeletal material and grave sediments. At the Bronze Age site of Koken in Kazakhstan, we also studied sediments from an adjacent settlement area. At Koken, the cemetery samples yielded a higher success rate in recovering ancient human DNA, with six samples allowing mitochondrial haplogroup assignment. Across all sites, nine out of eleven samples matched their bone counterparts. In population genetics analyses, we observed similar patterns in PCA, f-statistics and admixture. Together with the haplogroup inferences, these results suggest that the human DNA in grave sediments primarily originates from the buried individual. At Koken, we also found a high abundance of faunal DNA, primarily from domestic sheep and dogs, in both cemetery and settlement samples. Equids were present only in the settlement, while rodents and cervids were found only in the cemetery. Further analysis of Koken sheep DNA suggests connections with Western Eurasia. Overall, despite the lower yield of human DNA from sediments, it can be possible to achieve a comparable accuracy in genetic studies, and the incorporation of faunal data can allow for further insights via domestic animal movements.

Ancient genome analysis of human remains in the Kofun period on the Japanese archipelago.

Takashi Gakuhari

Kanazawa University

The Kofun period on the Japanese archipelago was characterized by the beginnings of the formation of an early state, and various cultural influences from the continent that differ from the archaeological artifacts seen in the Yayoi period have been found. In archaeological and morphological studies of human remains, human populations from the Yayoi period to the Kofun period have been considered highly genetically continuous, but no study has yet evaluated this hypothesis by genomic analysis of ancient human remains.

We have obtained genomic data from human remains excavated from two types of tombs in the Kofun period in Japan (large keyhole-shaped tombs and horizontal tunnel tombs) that have been radiocarbon dated to the Kofun period. The Kofun genomic datasets were compared with genomic data from the Jomon period, a previously inhabited Japanese population; the Yayoi period population, which migrated from the continent with the rice culture and admixed with the Jomon people; and an ancient population from the continent.

We found that there were genetically distinct characteristics between the keyhole-shaped and horizontal tunnel tombs. In addition, these individuals were found to have high genetic affinity with ancient people on the continent, each showing a different pattern. Individuals with high genetic affinity to ancient humans in the continent were detected, and they showed different admixture patterns among these individuals. Based on the results of comparative genomic analyses and archaeological information, we evaluated the genetic heterogunity in the Kofun period and the genetic continuity from the Yayoi period to the Kofun period.

Ancient DNA analysis unveils the genetic and demographic history of zoroastrians in India

Bhavna Ahlawat

IIT Gandhinagar

Parsis, the descendants of Zoroastrian migrants who fled religious persecution in Persia (modern-day Iran) during the early medieval period, settled along India's western coast. Despite their small population, they have preserved a distinct cultural and religious identity. This study presents the first ancient genome-wide data from 11 individuals from Sanjan, one of the earliest known Parsi settlements in India. Additionally, we generated 36 high-coverage whole-genome sequences from present-day Parsis across western India. In-depth genetic and demographic analyses reveal that ancient Parsis in India share a genetic affinity with Neolithic Iranians, with marginal yet significant gene flow from South Asia. These genetic findings align with literary evidence of the Parsis' migration to India during the 7th–8th century CE. Our study highlights the enduring genetic imprint and ancestry gradients of Neolithic Iranians among both ancient and modern-day Parsis in India. By integrating ancient genomics with historical and archaeological records, this research reconstructs the genetic history of the Parsis and underscores their resilience within a host society. Much like their own metaphor—"Like Sugar in Milk"—the Parsis have blended into the Indian subcontinent while maintaining their distinct identity. This study enhances our understanding of the resilience of micro-religions within larger host societies and sheds light on the complexities of migration, assimilation, and cultural continuity.

Out-of-Anatolia: Genomic and cultural dynamics during the neolithization of Western Eurasia

Dilek Koptekin

University of Lausanne

The genetic histories of the Fertile Crescent and of Europe during Neolithization have been studied extensively. However, the origin of subsequent European agricultural populations is vastly debated. Some argue is has been located in the Aegean around 9,000 BP. To study this question, we sequenced 30 new paleogenomes from early Holocene West and Central Anatolia to show that the emergence of farming settlements in the Aegean was a complex process: the first villages were established not only by migrant farmers from inland Anatolia but also by the descendants of local Aegean foragers. The admixture of these populations then gave rise to the frequently cited "Anatolian Farmer" genetic profile that expanded into Greece and Europe. We further compare the genetic profile among 16 settlements with their material culture, using a quantitative approach comprising a digitalized material culture dataset of 58 unique traits. This analysis reveals a lack of correlation between genetics and cultural traits. This decoupling of demographic and cultural processes at the local level suggests that ideas moved and mixed faster than the genetic signature left by people's movements.

Inferring castration status and age-at-death from sheepskin parchments

Ciarán O'Connor

Trinity College Dublin

Besides meat and dairy, secondary products such as traction from cattle, wool from sheep, and mobility from horses are invaluable outputs from rearing livestock. The innovative herd management practice of castration enabled non-breeding males to be managed safely beyond the typical age of slaughter, thereby greatly improving the productivity of livestock herds. Although osteological methods can identify some morphological differences between castrated and intact males, it is difficult to make clear distinctions as significant overlaps exist between the two groups. However, methylation values are affected by the considerable hormonal changes that occur as a result of castration. For example, castrated male sheep have been shown to have lower biological ages compared to age-matched intact rams, with androgen-sensitive CpG sites displaying marked differences in methylation values (Sugrue et al., 2021). Furthermore, age-at-death has been predicted from reconstructed methylation values in ancient horse specimens, informing on culling practices (Liu et al., 2023). Using bisulfite conversion and ssDNA library preparation methods, we have reconstructed CpG methylation values from historical sheepskin parchment samples. In addition, we have developed machine learning models trained on an existing dataset of modern sheep methylation values in order to infer traits of interest such as castration status and age-at-death. The informative CpG sites have been incorporated into a targeted methylation probe set to enable cost-effective higher coverage sequencing of additional samples. This will enable the characterisation of these traits in ancient sheep across different time periods, geographical locations, and archaeological contexts.

Reconstructing Jomon Kinship through Paleogenomic Analysis

Yoshiki Wakiyama

The University of Tokyo

Understanding kinship among human remains is crucial for reconstructing the social structures of prehistoric hunter-gatherers. The Jomon people, who inhabited the prehistoric Japanese archipelago, are known for their relatively sedentary lifestyle, with many of their remains interred in shell-mounds. While certain clusters of individuals buried together are observed in Jomon burial sites, the relationships among individuals within these clusters remain unclear. In this study, we analyzed the kinship relationships within and among three Jomon sites in Japan. We sequenced the genomes of 18 individuals, 6 from the Ikawazu, 4 from the Shijimizuka, and 8 from the Kosaku sites, analyzing both mitochondrial and nuclear genomes to infer biological relationships. At the Ikawazu site, second- and third-degree kinship relationships were identified. At the Shijimizuka site, all four individuals displayed kinship spanning three generations. At the Kosaku site, kinship relationships spanning three generations were identified in three individuals, two of whom were mother and child buried together. Additionally, a second-degree kinship relationship was observed between individuals in different group burials. At Shijimizuka, kinship was detected among three individuals buried adjacent to one another and one buried separately, suggesting that the burial area may have been organized based on kinship. Conversely, at the Kosaku and Ikawazu sites, burial proximity did not necessarily correspond to kinship relationships. These results provide new insights into the diversity of burial practices and social structures of prehistoric Jomon hunter-gatherers.

Tracing Denisovan ancestry in early modern human genomes

Jiaqi Yang

Max Planck Institute for Evolutionary Anthropology

During their dispersals into Eastern Eurasia, modern humans met and mixed with Denisovans, a sister group of Neanderthals. Analyses of Denisovan ancestry in present-day populations have shown that it is today heterogeneously distributed and comes from genetically distinct Denisovan groups. However, population movements and admixture after the initial dispersals of modern humans in Eastern Eurasia have obscured the origin of this heterogeneity. To address this, we investigate Denisovan ancestry in early modern human genomes.

We first evaluated the performance of commonly used local ancestry inference methods and applied them to detect introgressed Denisovan fragments in early modern human genomes. We found that the earliest modern humans in East Asia carried higher levels of Denisovan ancestry than present-day Eurasians, and some early East Asians already carried ancestry from multiple Denisovan groups before the last glacial maximum. We also provide evidence for an early-diverging East Asian lineage with at least six times less Denisovan ancestry than other ancient and present-day East Asians, suggesting that early modern humans in East Asia experienced distinct histories of admixture with Denisovans. Finally, we trace the Denisovan ancestry of Europeans back to admixture with early East Eurasians.

These findings highlight the utility of inferring archaic ancestry at the haplotype level in early human genomes, providing new insights into the complex migration and admixture history of modern humans in Eurasia.

Recent ancestry specific inference of archaic tracts in humans

Vladimir Shchur

HSE University

Archaic introgression is one of the most important sources of genetic diversity in modern humans. We developed a hidden Markov model based method DAIseg which allows to infer archaic tracts in modern humans under different population scenarios including multiple introgressions and recent admixture events. We use DAIseg to infer Neanderthal tracts in modern Mexicans together with the way of recent admixture. We discriminate between Neanderthal tracts which appear in Mexicans through the European or through the Native American component. Further, we run a local ancestry selection scan to study the recent post-admixture selection on the Neanderthal component with the specific way of admixture (European or Native American). Another novelty of our method is the simultaneous use of an unadmixed outrgroup (similarly to the HMMMix) and the available archaic genomes. Such a combination increases both precision and recall compared to the state-of-the-art methods. This work was supported by the RSF grant 22-71-10056.

Symposium 29: Open symposium session A

Leaf Masquerade Mimicry in Oakleaf Butterflies

Wei Zhang

Peking University

Oakleaf butterflies in the genus Kallima have polymorphic wing phenotypes enabling these insects to masquerade as dead leaves. This iconic example of protective resemblance provides an interesting evolutionary paradigm that can be employed to study biodiversity. We integrated multi-omic data analyses and functional validation to infer the evolutionary history of Kallima species and investigate the genetic basis of their variable leaf wing patterns. We find that Kallima butterflies diversified in the eastern Himalayas and dispersed to East and Southeast Asia, accompanied by genetic introgression. Moreover, we find that the leaf wing polymorphism is controlled by a wing patterning locus cortex, which involves a complex synergy of coding and non-coding regions and has been maintained in Kallima by long-term balancing selection. Our results provide macroevolutionary and microevolutionary insights into a model species originating from a mountain ecosystem.

High-quality vertebrate reference genomes for a new era in biology

Giulio Formenti

The Rockefeller University

On behalf of the Vertebrate Genomes Project

The Vertebrate Genomes Project (VGP) aims at generating a reference genome for all ~70,000 vertebrate species. The VGP is one of the largest multispecies genomics projects worldwide, and is supported by a large, international network of scientists interested in vertebrate biology and genomics. The project is organized in four phases, progressively targeting all vertebrate orders, families, genera, and species, and comprises multiple working groups focussed on different aspects of this endeavor. We have now completed our 80% threshold of Phase I species, generating a representative genome for most vertebrate orders and we are actively studying the genomes. We provide resources for all assembled genomes, including metadata, summary statistics, and annotation of genomic features. These new genomes already allowed the study of hard to sequence and assembly regions, such as promoters, enhancers, and repeats. They also provided greater resolution of gene family evolution, and more accurate identification of structural variants. VGP resources are particularly useful to support scientific studies on specific clades, loci or genome-wide. Resources are made available to the broader scientific community through our public GenomeArk database, and include multiple raw data types for all species, genome assembly and evaluation tools, pipelines, and training material. We also provide advice for similar projects aimed at generating high-quality reference genomes. We invite all scientists to join our efforts to study these high-quality resources and meet the overarching goal of generating reference genomes for all vertebrate species and eventually all of life.

Adaptive stop codon recoding outside of viruses boosts energy production in horses

Gianni Castiglione

Vanderbilt University

Horses are among nature's greatest athletes, yet the ancestral molecular adaptations fueling their energy demands are poorly understood. Within a clinically important pathway regulating redox and metabolic homeostasis (NRF2/KEAP1) we discovered an ancient mutation—conserved in all extant equids—that increases mitochondrial respiration while decreasing tissue-damaging oxidative stress. This mutation is a de novo premature opal stop codon in KEAP1 that is translationally recoded into a cysteine through novel mechanisms, producing an R15C KEAP1 that is more sensitive to electrophiles and reactive oxygen species. This enables increased NRF2 activity, which we demonstrate to enhance mitochondrial ATP production and cellular resistance to oxidative damage. This illustrates how recoding of a de novo stop codon, a strategy thought restricted to viruses, can facilitate adaptation in vertebrates

Extrachromosomal DNA driven oncogene spatial heterogeneity and evolution in glioblastoma

Magnus Haughey

Barts Cancer Institute

Extrachromosomal DNA (ecDNA) oncogene amplification is associated with treatment resistance and shorter survival in cancer. At present, the spatial dynamics of ecDNA, and the impact on tumor evolutionary trajectories, are not well understood. Here, we investigate ecDNA spatial-temporal evolution by integrating computational modeling with samples from 94 treatment-naive human IDH-wildtype glioblastoma patients. We developed a spatial-temporal computational model of ecDNA positive tumors that integrates whole-genome sequencing, multi-region DNA FISH, and nascent RNAscope, to provide unique insight into the spatial dynamics of ecDNA evolution. Random segregation in combination with ecDNA-conferred fitness advantages induce predictable spatial patterns of cell-to-cell ecDNA copy number variation that are highly dependent on ecDNA oncogenic makeup. EGFR-ecDNAs often reach high mean copy number (mean of 50 copies per tumour cell), confer strong fitness advantages and do not co-amplify other oncogenes on the same ecDNA particles. In contrast, PDGFRA-ecDNAs reach lower mean copy number (mean of 15 copies per cell), confer weaker fitness advantages and frequently co-amplify other oncogenes on the same ecDNA. Evolutionary modeling of EGFR-ecDNAs suggest ecDNA accumulation prior to clonal expansion, which we validate biologically in MYC-ecDNA induced mice. EGFR structural variants occur exclusively on ecDNA, arise from and are intermixed with wild-type EGFR-ecDNAs. Modeling suggests wild-type and variant EGFR-ecDNAs often accumulate before clonal expansion, even in patients co-amplifying multiple ecDNA species. In the first systematic study of the spatial dynamics of ecDNA, our results suggest a potential time window in which early ecDNA detection may facilitate more effective intervention.

Escalation of genome defense capacity enables control of expanding meiotic drivers

Peiwei Chen

Cornell University

From RNA interference to chromatin silencing, diverse genome defense pathways silence selfish genetic elements to safeguard genome integrity. Despite their diversity, different defense pathways share a modular organization, where numerous specificity factors identify diverse targets and common effectors silence them. In the PIWI-interacting RNA (piRNA) pathway, target RNAs are first identified by complementary base pairing with piRNAs and then silenced by PIWI-clade nucleases. Such a binary architecture allows the defense systems to be readily adaptable, where new targets can be captured via innovation of specificity factors. Thus, our current understanding of genome defense against lineage-specific selfish genes has been largely limited to specificity factor innovations, while it remains poorly understood whether other types of innovations are required. Here, we describe a new type of innovation, which escalates the genome defense capacity to control a recently expanded selfish gene in Drosophila melanogaster. Through a targeted RNAi screen for repressors of Stellate—a recently evolved meiotic driver—we identified a novel defense factor, Trailblazer. Trailblazer is a transcription factor that promotes the expression of two PIWI-clade nucleases, Aub and AGO3, to match Stellate in abundance. Recent genetic innovation in the DNA-binding domain of Trailblazer enabled it to elevate Aub and AGO3 expression, thereby escalating the silencing capacity of piRNA pathway to tame expanded Stellate and safeguard fertility. As copy-number expansion is a recurrent feature of diverse selfish genes across the tree of life, we envision that augmenting the defense capacity to quantitatively match selfish genes is a repeatedly employed defense strategy in evolution.

Extensive structural variation and longevity-associated adaptations in nearctic Myotis bats

Juan Vazquez

University of California, Berkeley

The genus Myotis is one of the largest clades of bats, and exhibits some of the most extreme variation in lifespans among mammals. To study the evolution of longevity-associated traits, we generated primary cell lines for over 28 species of bats, and assembled near-complete diploid genome assemblies for 8 Myotis species. Using genome-wide screens of positive selection, analyses of structural variation, and functional experiments in primary cell lines, we identify new patterns of adaptation contributing to longevity, cancer resistance, and viral interactions in bats. We find that Myotis bats have some of the most significant variation in cancer risk across mammals and demonstrate a unique DNA damage response in primary cells of the long-lived M. lucifugus. We also find evidence of extensive structural variation both within and between species, including in genes central to stress response and viral immunity. Together, our results demonstrate how genomics and primary cells derived from diverse taxa uncover the molecular bases of extreme adaptations in non-model organisms.

Net bootstrap support: Revising Felsenstein's bootstrap support for Phylogenomics

Sudhir Kumar

Temple University

Felsenstein's bootstrap method remains a cornerstone for assessing the robustness of phylogenetic inferences, with support values increasing rapidly with the addition of genomic loci in concatenated super-alignments. However, in phylogenomics, high bootstrap support values can sometimes be assigned to spurious relationships due to phylogenetic incongruence across the data set. Addressing this challenge often requires application of computationally expensive approaches, such as partitioned analyses and multi-species coalescence methods. Here, we introduce Net Bootstrap Support (NBS), a novel metric that accounts for overconfidence in Felsenstein's bootstrap support for clades inferred from concatenated super-alignments. NBS accounts for phylogenetic discordance across the data set using the phylogenomic site subsampling and upsampling approach. Analyses of empirical and simulated datasets demonstrate that NBS achieved high accuracy in maximum likelihood phylogenetics, matching or exceeding the performance of partitioned and multi-species coalescence methods. NBS offers a computationally efficient solution that combines low memory requirements and rapid calculations for long phylogenomic datasets. Its application streamlines phylogenomic workflows by eliminating the need for the knowledge of predefined or inferred data partitions and associated phylogenetic analyses. As a result, NBS provides a practical and reliable tool for reconstructing organismal relationships with confidence and efficiency.

Denisovan ancestry sheds new light on modern human dispersals into Asia

Stéphane Peyrégne

Max Planck Institute for Evolutionary Anthropology

While it is widely accepted that modern humans arose in Africa, the timing of the initial arrival and survival of modern humans in different regions of the world remain unresolved. Archaeological evidence suggests multiple dispersals of modern humans out of Africa, with some fossil finds supporting early modern human presence in China by 100 thousand years ago (ka) and in Southeast Asia by 70ka. However, genomic data suggests that people living today descend primarily from a single rapid dispersal after 50ka, as all non-Africans carry Neandertal ancestry from a shared admixture event. Genetic contributions from earlier migrations, if any, were minimal (<1%), but uncertainties persist regarding the relationships and migration routes of ancestral populations across Eurasia and Oceania. Investigating interactions between early modern humans and Denisovans, archaic humans who lived in Asia before modern humans, can shed light on these dispersal(s). Previous work has shown that the ancestors of present-day Asians and Oceanians met and mixed with multiple, genetically distinct Denisovan populations. Using a second Denisovan genome we reassess the distribution of different Denisovan ancestry components in present-day people. The geographic distribution of these components today offers clues about where Denisovans lived and the routes taken by modern humans as they colonized Asia. Furthermore, comparisons of introgressed Denisovan DNA segments across populations and estimates of their ages in present-day genomes shed light on the number and timing of these migrations. These findings offer a new perspective on the history of early modern human dispersals.

Pathogen Evolution and Social Dynamics in Colonial Mexico through Paleogenomics

Laura Carrillo-Olivas

National Autonomous University of Mexico (UNAM)

The European colonization of the Americas caused profound demographic changes and epidemic outbreaks, yet the diversity of pathogens affecting marginalized colonial populations in Mexico remains poorly understood at the molecular level. To address this, we analyzed paleogenomic data from 96 individuals and seven dental calculus samples excavated from the Hospital Real San José de los Naturales (HSJN) and the Templo de la Inmaculada Concepción (La Conchita) in Mexico City, dated to the colonial period. Mitochondrial DNA (mtDNA) and nuclear ancestry analyses revealed predominantly Indigenous genetic ancestry, with one individual from HSJN exhibiting both matrilineal and autosomal ancestry consistent with a recent origin in Africa. These findings reflect the hospital's role in serving marginalized Indigenous and African populations as well as the diverse ancestries of people in colonial Mexico. Molecular sex analysis revealed balanced XX and XY ratios, while one individual from La Conchita exhibited a rare XYY karyotype, offering a unique perspective on sex chromosomal variation in ancient populations. We reconstructed three S. enterica Paratyphi C genomes and found that two were basal to ancient strains from Asian lineages, while the other one fell within European variation, suggesting multiple introductions. Additionally, we recovered the first ancient genome of Streptococcus pneumoniae, which clustered within a predominantly African lineage, highlighting a likely transatlantic introduction during colonization. These findings integrate archaeogenomic, archaeological, and historical evidence to explore pathogen evolution and social organization in colonial Mexico. This study shows the interdisciplinary potential of paleogenomics to illuminate the biological and cultural complexities of the past.

Pan-Evo: The evolution of information and biology's part in this

William (Bill) Sherwin

EERC, BEES, UNSW-Sydney

This talk merges molecules, biology and evolution, in a novel way. Will biology, including humans, benefit or suffer from new developments in information such as artificial intelligence (AI)? I propose that biological and non-biological information might be components of a unified process, 'Panevolution' or 'Pan-Evo', based on four basic operations—innovation, transmission, adaptation, and movement. Pan-Evo contains many types of variable objects, from molecules to ecosystems. Biological innovation includes mutations and behavioural changes; non-biological innovation includes naturally occurring physical innovations and innovation in software. Replication is commonplace in and outside biology, including autocatalytic chemicals and autonomous software replication. Adaptation includes biological selection, autocatalytic chemicals, and 'evolutionary programming', which is used in AI. The extension of biological speciation to non-biological information creates a concept called 'Panspeciation'. Panevolution might benefit or harm biology, but harm might be minimal if AI and humans behave intelligently because humans and the machines in which AI resides might split into vastly different environments that suit them. Such Panspeciation would be the first speciation event involving humans for thousands of years. This event will not be particularly hostile to humans if humans learn to evaluate information and co-operate better to minimize both human stupidity and artificial simulated stupidity (ASS—a failure of AI).

NB: This presentation's scope is wide, needing 30+ minutes.

- Full: Sherwin, WB. 2024. Biology 13:507. doi.org/10.3390/biology13070507

- Brief: Sherwin WB 2023. 'Singularity or Speciation?' A comment on Lazar S & Nelson A. 2023. 'AI safety on whose terms?' Science 381(6654):138 DOI: 10.1126/science.adi898

Symposium 30: Open symposium session B

Genetic and epigenetic insights into mandibular tori in a medieval mongolian population

Odongoo Ravdandorj

Kanazawa University

Mandibular tori are bony growths that appear on the inner surface of the lower jaw. While they are frequently found in skeletal remains, the reasons behind their formation are still not fully understood. Both genetic and epigenetic factors may play a role, but further research is needed. This study investigates these factors in a medieval Mongolian population from the Oortsog Mountain site, offering insights into skeletal variation and adaptation.

Ancient DNA was extracted from 27 excavated individuals, 8 of whom had mandibular tori. Well-preserved cochlear and mandibular bone samples provided promising levels of endogenous DNA, allowing for initial genetic and epigenetic analysis.

This study focuses on (1) identifying genetic variations that may be linked to mandibular tori, (2) examining epigenetic patterns, particularly DNA methylation, that could influence bone growth, and (3) comparing individuals with and without mandibular tori to uncover biological factors contributing to their formation. By analyzing genetic and epigenetic markers in this medieval Mongolian group, this research seeks to better understand how mandibular tori develop.

The findings may contribute to broader discussions on skeletal variation, bone remodeling, and human adaptation, but further studies will be needed to confirm any genetic or epigenetic links.

Native American druggable genome diversity and its impact on drug response in the Brazilian population.

Tábita Hünemeier

Institute of Evolutionary Biology (CSIC/UPF, Spain) and USP (Brazil)

The advancement of genetic sequencing technologies, along with cost reduction, has enabled precision medicine, particularly in drug use. However, its progress is not uniform. This genomic underrepresentation is especially evident among Native Americans. The objectives of this study are: a) Identify genetic variants associated with Brazilian Indigenous populations compared to controls (Europeans, Africans, and Asians) in druggable genes. b) Analyze which drugs interact with genes containing alleles associated with Indigenous populations. c) Understand, from an evolutionary perspective, why these alleles are at high frequency in Indigenous groups.d) Quantify the Indigenous contribution to the genetic background of the Brazilian population.

The Indigenous samples consist of whole-genome data from 50 Brazilian Indigenous individuals, plus 20 from the HGDP/Simon project, also part of the control group. The Brazilian population samples include 2,721 individuals from the “DNA do Brasil” project. Comparing Indigenous and control groups, we identified 27 alleles at high frequency in Indigenous individuals, belonging to 26 genes. These genes interact with 409 drugs or small molecules, of which 209 have their use approved.

Using Fine-Mapping of Adaptive Variation (FineMAV) and Population Branch Statistic (PBS), 7 alleles at high frequency in Indigenous populations showed signals of positive selection in both methods. Of these 27 alleles, 1 was also frequent in the Brazilian population, and 12 were frequent in northern Brazil, the region with the highest Indigenous ancestry.

Species complexes of malaria mosquitoes in North America

Yuttapong Thawornwattana

University College London

cryptic sibling species known as species complexes. Members of many Anopheles species complexes are dominant vectors of human malaria, yet the nature of species in many species complexes remain poorly understood, especially those in North America. Here, we use a combination of phylogenetic and population genomic analyses of new de novo genome assemblies and whole genome resequencing data to establish the identity and evolutionary relationship of Anopheles species in North America. Population genomic analysis confirms putative species and uncovers several potential new cryptic species. We find substantial gene flow among sibling species throughout the divergence history despite a lack of prior evidence for gene flow, suggesting that gene flow is a common feature of species divergence rather than an exception. Our findings has implications for vector control and surveillance as well as laying the groundwork for studying the evolution of key traits such as vectorial capacity and vector responses to anthropogenic changes such as insecticide use and gene drives.

Joint inference of continental and archaic ancestry: A deep learning approach

Jazeps Medina Tretmanis

Brown University

Local Ancestry Inference (LAI) has been useful to study the evolution of recently admixed populations in the Americas. Similarly, detecting archaic introgressed tracts in Eurasian populations has provided insights into our evolutionary history and interactions with archaic humans. Currently, the landscape of archaic introgression in recently admixed populations remains understudied, leaving questions about how recent admixture has altered patterns of archaic ancestry. This is partly due to a lack of methods that simultaneously detect LAI and archaic ancestry. We present the first deep learning method capable of jointly inferring intercontinental human ancestry and archaic introgressed regions.

A major challenge for deep learning models is the need for extensive training data. While high-quality unadmixed datasets exist for many modern populations, reference panels for admixed groups are often limited. Similarly, the available archaic genomes fail to capture their full diversity. We present a reference-free method for archaic introgression detection. With a multitask architecture, our approach displays improved performance for both modern and archaic ancestry inference tasks.

The model can be trained on synthetic data alone or a mix of real and synthetic data. Using purely synthetic training data, it achieves ~85% precision and recall of archaic introgression detection on real datasets, comparable to state-of-the-art archaic detection methods without requiring real data for training. We also show state-of-the-art performance for intercontinental ancestry inference while training on completely synthetic datasets. Additionally, the method can be successfully trained on one real dataset and applied to another, even if SNP positions between datasets are different.

Constitutively active GCGR: A key factor for birds in maintaining high blood glucose and lean body weight

Cheng Deng

West China Hospital, Sichuan University

As the body's primary source of energy, the maintenance of blood glucose and lipid are indispensable for overall health and metabolic homeostasis. It is predominantly regulated by the glucagon receptor family which is highly conserved in vertebrates. Compared to other vertebrates, avian blood glucose levels are relatively high and birds exhibit modest weight gain and limited fat accumulation due to rapid fat utilization as flight fuel, yet the regulatory mechanisms have remained obscure for more than a century. We show that high hepatic expression of the avian glucagon receptor (GCGR) in association with constitutively active Gs signaling was dependent upon the interaction of different domains. In vivoexperiments focusing on the regulation of constitutively active GCGR expression in hepatocytes led to correspondingly high blood glucose, rapid hepatic lipid utilization and high metabolic rates via downstream signaling pathway activation in fish, reptiles, birds, and mammals. Furthermore, cross-species single-nucleus RNA-sequencing analysis revealed high GCGR expression levels in avian adipocytes, while minimal in the white adipose tissue (WAT) of other vertebrates. Avian GCGR employs a unique high glucose, rapid fat utilization and energy metabolism for retaining a light body in mice. Transgenic avian GCGR or other constitutive active GCGR variants (human GCGRH339R) expressed in mouse liver or WAT effectively facilitated rapid glucose metabolism or fat utilization, lowered and maintained body weight. We suggest that GCGR plays this crucial role for maintaining high blood glucose and lean body weight in birds, which might be employed in therapeutic approaches to treat obesity in humans.

Prior probability of clade splits under serial-sampled coalescent

Zijing Yang

University of Auckland

Conditional Clade Distributions (CCDs) provide an effective way to represent posterior tree distributions in Bayesian phylogenetic analysis. The prior distribution of clade splits is an important element of CCDs, which is useful for combining multiple datasets in large-scale phylogenetic studies. While the prior probability of clade splits can be calculated easily for contemporaneously sampled data, the case of serial-sampled data is more complex. Here we introduce a dynamic programming algorithm to calculate the prior probability of clade splits under the coalescent process with serial-sampled data. To overcome computational challenges, we develop a heuristic method that improves running time while maintaining high accuracy. Additionally, we propose a simulation-based approach to integrate over various population sizes. Our results highlight the influence of serial sampling intervals on clade split probabilities, enhancing our understanding of prior clade distributions. This work provides a foundation for further optimization and scaling in complex phylogenetic analyses.

A hotspot-aware map of structural variation in 252 haplotype-resolved assemblies

Ziyi Yang

Fudan University

The genetic toolkit mediating metabolic division of labour across ant social networks

Yuqi Wang

University of Cambridge

Metabolic division of labor is central to cooperative biological systems, from multicellular organisms to mutualisms. Ant colonies exemplify this, functioning as superorganisms where individuals perform specialized roles, facilitated by social transfers of materials like regurgitates, secretions, and excretions. These fluids often contain colony-tissue-specific proteins, hinting at distinct metabolic roles for different colony tissues. However, the genetic regulators and "currencies" driving this colony economy remain poorly understood.

We studied five ant species from four subfamilies (Formicinae, Myrmicinae, Ponerinae, Amblyoponinae) with diverse social transfer systems, analyzing tissue-specific gene expression and protein abundance in individuals and transferred materials. Across species, transferred materials shared orthologous groups of storage proteins, juvenile hormone esterases, and digestion enzymes, suggesting a conserved ancestral toolkit despite varied transfer networks. Proteins in social fluids exhibited more secretion signals than tissue proteins, and larval excretion fluid emerged as a key, previously undescribed social transfer, likely conveying metabolic information to workers.

Transfer patterns differed between major ant lineages: Formicoid ants, like honeybees, transfer costly major royal jelly-like proteins to larvae, while Poneroid ants derive larval storage proteins from larvae for adult use. These findings reveal conserved mechanisms in ant social organization while highlighting species-specific metabolic adaptations shaped by social transfer networks.

Predominant subfunctionalization of ohnologs in teleost and salmonid genomes following whole genome duplication

Yuanshuo Li

Trinity College Dublin

Whole genome duplication (WGD) events represent pivotal drivers of eukaryotic genome evolution, generating abundant genetic material that facilitates the functional divergence of duplicated genes. While most duplicates are rapidly lost due to redundancy, retained "ohnologs" (both copies preserved following WGD) exhibit distinct evolutionary trajectories in their expression patterns. However, the predominant mode of expression divergence in ohnologs remains contentious: studies in spotted gar support subfunctionalization through tissue-specific regulatory partitioning (both copies alter their expression patterns such that their combined expression remains similar to that of the pre-WGD single copy), whereas salmonid research suggests neofunctionalization via asymmetric regulatory evolution (one copy retains its ancestral expression pattern while the other diverges).

To resolve this debate, we systematically investigated the regulatory fates of ohnologs following the teleost-specific TGD (~320 MYA) and salmonid-specific Ss4R (~88 MYA) WGD events. Our study integrates two key innovations: (1) a probabilistic macro-synteny model to reconstruct ohnologs across 15 fish species, creating the most comprehensive database of TGD- and Ss4R-derived duplicates to date; and (2) a meta-expression atlas combining RNA-seq data from 10 tissues spanning these species. Through tissue-specific expression divergence simulations based on the Ornstein–Uhlenbeck process, and the development of neural network classifiers, we found that subfunctionalization is the dominant mode (average 14.70%), significantly exceeding neofunctionalization (average 0.88%).

Mechanistically, this pattern arises from dosage constraints and the coordinated duplication of regulatory elements, which collectively preserve the ancestral expression breadth through complementary tissue specialization. Our findings reconcile previous conflicting observations by establishing a unified framework for understanding WGD-driven regulatory evolution.

GP4PG: A metaheuristic approach to demographic modelling in population genetics

Jose Miguel Serradell Noguera

IBE-UPF-CSIC

Traditional demographic modelling describes each tested model as a combination of topology and demographic events. However, the ascertainment of such models is often subject to bias, primarily due to the subjective selection of demographic events by researchers, particularly when population substructure is involved. To address these limitations, we developed the Genetic Programming for Population Genetics (GP4PG) methodology, an innovative approach to modelling complex demographic scenarios by applying genetic programming. GP4PG automates the exploration of parameter-free demographic models by coding demographic relationships as graphical structures, enabling optimization through mutation and selection. GP4PG minimizes model bias by only relying on the topology of demographic models rather than predefined parameter assumptions.

A key feature of GP4PG is its organization of populations into hierarchical substructures known as "topodemes," nested within "ecodemes." This organization accounts for population heterogeneity and spatial distribution, providing a more accurate framework for modelling demographic dynamics. Migration between populations is modelled under an isolation-by-distance principle, with migration rates decaying exponentially over time. GP4PG's iterative optimization process evaluates potential demographic scenarios by comparing simulated data to observed genetic data, using Site Frequency Spectrum (SFS), refining solutions until an optimal fit is achieved.

We applied GP4PG to North African human population genomics revealing a back-to-Africa expansion with differential origins for Arab and Amazigh groups, emphasizing continuous migration decay rather than discrete admixture pulses as primary driver of genetic diversity. The demographic model suggested by GP4PG exhibited a closer alignment with the observed data compared to Approximate Bayesian Computation with Deep Learning (ABC-DL) approach.

Phylogenomic insights into trait evolution and biogeography in the liverwort family Lepidoziaceae

Antonio Rayos

University of Sydney

Liverworts (Marchantiophyta), one of the three groups of non-vascular plants, include the Lepidoziaceae as its third-largest family. Having originated in the early Jurassic, this family has spread to all continents. The long evolutionary history of this family is accompanied by notable phenomena such as morphological convergence, but its phylogeny has been estimated only through analyses of multilocus data sets. We present a phylogenomic analysis of the family based on 400 nuclear exons from 141 species, finding broad agreement between trees inferred using concatenation and coalescence approaches. We identified a case of parallel morphological convergence in Meinungeria from New Caledonia. Moreover, we found different dispersal routes from Australasia to the Neotropics. Our study demonstrates the power of using a phylogenomic approach to resolve the evolutionary history of a megadiverse family of plants.

Coupling of habitat-preference barriers leads to reproductive isolation in sympatric speciation

Karyn How

Evolution Research Inc.

Habitat preference is a widely recognized mechanism of reproductive isolation (RI), yet its role in initiating premating barriers and coupling with other barrier mechanisms to establish robust RI in sympatric speciation remains unclear. Here, we developed mathematical models and computer applications to investigate one- and two-allele models of habitat-preference barriers in sympatric populations under disruptive ecological selection. We examined two spatial arrangements: an open-space model, inspired by cichlid fishes with small niche habitats relative to lake size, and a no-open-space model, inspired by hawthorn and apple maggot flies that move directly between trees. Coupling dynamics between habitat-preference barriers and mating-bias barriers were explored. Our findings confirm that habitat preference can establish initial premating barriers and couple with additional barriers to enhance overall RI. Because habitat preference is an adaptive barrier mechanism, its invasion and coupling are driven by maladaptive hybrid loss, and it is readily reversed when disruptive ecological selection weakens. As long as hybrid loss persists, both one- and two-allele models can evolve to mitigate it, though the one-allele model can evolve more readily because it is not affected by gene flow. By reducing gene flow, habitat preference facilitates the emergence of mating-bias barriers. Open-space systems limit inter-niche encounters and generate faster and stronger RI than no-open-space systems. These results reinforce the important role of habitat-preference barriers in sympatric speciation and provide insights into a wide range of premating isolating mechanisms—temporal, behavioral, and mechanical—that function similarly by reducing mating encounters.

Fitness Costs and Trade-offs of Pseudomonas aeruginosa in Spatially Complex Environments

Emma Helman

McGill University

The continued misuse and overuse of antibiotics have accelerated the rise of antibiotic resistance,presenting a major threat to our healthcare system. One of the leading pathogens of concern is the opportunistic bacterium Pseudomonas aeruginosa, owing to its high intrinsic resistance andability to infect immunocompromised patients. Although widely studied in well-mixed bacterial populations, less is known about how spatial network structures influence the evolution and fitness costs of resistance in pathogens such as P. aeruginosa. Star-shaped metapopulations, characterized by a central hub with peripheral nodes, have been shown to accelerate the spread of beneficial mutations due to a reduced probability that rare beneficial mutations are randomly lost from genetic drift.3 Here, we compare the underlying mechanisms of ciprofloxacin resistance in well-mixed versus star-shaped metapopulations of P.aeruginosa. Carrying capacity fitness assays were conducted to compare costs of fitness and tradeoffs. Whole-genome, single-isolate sequencing will be used to identify known antibiotic resistance mutations and putative compensatory mutations linked to fitness costs. Our preliminary results indicate that isolates from star-shaped metapopulations exhibit higher costs of fitness and tradeoffs, possibly due to large-effect costly mutations. This study elucidates the role of metapopulation structure in evolutionary dynamics of resistance, specifically how network topology contributes to the spread and maintenance of antibiotic resistance. These findings have broader implications for informing strategies to manage antibiotic resistance in healthcare settings, such as hospital networks, where patient transport patterns may closely resemble star-shaped bacterial metapopulations.

Symposium 31: Open symposium session C

Genetic architecture of pheromone variation and prezygotic isolation in two parasitoid wasp species

Weizhao Sun

Institute for Evolution and Biodiversity, Muenster University

One of the most important questions in biology is how new species originate. Variations in sexual signals contribute greatly to reproductive isolation by enabling species-specific recognition. Cuticular hydrocarbons (CHCs), major recognition molecules on insects' cuticles, are shown to encode species-specific sexual signals in many taxes. However, there is still limited knowledge on which specific compounds in complex CHC profiles mediate species-specific mate choice. Similarly, the genetic basis governing CHC variations involved in reproductive isolation remains largely elusive. The present study characterizes the chemical and genetic basis of species-specific CHC variation and its consequences for prezygotic reproductive isolation between two closely related parasitoid wasp species. Exploiting their haplo-diploid sex determination and cross-fertility under laboratory conditions, we generated female recombinant inbred lines (RIL), constituting both genetic and phenotypic hybrids of the parental species. We conducted quantitative trait loci (QTL) analyses of all individually identified CHC compounds separately for the diploid female recombinant lines as well as their haploid male fathers. Correlating CHC quantities with mate preferences, we identified three key compounds significantly correlated with male mating behavior frequencies. Further zooming into the QTL for the species-specific variation of our three key pheromonal compounds, we were able to localize several CHC biosynthesis candidate genes. Further knockdown of one of the candidate genes largely impacted the newly identified key compounds as well as structurally similar CHCs. Thus, through identifying the species-specific key components out of complex pheromonal blends mediating mate preference, we determined the molecular and chemical basis of a prezygotic isolation mechanism.

Closing the 'Jurassic Gap': Integrated analysis of fossil occurrences and molecular divergence time estimation infers a latest Jurassic origin of flowering plants

Ruolin Wu

University of Bristol

Molecular timescales are based on the calibration of molecular evolution to geological time using fossil constraints, but conventional calibration strategies use limited and often subjectively interpreted fossil data. Here, we used the Bayesian Brownian Bridge model to objectively derive calibration densities based on extensive fossil occurrence data independent of phylogenetic relationships. This approach can integrate the uncertainty on extant and historical diversity to estimate clade origination and extinction. We transformed the estimated ages based on >25,000 fossil occurrences into calibration densities, which we used to constrain 110 node ages in a 644-species angiosperm phylogenetic tree inferred from a molecular alignment of 83 genes. The results refute a post-Jurassic origin of angiosperms, suggesting a cryptic early history. Our study demonstrates the utility of a mechanistic approach to establish node age constraints in molecular clock-dating analyses, resulting in a more objective method to integrate molecular and palaeontological data when inferring evolutionary timescales.

Bloodhound: A deep learning approach to bacterial taxonomic classification

Wytamma Wirth

The Peter Doherty Institute for Infection And Immunity

Deep learning has revolutionized fields such as protein folding through tools like Alphafold. Here, we introduce Bloodhound—a novel tool that leverages state-of-the-art protein language models and a hierarchical neural network architecture to classify bacterial genomes with unprecedented accuracy. Trained on hundreds of thousands of bacterial genomes, Bloodhound achieves state-of-the-art performance in bacterial taxonomic classification, significantly outperforming previous deep learning approaches. Our work presents a comprehensive overview of the design, training, and validation processes involved in developing Bloodhound, along with key insights gained during its development. Additionally, we discuss the broader implications of our findings for future research in microbial genomics. Bloodhound is available as an open-source, easily installable tool at https://github.com/bloodhound-devs/bloodhound.

TOGA2 improves speed and accuracy of comparative gene annotation

Yury Malovichko

Senckenberg Research Institute and Natural History Museum

Gene annotation remains a major bottleneck for modern genomics, evident by the growing gap between the numbers of assembled and annotated genomes. For comparative and evolutionary genomic studies, gene annotation as well as inferring evolutionary relationships for homologous sequences across different genomes is a fundamental requirement. TOGA (Tool to infer Orthologs from Genome Alignments) is the first method that jointly addresses both problems by using whole genome alignments to infer and simultaneously annotate orthologous genes.

Here we present TOGA2, which undertakes the few known limitations of the original TOGA and further expands its functionality. By implementing a new exon-wise instead of locus-wise annotation procedure, TOGA2 improves exon annotation accuracy, and massively reduces memory requirements 300 times on average. Furthermore, we show how incorporating deep-learning based splice site predictions enables TOGA2 to detect evolutionary gene structure changes, thus further improving annotation correctness.

A gene tree-based reconciliation step further improves orthology resolution and detects additional single copy orthologs. Finally, TOGA2 can also predict untranslated regions. We demonstrate the capability of TOGA2 by generating the largest comparative genomic datasets for mammals and birds, comprising whole genome alignments, gene annotations and classifications, orthologs and codon alignments for more than seven hundred species each. Finally, we demonstrate how the application of TOGA2 to these datasets reveals patterns of gene loss and phenotype evolution across mammalian and bird biodiversity.

The role of gene family size in cancer resistance among wild mammals

Louise Maille

LIENSs, University of La Rochelle

Cancer is a disease associated with mutations in genes that regulate essential cellular processes such as growth, division, and functioning. Cancer impacts most multicellular organisms, including a wide variety of mammals. Notably, the prevalence of cancer varies significantly across mammalian species, suggesting underlying evolutionary adaptations aiming to combat cancer in resistant taxa. Gene duplications are a key mechanism driving the evolution of new traits and could explain the evolution of resistance in some taxa. For instance, the African elephant has 19 copies of the tumor suppressor gene TP53, two of which are functional, providing enhanced protection against cancer compared to humans harbouring a single copy of TP53. Moreover, studies have revealed a correlation between the duplication of cancer-related genes and longevity in mammals and fish, suggesting that gene duplications may play significant roles in enhancing cancer resistance, especially in long-lived, large species where the problem of carcinogenesis is more pronounced.

To test this hypothesis, we conducted a comprehensive analysis of gene family sizes across 68 wild mammalian species with known cancer prevalence. Our analyses were not restricted to cancer-related genes, providing a broader view and thorough understanding of the genetic factors that may influence cancer resistance across mammals. Our study revealed a handful of gene families whose expansion predicts cancer prevalence across species, including some involved in cell cycle regulation, some of which were not previously associated with cancer resistance. These findings highlight the role of gene family expansions in shaping resistance, providing insights into cancer resistance mechanisms in wildlife.

A large set of nuclear genes brings gastrotrichs into the phylogenomic era

Oscar Wallnoefer

University of Bologna

The phylum Gastrotricha remains largely unexplored from a molecular phylogenetic perspective. This small taxon, comprised of approximately 890 species of tiny and microscopic worms (ranging from 80 μm to 3.5 mm), inhabits both freshwater and marine environments and is considered one of the most common invertebrate groups within meiofaunal communities. To date, extensive knowledge has been accumulated on morphological characters, and in-group relationships are often inferred from them. Recent efforts have incorporated molecular markers, primarily mitochondrial genes; however, the limited number of loci analyzed has led to inconsistencies among taxa previously grouped based on morphology. Moreover, the phylogenetic placement of Gastrotricha within the metazoan tree remains unresolved and their precise position is still debated. The incomplete representation of taxa in most spiralian phylogenetic studies further complicates the resolution of their exact placement.

The results we present are based on more than 30 species whose nuclear genomes have been sequenced and assembled for the first time. In order to maximize the number of species included, we prioritized species diversity over coverage depth. Despite this, our carefully curated pipeline allowed us to extract hundreds of phylogenomic markers.

We extracted a significant number of nuclear genes (over 250, with an 80% complete matrix), and phylogenetic analyses have been carried out. In addition to the in-group analysis, we have also tested the position of Gastrotricha within Spiralia. This study represents a pivotal step toward establishing a reference phylogeny for this phylum.

Exploring migration patterns and metastatic evolution in colorectal cancer: A dense sampling strategy

Qihang Chen

Guangzhou Medical University

Tumor metastasis, the leading cause of cancer-related mortality, remains a major therapeutic challenge due to its often undetectable early stages.

This study explores the genomic and spatial heterogeneity of colorectal cancer (CRC) and its liver metastases using a high-resolution micro-sampling approach, allowing for an in-depth analysis of intra- and inter-tumoral migration patterns. We sampled 698 tissue samples from two patients, with each sample containing approximately 3,000 cells and 3D spatial coordinates.

We utilized whole genome sequencing and targeted sequencing to track the evolution of tumor clones across multiple tumor regions. Clonal deconvolution, phylogenetic analysis, and computational modeling were employed to reconstruct migration pathways between primary and metastatic sites.

Our findings reveal considerable genetic heterogeneity both within and between tumors. Specifically, we observed polyclonal dissemination in both cases and subsequent spreading among liver metastases. Several clonal transfers occurred between metastatic tumors, highlighting the complex inter-tumoral interactions.

Spatial analysis identified distinct modes of migration in the two cases: Case 1 exhibited limited intra-tumor clonal migration, with genetic distances strongly correlating to physical distances, while Case 2 showed no significant correlation between genetic and physical distances.

Phylogenetic analysis further uncovered early dissemination in both cases, occurring several years before the primary tumor was clinically detected. Integration of computational simulations of tumor cell migration dynamics suggests that early migration events likely precede clinical detection, complicating metastatic spread.

These findings emphasize the role of spatially distinct clonal populations in driving metastatic progression and offer new insights into the temporal and spatial mechanisms underlying tumor development.

The tempo of molecular and morphological evolution across the Tree of Life

Estefany Karen López Estrada

Instituto de Biología, UNAM

To explain the evolution of Earth's vast genetic and morphological diversity an array of models have been described. However, two contrasting models, phyletic gradualism and punctuated equilibrium are usually under the spothlight. Under a phyletic gradualism framework, species are steadily transformed into new forms over time. In contrast, under punctuated equilibrium, evolutionary change tends to occur in bursts at speciation events. Previous studies of morphological and molecular data have found varying levels of support for the two evolutionary models. We examined these models using comprehensive morphological and molecular data sets from 40 clades across the Tree of Life. Testing for associations between species richness and the amount of evolutionary change in sister clades, we find little evidence to support the punctuated equilibrium model. However, we found high levels of among-lineage rate variation in molecular evolution and particularly morphological evolution. Our comparison of coding and non-coding DNA regions revealed contrasting patterns of among-lineage rate variation, without clear trends across taxa. Our study suggests that heterotachy could be a dominant feature in macroevolution and that molecular and morphological evolution cannot simply be described by either a gradual or punctuated model.

Adaptive Splicing Responses to Temperature Deviations in Soybean Flowering Genes

Tanya Pelayo

California State University, Northridge

Understanding the distribution of fitness effects (DFE) of new mutations is crucial for predicting evolutionary trajectories and adaptation rates across species. While DFE studies have largely focused on microorganisms and animals, comprehensive comparisons across plant species remain limited. This project seeks to analyze genome-wide data from phylogenetically diverse plant species to characterize and compare their DFEs. Here, we present preliminary data on three plant species, tomato (Solanum lycopersicum), black cottonwood (Populus trichocarpa), and maize (Zea mays). We observe a smaller nonsynonymous-to-synonymous mutation ratio (ranging from 1.32 to 1.37) than expected in neutrality (~2.5), evidencing the action of purifying selection on amino acid-changing mutations. To infer the full DFE for these species, we first used dadi to fit a demographic model to the synonymous allele frequency spectrum (AFS) and then fitdadi to fit a gamma-distributed DFE to the nonsynonymous AFS. We find the DFE to be statistically similar across the studied species, with maize having slightly lower mean fitness effects. We discuss our findings in the context of differences in life history traits, particularly breeding systems and effective population sizes.

DNA methylation reshapes cis- and trans-regulatory landscapes during natural hybridization: A case study of F1 hybrid walnut

Weihao Wang

Beijing Normal University

Transcriptional regulation changes frequently underlie trait evolution in plants and can lead to reproductive isolation. These changes primarily arise from cis-regulatory DNA sequence variations and evolution of trans-acting transcription factors, resulting in altered gene expression levels and timing. However, this approach often overlooks the potential role of epigenetic modifications, which can change during hybridization and affect gene expression. We used the natural hybrid F1 tree Juglans hopeiensis and its parental species as a model system to compare methylation and gene expression levels across multiple tissues. Our findings reveal that hybridization can trigger extensive genome-wide methylation changes. Notably, we identified specific genes for which changes in DNA methylation patterns strongly correlated with altered expression levels. This finding suggests that, under certain conditions, epigenetic modifications can selectively override or modulate expected cis-regulatory control in specific genomic regions, offering new insights into how both genetic and epigenetic changes drive trait evolution during plant speciation.

Ancient proteins preserve genetic sequence data >20 million years old, and offer robust, but limited, phylogenetic information

Ryan Sinclair Paterson

University of Copenhagen

Ancient proteins can persist across geological timescales, particularly in biomineralized tissues such as enamel and eggshell, while retaining phylogenetic information. However, ancient peptides from these tissues tend to derive from the most conserved regions of these proteins, limiting their utility for resolving evolutionary relationships of extinct taxa. Despite this constraint, ancient proteins provide the only direct source of genetic data in deep-time contexts where DNA is no longer recoverable, offering a crucial complement or alternative to morphology-based phylogenetics.

Here, we apply protocols for extracting ancient protein sequences from dental enamel to evaluate protein survival across time and space. We sample broadly across geographic localities and geological epochs, focusing on sites from high latitudes and altitudes that favor biomolecular preservation. Specifically, we aim to push back the oldest verifiable evidence for ancient protein survival by sampling dental enamel from high-latitude sites in the Miocene, Oligocene, and Eocene of Antarctica and Canada's High Arctic.

Our results indicate that in tropical and subtropical regions, ancient peptides from biomineralized sources do not appear to preserve beyond the Late Miocene, and phylogenetically informative peptide sequences do not persist beyond the Pliocene. However, we successfully extracted informative peptide sequences from dental enamel from the Early Miocene (>20Ma) of Canada's High Arctic. Beyond this timeframe, we detect only limited evidence of protein preservation. Although recovered ancient protein sequences offer limited resolution for reconstructing phylogenetic hypotheses, they provide robust divergence time estimates, enabling direct calibration of key nodes deep within Mammalia.

Symphony of conflict and cooperation as revealed by dynamic evolution of gene expression in sexual reproduction

Aimei Dai

Sun Yat-sen University

The evolutionary trajectory of sexual reproduction has been shaped by intricate interactions between intra- and intersexual conflict and cooperation, yet the transcriptional regulatory mechanisms underlying these dynamics remain poorly resolved. Here, we integrate comparative transcriptomics and chromatin accessibility analyses of male accessory glands, testes, and ovaries across closely related Drosophila species to elucidate organ-specific evolutionary dynamics in sexual reproduction. Cross-organ correlation analyses of gene expression divergence reveal that approximately half of genes exhibit concordant evolutionary trajectories across reproductive tissues, while only 9% display divergent evolution linked to accessory gland proteins (Acps), which emerge as central mediators of intersexual conflict. Notably, highly expressed genes show opposing expression trends: decreased in testes but elevated in accessory glands and ovaries of D. melanogaster versus D. pseudoobscura. Despite similar chromatin states between testes and accessory glands, open chromatin regions correlate with increased expression divergence in accessory glands and ovaries, while being depleted in testes. These findings suggest that transcriptional networks maintain global coordination across reproductive organs despite episodic conflicts driven by sexual selection. By linking chromatin dynamics to gene expression divergence, our study positions chromatin remodeling as a key driver of organ-specific transcriptional evolution, offering a molecular framework to reconcile cooperation and conflict in reproductive adaptation.

Reference-quality genomes reveal the family-level phylogeny of bats

Yan Liang

Wellcome Sanger Institute

Bats (Chiroptera) represent the second-largest order of mammals, comprising over 1,400 species distributed across six continents, excluding Antarctica. They exhibit remarkable evolutionary diversity and are classified into two suborders: Yinpterochiroptera (which includes megabats and several microbat families) and Yangochiroptera (comprising all remaining microbat families). Bats possess extraordinary adaptations, including powered flight, echolocation, extreme longevity, and unique immune responses. However, their deep evolutionary history remains contentious. Given this complexity, high-quality genomes are essential for resolving bat phylogeny and understanding the genetic basis of their adaptations.

As part of the Bat1K initiative, which aims to sequence the genomes of all living bat species, we employed state-of-the-art sequencing technologies and annotation pipelines to generate a high-resolution phylogeny of all 21 bat families using 103 newly assembled reference-quality genomes. Our analyses integrate whole-genome alignments, protein-coding genes, non-coding regions, and mitochondrial genomes to reconstruct robust phylogenetic trees. We resolve long-standing phylogenetic conflicts and provide strong support for the Yangochiroptera/Yinpterochiroptera classification over the traditional Megachiroptera/Microchiroptera division. By identifying phylogenetic discordance across genomic regions and evolutionary timescales, we offer novel insights into bats evolution.

Beyond phylogeny, these genomic resources provide a foundation for investigating key evolutionary innovations in bats, including immune system evolution, viral tolerance, sensory adaptations, and developmental biology. Future work will integrate transcriptomic, epigenomic, and functional validation approaches to further elucidate the genetic basis of bats' extraordinary traits.

Neutral processes shape the genetic diversity of a marine microbial community

Julio Cesar Ignacio Espinoza

Keck Graduate Institute

Virus–host interactions are key to ecosystems, including marine microbial communities, where they influence host composition, evolution, and global biogeochemical cycles. Phage infection begins with the recognition of host surface proteins (receptors) by receptor-binding proteins (RBPs) on viral particles, subjecting both to continuous evolutionary pressures, as exemplified by our previous work. Here, we analyze prokaryotic genetic variation in a marine ecosystem using metagenomes from near-surface (~10 m) samples collected over four years (Dart et al., 2023). We processed 39 metagenomes (median: 34 million reads per sample) and aligned them against a custom marine prokaryote database. We identified 1,862 highly abundant prokaryotic MAGs (>10× coverage), ~93% of which were detected across all samples, demonstrating persistence despite relative abundance variations spanning two- to three-fold changes. This mirrors previous findings in the same ecosystem. While stable at the population level, these contigs maintained consistent genetic variation across all time points, with at least 70% of single-nucleotide polymorphisms (SNPs) shared among samples, independent of collection date. This contrasts with prior work showing temporal population differentiation in viral genomes. The observed long-term genetic stability suggests that fluctuating selection may not be a major force shaping variation in these host populations. Instead, neutral processes may dominate, maintaining genome-wide genetic diversity. These findings challenge traditional views of virus–host dynamics, where selective pressures are often assumed to drive genetic change, and emphasize the role of neutral processes in structuring marine microbial communities.

Functional olfactory polymorphisms in bark beetles

Jibin Johny

Czech University of Life Sciences Prague

Eurasian spruce bark beetle Ips typographus, a natural part of forest ecosystems, is a major threat to Norway spruce forests during outbreaks. Olfaction plays a crucial role in the survival and range expansion of these beetles and thus reverse chemical ecology-based management approachs are highly considered. Interestingly, our attempts to characterize one of the highly expressed odorant receptors, ItypOR33, located at chromosomal inversions, led to the discovery of population-level variants distributed with similar frequency across 18 European populations. Deorphanizing ItypOR33 and its variant ItypOR33a using the Drosophila empty-neuron system (DNS) revealed ItypOR33 tuned to amitinol, a heterospecific pheromone component in Ips spp., whereas its variant tuned to (S)-(−)-ipsenol, a conspecific pheromone component of I. typographus. The in silico approaches revealed the structural basis of variations and ligand-receptor interactions. However, no sex-specific differnces were observed in ItypOR33 expression, and the behavioral and responses to the ligand amitinol. Reporting population-level functional olfactory polymorphisms in a non-model organism, for the first time in bark beetles, provides the essential primary evidence to explore the survival and adaptation of this species in forests and possible eco-evolutionary dynamics of these variations.

Symposium 34: Open symposium session D

On the role of neutral diversity in interpreting Fst and LD

Marcy Uyenoyama

Duke University

Wright's F-statistics (especially Fst) and linkage disequilibrium (LD) remain widely-used indices of population structure and associations among loci. Many methods used to infer Fst and LD presume the existence of genetic variance, but view the dependence of these summary statistics on the magnitude of diversity as a shortcoming. To address this concern, a variety of scalings of Fst and LD have been proposed. Even so, none appear to have succeeded in removing the problematic dependence.

Adoption of a sample-based coalescence approach has permitted derivation of exact expressions for expected joint identities among loci and populations. For example, these results clarify the influence of tight linkage and inbreeding on measures related to LD. These results also serve to clarify the significance of the rate of neutral mutation to these fundamental indices.

The widespread practice of ignoring the rate of mutation of the genetic variation that serves as the basis for inferring LD and Fst may reflect a presumption that the rate of mutation is negligible relative to the rates of crossing-over and migration. However, under reasonable assumptions regarding effective number, the rate of mutation may be of similar magnitude to the rates of crossing-over and genetic drift. As a consequence, the pattern of association among nucleotide sites on the order of tens of kilobases apart reflect the joint influence of these processes. In addition, the nature of the mutational process (as reflected in the K-allele and infinite-alleles models of mutation, for example) affects the magnitude of Fst.

Lessons from ancient dog (kurī) palaeofaeces about the settlement of Aotearoa New Zealand

Meriam van Os

University of Otago

Approximately 750 years ago, Polynesian explorers became the first humans to settle Aotearoa New Zealand. Archaeological evidence suggests that upon their arrival, Polynesians and their dog companions quickly spread across much of the coastal North and South Islands. Coming from the tropical Pacific, they rapidly adapted to Aotearoa's temperate climate and mastered its resources. A distinct new culture emerged, known as Māori culture.

The diverse landscapes and environmental conditions across Aotearoa resulted in varying lifestyles. For instance, cultivation of staple Māori crops like kūmara (Ipomoea batatas or sweet potato) was more productive in the warmer North Island compared to the South. The relative abundance of palaeofaeces from kurī (the now-extinct Māori dog) across archaeological sites in Aotearoa, provides an opportunity to study these human-animal-environment interactions across diverse landscapes in greater detail.

This presentation will discuss preliminary findings from metagenomic shotgun analyses of sixteen kurī palaeofaeces. This comparative study includes samples from both the North and South Islands and covers different time periods, including initial Māori settlement/the Early period (1250-1450 AD), the Middle period (1450-1650 AD), and the European contact period (late 18-19th century AD). The sites studied are the Long Bay Restaurant site in the upper North Island (Early to Middle period, n=4), Kahukura in the lower South Island (Middle period, n=4), and two occupation phases from Sealers Bay Camp on Codfish Island/Whenua Hou in Southern Aotearoa (Early period and European contact, n=8). The findings will be placed in a broader archaeological context by integrating them with other archaeological results.

Bayesian Phylogenetic Inference of Protein Structure Evolution Using Angular Diffusion Model

Clementine Yan

The University of Auckland

Protein structures are more conserved than sequences over long evolutionary timescale. Incorporating protein structural information alongside amino acid sequences within a probabilistic inference model provides a novel way to improve the accuracy of estimating evolutionary trees. The angular diffusion model proposed by Golden et al., detects the dependencies between adjacent Cα atoms. Despite its potential, the angular diffusion model has not yet been integrated into a phylogenetic framework. Here, we propose to implement the angular diffusion model for phylogenetic inference. The model will be integrated into the Bayesian phylogenetics software BEAST 2 by extending the tree likelihood function to compute the pseudo-transition probability density described by Golden et al. The implementation of the angular diffusion model allows us to improve the phylogenetics inference of protein structural evolutionary trees. The new structural evolution model incorporates dependencies between neighbouring Cα atoms, which is a distinct improvement from previous structural phylogenetics inference models. Validation of the model are based on simulation tests, and the model is further evaluated using real protein structural datasets. We anticipate the model will capture a more realistic evolutionary trajectory, providing better estimates of tree topology and evolutionary parameters.

Drivers of avian genomic change revealed by evolutionary rate decomposition

Simon Ho

University of Sydney

Modern birds have diversified into a striking array of forms, behaviours, and ecological roles. Analyses of molecular evolutionary rates can reveal the links between genomic and phenotypic change, but disentangling the drivers of rate variation at the whole-genome scale has been difficult. Using comprehensive estimates of traits and evolutionary rates across a family-level phylogeny of birds, we find that genome-wide mutation rates across lineages are predominantly explained by clutch size and generation length, while rate variation across genes is driven by the content of guanine and cytosine. To find the subsets of genes and lineages that dominate evolutionary rate variation in birds, we estimated the influence of individual lineages on decomposed axes of gene-specific evolutionary rates. We find that most of the rate variation occurs along recent branches of the tree, associated with present-day families of birds. Additional tests on axes of rate variation show rapid changes in microchromosomes immediately after the Cretaceous–Palaeogene transition. These apparent pulses of evolution are consistent with major changes in the genetic machineries for meiosis, heart performance, and RNA splicing, surveillance, and translation, and correlate with the ecological diversity reflected in increased tarsus length. Collectively, our analyses paint a nuanced picture of avian evolution, revealing that the ancestors of the most diverse lineages of birds underwent major genomic changes related to mutation, gene usage, and niche expansion in the early Palaeogene period.

Adaptive coupling of chromosomal inversions with multi-locus ecological and mating-bias genotypes facilitates sympatric speciation

David Hernandez

Evolution Research Inc

Sympatric speciation relies on mechanisms that promote differential assortment of adaptive traits between distinct niche ecotypes. Most phenotypic traits in nature are polygenic. Chromosomal inversions (CIs), which can link adaptive alleles into supergenes and protect them from recombination, are thought to facilitate speciation in polygenic environments. However, their precise role in sympatric speciation remains unclear. To explore this, we extended our previous two-allele model of a mating-bias barrier by developing computational tools to simulate the invasion and population dynamics of multi-locus ecological and mating-bias genotypes. We then examined how the invasion of CIs can facilitate or hinder sympatric speciation in polygenic systems. Our results confirm that disruptive ecological selection promotes the establishment of premating reproductive isolation (RI) as the number of ecological gene loci increases, while the presence of viable hybrids hinders RI establishment. CIs that capture only locally adaptive ecological alleles weaken preexisting premating RI through negative coupling. In contrast, CIs capturing both adaptive ecological and mating-bias alleles can strengthen RI across entire genomes. Viable hybrids lead to CI polymorphisms within niches, maintained through immigration-selection balance. Reducing the number of mating-bias gene loci minimizes hybrid production and facilitates the emergence of premating RI. Additionally, the invasion of CIs carrying advantageous combinations of high mating-bias alleles enhances premating RI by reducing the effective number of mating-bias gene loci and limiting mating-bias hybrid production. By integrating mathematical modeling with computer simulations, this study advances our understanding of how polygenic traits and CIs contribute to the processes underlying sympatric speciation.

STICI: Split-Transformer with integrated convolutions for genotype imputation

Xinghua Shi

Temple University

Missing genomic data undermines accuracy and statistical power of evolutionary inferences and association studies. This missing data issue can be addressed through computational methods, often referred to as genotype imputation. Due to the importance of this issue, many methods have been proposed over the last decade, providing a high degree of accuracy in recovering the missing single nucleotide variants. However, it remains an open challenge to accurately impute structural variants due to the high complexity of structural mutation patterns. To address this challenge, we propose a novel deep learning model named STICI, which is a transformer-based convolutional framework for accurate genotype imputation of both structural and single nucleotide variants across the human genome. Through empirical studies on the 1000 Genomes project datasets, we showcase STICI's comparative imputation performance to existing state-of-the-art methods. Notably, STICI outperforms existing methods at imputing structural variants and does not require any conversions for multi-allelic variants, removing a layer of complexity from the imputation process of such variants. Moreover, the self-supervision training regimen of STICI enables improved genotype imputation by capturing intricate patterns across species, as evidenced by our experimental results on chicken and rat genotyping datasets.

Repeatable phenotypic but not genetic response to selection on body size of the black soldier fly

Wenjun Zhou

University of Cambridge

Polygenic traits are expected to show high genetic redundancy and therefore low repeatability in the genomic response to selection. We tested this prediction by selecting for large body size in the black soldier fly (Hermetia illucens). Over three replicate experiments selected for large body size, we found a strong and repeatable phenotypic response, with a 15% increase in body size. However, whole genome sequencing identified a highly polygenic and non-repeatable genomic response, with 120, 301 and 157 outlier genomic regions showing signal of selection identified across replicates, but there is high redundancy with only four shared regions across all three replicates. Among 12 candidate genes found in these regions, the insulin-like receptor gene was confirmed to be involved in larval growth regulation using a CRISPR knockout experiment. In summary, our study shows that artificial selection can generate rapid improvement for certain trait of industrial interest, and the highly polygenic quantitative trait body size shows high genetic redundancy.

Determinants of de novo mutation rates

Chaowei Charlene Zhang

the University of Hong Kong

Rates of germline mutations (μ) exhibit large variation among species, but less is known about their variability at the intraspecific level. Using a large dataset (n=464 individuals) from eight nine-spined stickleback (Pungitius pungitius) populations, we aimed at testing the prediction of the "drift-barrier hypothesis" that populations with smaller effective population (Ne) size should evolve higher μ than those with large Ne. We found a significantly higher μ in small-Ne freshwater (4.47 × 10-9/bp/gen) than in large-Ne sea (3.06 × 10-9/bp/gen) populations. However, this difference did not remain significant after controlling for a positive correlation between μ and maternal age, the latter implying an influence of reproductive age on μ. Furthermore, we found a positive association between μ and recombination rate across populations as expected if recombination is mutagenic. We are in process of estimating μ in multiple different stickleback lineages and species to get insights on range of variability and evolution of μ across divergence times ranging from ca 10,000 to 26 million years.

Large-scale chromosome rearrangements reshape genome architecture and lead to high karyotype diversity in ants

Janina Rinke

University of Münster

Genome structure and chromosome number are typically conserved over long evolutionary timescales, with large-scale rearrangements being rare. However, such rearrangements can substantially shape the structure and function of genomes, potentially driving adaptation and speciation. The evolutionary implications of chromosomal rearrangements leading to drastic karyotype changes remain poorly understood, due to a general stability of karyotypes with only small-scale genomic rearrangements between closely related species.

Here, we present a detailed comparative genomics study of the ant genus Cardiocondyla and unravel an extraordinary karyotype diversity and amount of genomic rearrangements within a short evolutionary time frame (1-10 million years) and even between lineages of the same species. We find extensive chromosome rearrangements, resulting in a complete loss of macro- and microsynteny, together with substantial genome reduction, while other genomic features remain conserved across species.

Additionally, we discovered a single, uniquely conserved ~6 Mb region analogous to a supergene in the genus, suggesting an uncoupling of genomic rearrangement mechanisms in this region. Our findings challenge the notion of large-scale structural conservation and position Cardiocondyla as a compelling system for studying drastic shifts in genome architecture evolution.

Horizontal gene transfer and predation facilitate kin-selected social digestion in the nematode Pristionchus pacificus

Ziduan Han

Northwest A&F University

Complex behaviors can emerge from newly gained traits cooperating with existing abilities. However, few examples reveal how these behaviors have evolved. In the nematode Pristionchus pacificus, developmental plasticity and horizontal gene transfer together may facilitate social behaviors. P. pacificus has polyphenic mouth morphs that can develop into either a predatory form, capable of preying on distantly related nematodes, or a bacterial-feeding form, suggesting different ecological roles in nature. It has also later, horizontally acquired cellulase genes and their cellulolytic ability expands the nematode's dietary range via digesting cellulosic bacterial biofilms. Unlike known plant-parasitic nematodes, which secrete cellulases directly into host cells, P. pacificus secretes cellulases into the environment and the digested cellulosic materials might be exploited by other P. pacificus individuals. When CRISPR-knockout cellulase-null mutants are grown on cellulosic bacteria with wildtype animals, our preliminary data show that the fitness of cellulase-null mutants increases. This indicates the non-cellulolytic individuals benefit from enzymes secreted by wildtype animals. However, the cellulolytic ability alone might be deleterious in the presence of interspecies competition. We hypothesize that kin-selection through predatory behavior is necessary. In addition, the non-predatory form displays higher levels of cellulase expression than the predatory form. Through genetic manipulation in P. pacificus, we are testing how different forms could affect the social behaviors. We are using forward genetic methods to identify regulators of cellulase genes and exploring their interactions with the developmental plasticity gene regulatory network. Thus, our work may reveal new insights into the evolution of complex social behaviors in nematodes.

Exact path-integral representation of the Wright-Fisher model with mutation and selection

David Waxman

Fudan University

The Wright-Fisher model describes a biological population containing a finite number of individuals. I consider a Wright-Fisher model for a randomly mating population, where selection and mutation act at an unlinked locus. The selection acting has a general form, and the locus may have two or more alleles. I determine an exact representation of the time dependent transition probability of such a model in terms of a path integral. Path integrals were introduced in physics and mathematics, and have found numerous applications in different fields, where a probability distribution, or closely related object, is represented as a 'sum' of contributions over all paths or trajectories between two points. Path integrals provide alternative calculational routes to problems, and may be a source of new intuition and suggest new approximations. For the case of two alleles, I relate the exact Wright-Fisher path-integral result to the path-integral form of the transition density under the diffusion approximation. I determine properties of the Wright-Fisher transition probability for multiple alleles and show how, in the absence of mutation, the Wright-Fisher transition probability incorporates phenomena such as fixation and loss. This work leads to insight into the interplay of random genetic drift and selection, and paves the way for novel approximations.

Harnessing AI for the problem of multiple sequence alignment

Tal Pupko

Tel-Aviv University

Multiple sequence alignments (MSAs) are extensively used in biology, from phylogenetic reconstruction to structure and function prediction. Here, we suggest an out-of-the-box approach for the inference of MSAs, which relies on algorithms developed in the field of natural language processing (NLP), when processing natural languages. We show that our artificial intelligence (AI)-based methodology, BetaAlign, can be trained to align sequences by processing alignments that are generated via simulations, and thus different aligners can be easily generated for datasets with specific evolutionary dynamics attributes. BetaAlign is the first deep learning aligner, which draws on NLP techniques. It trains transformers to map a set of unaligned biological sequences to an MSA. We show that our approach is highly accurate, comparable and sometimes better than state-of-the-art alignment tools. We characterize the performance of BetaAlign and the effect of various aspects on accuracy; for example, the size of the training data, the effect of different transformer architectures, and the effect of learning on a subspace of indel-model parameters (subspace learning). In addition, we also show how deep-learning networks can help rank alternative alignments, and thus improve alignment quality. Our findings uncover the potential of NLP-based methods for sequence alignment, highlighting that AI-based algorithms can substantially challenge classic approaches in phylogenomics and molecular evolution.

Genotyping Short Tandem Repeats Across Copy Number Alterations, Aneuploidies, and Polyploid Organisms

Maria Anisimova

Zurich University of Applied Sciences (ZHAW)

Short tandem repeats (STRs) are a rich source of genetic variation, but are difficult to genotype. While specialized repeat variant callers exist, they typically assume a euploid human genome. This means recent findings regarding phenotypic effects of STR variants in human health and disease cannot be readily extended to polyploid organisms or cancer, which is characterised by copy number alterations (CNAs). Here we present ConSTRain, a novel STR variant caller that explicitly accounts for the copy number of loci in its genotyping approach. We benchmark ConSTRain using a euploid human 100X whole genome sequencing sample where it calls STR allele lengths for over 1.7 × 106 loci in under 20 minutes with an accuracy of 98.28%. Subsequently, we show that ConSTRain resolves complex STR genotypes in an artificial trisomy 21 sample and a polyploid Dwarf Cavendish banana harbouring a large duplication. Finally, we analyse a microsatellite instable colorectal cancer tumoroid, where ConSTRain tackles CNAs and whole-genome duplications. ConSTRain is the first STR variant caller that allows for the investigation of repeats affected by CNAs, aneuploidies, and polyploid genomes. This unlocks the investigation of STRs across a wide range of contexts and organisms where they previously could not be easily studied.

Symposium 35: Open symposium session E

Tracing the evolution of bacterial motility through protein structure

Caroline Puente-Lelievre

University of Auckland

Most bacterial swimming is powered by the flagellar motor, a nanomachine that self-assembles from up to 45 proteins into a membrane-spanning complex. Despite their importance, the evolutionary origins and structural diversity of flagellar proteins are not well understood. Solving these questions is essential to understanding bacterial diversity and the role of motility in key processes such as pathogenicity and biofilm formation. We conducted model-based phylogenetic analyses of flagellar proteins across 420 bacterial genomes from 44 phyla and used three-dimensional protein structures to trace evolutionary histories and explore protein diversity. Our results highlight the structural and functional diversity of the flagellar proteins and relatives and the potential of structural phylogenetics to uncover deep, sometimes obscure, evolutionary relationships and functional insights. Furthermore, we provide a framework to categorise flagellar proteins based on their structural features and functions.

On the origin of de novo protein-coding genes in Arabidopsis

Yalong Guo

Institute of Botany, CAS

The origin of gene from non-genic sequences is a long-term and fundamental biological question. In particular, how de novo genes and integrate into the existing pathways to regulate phenotypic variations is largely unknown. Here, we studied the origin of de novo gene in the model plant Arabidopsis at the natural populations level. We selected seven candidate de novo genes. Subsequently, we revealed that SWK, a de novo gene that originated from a non-coding sequence in Arabidopsis, plays a role in seed germination under osmotic stress. SWK can be fully translated into an 8 kDa protein, which is mainly located in the nucleus, and is primarily expressed in dry seed, imbibing seed and silique. Intriguingly, SWK was integrated into an extant pathway of hydrogen peroxide content (folate synthesis pathway) via the upstream gene cytHPPK/DHPS, an Arabidopsis-specific gene that originated from the duplication of mitHPPK/DHPS, and downstream gene GSTF9, to improve seed germination in osmotic stress. In addition, we demonstrated that the presence of SWK may be associated with drought tolerance in natural populations of Arabidopsis. Overall, our study highlights how a de novo gene originated and integrated into the existing pathways to regulate stress adaptation.

The changing mechanism and evolvability of a conserved function

Yeonwoo Park

Seoul National University

Are molecular mechanisms uniquely optimal or accessible solutions to functional needs or random and temporary choices among many possibilities? No study has comprehensively traced the mechanism of a conserved function across evolution to address this question. By assaying hundreds of extant and ancestral sequences and all of their point mutants, we reconstructed the structure, affinity, and evolvability of a protein-binding noncoding RNA across 440 million years of vertebrate evolution. Historical substitutions had negligible functional effects because selection acted throughout evolution to preserve an optimal affinity. They remodeled the structural basis of affinity, however, by continuously shuffling the distribution of pairings, wobbles, and mismatches and inducing occasional base-pairing frameshifts and conformational rearrangements affecting up to dozens of nucleotides. Slight strengthening of a part of the interface was followed by the weakening of another, gradually and randomly reshaping the pattern of affinity across the interface. Every substitution changed the effects of neighboring mutations, and these epistatic effects were often greater than the substitution's functional effect. Consequently, the set of alternate sequence states compatible with structure was quickly randomized between orthologs. Our findings highlight natural selection's blind spot: mechanistic and variational diversity evolves irrespective of selection because mutations keep arising and selection cannot distinguish functionally equivalent structures.

Diversity of vertebrate transposable elements in anurans

Ravinder Kanda

Oxford Brookes University

Most genomes consist of a substantial amount of sequences derived from transposable elements, including viruses (termed Endogenous Viral Elements, EVEs). While their diversity and impact on genome evolution has been well studied in humans and larger primates, considerably less is known about these insertions in the genomes of other species, especially outside of mammals. With a similar number of extant species and comparable genome sizes, Anurans represent an ideal group for comparison to mammals. We conducted a search for novel Endogenous Retroviruses (ERVs) through publicly available Anuran genomes. We identified 57 novel ERV families, represented by multiple loci, with several families shared across multiple species. The genomic structure of these insertions differs greatly to those currently recognised with retroviral classification, illustrating the mammal-centric bias in our current understanding of these genomic elements, and the previously unknown rich diversity of these genomic elements outside of mammals.

MAST-HMM: Fitting multiple trees from a single alignment with a Hidden Markov Model

Thomas Wong

Australian National University

Phylogenomic studies increasingly analyze thousands of loci, yet they often assume a single tree topology across loci. This assumption often fails to capture the complexity of evolutionary processes, as loci frequently exhibit distinct evolutionary histories due to biological processes such as incomplete lineage sorting (ILS), introgression, or horizontal gene transfer.

To address these limitations we recently introduced the Mixture Across Sites and Trees (MAST). While the MAST model improves on traditional approaches, it does not account for dependencies between sites within an alignment, even though such interdependencies are expected for the evolutionary history of neighboring sites. Here, we develop a Hidden Markov Model (MAST-HMM) extension of the MAST model, enabling the consideration of inter-site correlations in phylogenomic analyses, i.e., neighboring sites have a higher probability of sharing the same tree.

Simulations and empirical data analyses demonstrate that this enhanced approach captures underlying biological processes, suggesting that it will be useful for future empirical studies.

Comparative phenotypic analysis of rhodopsin supports stepwise evolution of dim-light vision in early vertebrates

Yang Liu

Shaanxi Normal University

The origin of dim-light vision in vertebrates has been attributed to an early duplication event in RH, which led to the formation of RH1 (rhodopsin) and RH2 (rhodopsin-like) opsins. Two sites (122 and 189) are known to critically underlie rhodopsin's high sensitivity at low light, however the evolutionary trajectory leading to dim-light vision remains unclear. Here we built a large comparative dataset of vertebrate RH1 and RH2 sequences, and inferred their respective ancestral protein sequences alongside their common ancestor. Sequence comparisons revealed that known critical sites for RH1 kinetics did not evolve at the origin of this molecule, and instead first evolved in RH (M122E) as well as later in the RH1 of ancestor of jawed vertebrates (P189I). To assess the phenotypic impacts of these changes, we performed in vitro protein expression. We first compared ancestral pigments and found that the retinal release rate of RH1 became significantly slower following its divergence from RH2, with further deceleration at the origin of jawed vertebrates. We also constructed single RH1 mutants, and found that the replacement S295A accounted for the observed difference in retinal release between RH and RH1. Finally, to validate these phenotypic consequences in vivo, we compared zebrafish overexpressing RH, RH1 or the mutant RH-S295A, and recorded similar swimming behaviour in the latter two treatments. We conclude that highly-sensitive dim-light vision in vertebrates evolved via three main stages of molecular adaptation, at the origins of RH and RH1, and subsequently in the ancestral RH1 of jawed vertebrates.

Gene duplication and the diverse diet of the black soldier fly

Daniele Kunz

University of Cambridge

Gene duplication is an important source of genetic novelty. Here I explore the role of gene duplication in the generalist diet of the black soldier fly, Hermetia illucens.

I developed a novel workflow using high-throughput microfluidics screening. Total RNA extracted from larval gut tissue was used to synthesize cDNA, which we cloned into E. coli for recombinant enzyme production. Activity screening employed a microfluidic droplet-based method, enabling quick and precise assessments of enzyme performance on various substrates. This study represents one of the few uses of a microfluidics-based approach in eukaryotic systems, highlighting its potential for exploring complex organisms and their enzymatic activities.

I identified a family of seven glycoside hydrolase enzymes in H. illucens that are duplicated and expressed in a tissue-specific manner—notably in the gut, fat body, and salivary glands. These enzymes exhibit broad substrate specificity, hydrolyzing fucopyranoside, glucopyranoside, xylobiose, and galactopyranoside. One enzyme is especially generalist, and can degrade both alpha- and beta-linked glycosides, as well as various sugars such as glucose, fucose, rhamnose, and arabinose in their furanose and pyranose forms. This is novel, as no Kochland glycosidase has previously been shown to degrade both types of linkages, so this is a uniquely generalist enzyme.

In summary, older duplication were associated with regulatory evolution leading to tissue-specific roles, while subsequent further duplications were associated with coding changes that give substrate versatility, and contribute to the species' dietary adaptability.

Evolution in metazoans of the temperature-sensing transient receptor potential (TRP) ion channel families TRPV and TRPM

Dan Larhammar

Guangdong Institute of Intelligence Science and Technology

The TRP superfamily of ion channels contains several members that are involved in temperature sensing, such as the Nobel Prize-winning TRPV1 (hot) and TRPM8 (cold) channels. Many members have other functional roles. We have investigated the evolution of these two families in metazoans. Both display ancestral complexity followed by loss of members and subsequent expansion in vertebrates, largely through their polyploidizations. The common ancestor of Cnidaria and Bilateria had three TRPV genes that became four in the deuterostome ancestor, two of which were lost in the vertebrate ancestor. The family expanded to nine members in early jawed vertebrates by local duplications and two polyploidizations, whereupon losses or additional duplications happened in some lineages. Heat-sensitive TRPV1 is present in all jawed vertebrates. The ancestor of the TRPM family added a melastatin and a nucleotide-binding domain. Five TRPM-like genes existed before the origin of bilaterians, but only two of these remained in the vertebrate ancestor. The vertebrate polyploidizations and a local duplication brought the number to eight in gnathostomes spanning a broad range of functions, five of which are temperature-sensitive. Two members have a kinase domain duplicated from a syntenic kinase gene. The nucleotide-binding domain has been lost multiple times independently. Interestingly, the cold sensor TRPM8 has been retained only in sarcopterygians. Whereas Drosophila has a single TRPM gene, the Atlantic salmon has as many as 16 after the salmonid tetraploidization. Thus, both TRPV and TRPM originated in early metazoans and display dynamic evolution with regard to gene duplications and losses.

Experimental evolution of gene regulatory networks in Drosophila

Xueying Li

Beijing Normal University

Gene regulation plays a critical role in development and evolution. Transcription factors and their target genes establish complex networks that define the spatial and temporal pattern of gene expression. However, how regulatory information is encoded in these networks and how the networks can change over evolutionary timescales remain open questions. Previsouly, we examined flies' short-term evolutionary response to extra copies of bicoid, a key morphogen in embryogenesis, by chemical mutagenesis and experimental evolution. We found that the compensation for extra copies of bicoid occurred through rapid increases in embryo size driven by maternal changes, outside of the focal developmental network. In order to perturb and evolve developmental networks in a more targeted manner, we designed TF-High-Evolutionary (HighEvo), a transcription factor (TF) fused with a base editor (activation-induced deaminase), to continuously induce germline mutations at TF-binding sites across regulatory networks in Drosophila. Populations of flies expressing TF-HighEvo in their germlines accumulated mutations at rates an order of magnitude higher than natural populations. Importantly, these mutations accumulated around the targeted TF-binding sites across the genome, leading to distinct morphological phenotypes consistent with the developmental roles of the tagged TFs. As such, this TF-HighEvo method allows the interrogation of the mutational space of gene regulatory networks at scale and can serve as a powerful reagent for experimental evolution and genetic screens focused on the regulatory genome.

Discovery of zymocin-like killer toxin gene clusters in the nuclear genomes of filamentous fungi

Padraic Heneghan

Conway Institute, University College Dublin

Zymocin-like killer toxins are anticodon nucleases secreted by some budding yeast species, which kill competitor yeasts by cleaving tRNA molecules. They are encoded by virus-like elements (VLEs), cytosolic linear DNA molecules that are also called killer plasmids. To date, toxins of this type have been found only in budding yeast species (Saccharomycotina). We discovered that the nuclear genomes of many filamentous fungi (Pezizomycotina) contain small clusters of genes for a zymocin-like ribonuclease (γ-toxin), a chitinase (toxin α/β-subunit), and in some cases an immunity protein. The γ-toxins from Fusarium oxysporum and Colletotrichum siamense abolished growth when expressed intracellularly in S. cerevisiae. Phylogenetic analysis of glycoside hydrolase 18 (GH18) domains showed that the chitinase genes in the gene clusters are members of the previously described C-II subgroup of Pezizomycotina chitinases, revealing their role and origin in these fungi. We propose that the Pezizomycotina gene clusters originated by integration of a yeast-like VLE into the nuclear genome, but this event must have been ancient because (1) phylogenetically, the Pezizomycotina C-II chitinases and the Saccharomycotina VLE-encoded toxin α/β subunit chitinases are sister clades with neither of them nested inside the other, and (2) many of the Pezizomycotina toxin cluster genes contain introns, whereas VLEs do not. We have also identified VLEs in some Early Diverging Fungus (EDF) lineages and present evidence supporting the hypothesis that all yeast cytosolic linear dsDNA VLEs originated from an EDF donor.

Flight style associated with energy metabolism predict genomic evolutionary rates in birds

Yanzhu Ji

Institute of Zoology, CAS

Evolutionary rates have been repeatedly identified to correlate with life history traits such as body mass and generation time, suggesting roles of basal metabolism or DNA replication that contribute to the rate variation. Nevertheless, the relative roles of metabolism and DNA replication on evolutionary rate heterogeneity are elusive. Studying other traits that relate to total metabolism may help to further shed light on this issue. For example, birds have evolved a range of flight styles, of which particular flight styles such as soaring and gliding flight are more efficient than others. Additionally, exploring different evolutionary rates with different mutagenesis mechanisms could help to stratify the impact of energy expenditure (if any) on genome evolution. Here by leveraging the bird genomes produced by Bird 10K project, we systematically evaluated three evolutionary rates and related them to life history traits and flight characters (including flight style) via a model selection approach. We found the best models that predict substitution rate and deletion rate, the rates with a metabolic component in mutagenesis, included body mass and flight style as predictors. The rate of length change of microsatellites, which did not include a metabolic component in mutagenesis, was related to body mass and/or generation time. Our results highlighted the role of flight styles in determining deletion rates and substitution rates by influencing the energy expenditure of flight.

Taming the beast: adaptive introgression from dogs might explain unique behavior in Iberian gray wolves

Carlos Sarabia

Autonomous University of Barcelona (UAB)

Introgression is one of the most important and rapid drivers of evolution, introducing novel alleles from one species or population into another, which can significantly influence adaptive capacity. While introgression from domesticated species into wild counterparts is often considered a source of maladaptive alleles that hinder adaptation, it can occasionally provide adaptive advantages by introducing beneficial variants, a topic that remains understudied. Gray wolves (Canis lupus) and domestic dogs (C. familiaris) share a close evolutionary relationship, yet hybridization between the two is rare in most populations. We investigated the Iberian Peninsula's geographically isolated gray wolves, which have coexisted with a large, loosely managed dog population for millennia in a human-modified landscape. Using whole-genome sequencing of 150 Iberian and Eurasian wolves and European dogs, we found minimal recent dog introgression despite significant demographic bottlenecks. We also identified low levels of ancient dog introgression (~5%) that have persisted for millennia. Through single-scan statistics, novel Local Ancestry Inference techniques and ancestry enrichment analyses, we identified six genes (DAPP1, NSMCE4A, MPPED2, PCDH9, MBTPS1, and CDH13) harboring introgressed dog alleles under positive selection in Iberian wolves. These genes from dogs under selection are associated with immune response, inflammation, and brain function, with one potentially explaining very unique behavioral traits—such as reduced dispersal—observed in Iberian wolves but not in their European counterparts. Our findings underscore the fascinating role of adaptive introgression in shaping the evolutionary trajectory of wild populations living in human-altered environments and stimulate discussions about the significance and implications of wild-domestic hybrids.

Tracing cell lineage in arabidopsis reveals common mutations in progeny from different branches, rare in somatic cells of parental plant

Wenfeng Qian

Institute of Genetics and Developmental Biology, CAS

The timing of germline cell segregation—when cells destined to become gametes are separated from somatic cells—plays a critical role in many evolutionary biology theories. It can influence how genetic material is inherited and how species adapt over time. However, the precise timing of germline cell segregation in plants remains uncertain. To address this, we traced cell lineage in Arabidopsis thaliana using a targeted dynamic editing strategy. A barcode sequence was continuously edited, and any mutations occurring in a cell were inherited by its daughter cells. We found that progeny from different branches of the plant tend to cluster together in the reconstructed barcode tree, and these progeny share mutations that are rare in the parental somatic cells. These findings suggest that some ancestral cells in the parent are more likely to become germline cells, indicating an early segregation of the germline. This hypothesis was further supported by observations in the Atmsh1 mutant, which results in chloroplast loss during development, leading to an albino phenotype. We observed that some offspring from the white pod exhibited green tissues, indicating they used different cell lineages. Our study has profound implications for understanding plant development and evolution.

Poster

Session 1

Symposium 1: Crop evolution: from Darwin's legacy to genomics of admixture and climate resilience

S1-P01-Analysis of genetic variation and complex traits regulation of tobacco on subgenome level

Le Yu

Tobacco Research Institute of Chinese Academy of Agricultural Sciences

Tobacco (Nicotiana tabacum) is an important model plant. Different types of tobacco show phenotypic diversity in agronomic traits, such as plant height, leaf shape, flower crown and flower color. Previous studies have proved that in many tetraploid plants, such as cotton and Brassica species, parallel selection on sub-genome level is associated with morphological diversity and convergent evolution. This makes tobacco a perfect model plant to investigate this question. We carried out whole genome re-sequencing analysis of 1,036 tobacco germplasm. 2,121,082 high-quality single nucleotide polymorphisms (SNPs) were identified. Population structure analysis shows tobacco germplasm could be divided into 7 clusters. Furthermore, we also carried out transcriptome sequencing and phenotypic measurements of plant type, leaf shape, flower crown and flower color. Genome-wide association study (GWAS) combined with expression quantitative trait loci (eQTL) analysis on subgenome level provides a comprehensive understanding of subgenome evolution, complex traits regulation in allotetraploid tobacco.

S1-P02-Recent geographic dispersal of indica rice post domestication

Yu-Chen Lin

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 weaker 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. The estimated split time of these populations, based on cross coalescent rate analysis, ranged from 500 to approximately 2,000 years ago, suggesting a relatively recent divergence. Notably, we observed evidence of gene flow between japonica and indica in populations from island Southeast Asia. Finally, we identified regions under potential positive selection in each population, which may point to local adaptation. In summary, we provide new insights into the evolutionary history of indica rice post domestication. The dispersal history would aid further studies on the geographical adaptation of indica rice.

S1-P03-Population genomic studies of wild cotton (Gossypium hirsutum)

Weixuan Ning

Iowa State University

Assessing the patterns and extent of genetic diversity in wild relatives of crop plants is essential to understanding their common origin and diversification. Although cultivars of Gossypium hirsutum produce most of the world's cotton fiber, diversity among wild cottons remains largely unexplored. Identifying G. hirsutum wild forms is challenging due to a complex domestication history combined with reciprocal introgression with a second domesticated species, G. barbadense, which generated numerous morphological forms and feral derivatives of both species and their interspecies recombinants. Multiple wild and naturalized populations are collectively scattered across a large geographic range in arid regions of the Caribbean basin. Using population-level sampling and whole genome resequencing data for G. hirsutum from populations scattered across the range (158 Yucatán peninsula, 166 Florida, 43 Puerto Rico and 25 Guadeloupe), we ask the following questions: 1) Can genomic data distinguish and identify truly wild cottons?; 2) What are the localized and regional population structures and genetic diversities among wild cotton population(s)?; 3) What are the genetic relationships between wild cottons from different localities?; 4) Do genomic data support the previous suggestion that the northern Yucatán is the ancestral home of domestication of G. hirsutum; and 5) What portion and proportion of wild diversity has been captured in modern cultivars that comprise the global crop plant today? Our study has broad relevance to our understanding of the status and ecology of wild cottons and implications for agronomic improvement.

S1-P04-Investigating genome-wide signatures of environmental adaptation in O. sativa ssp Indica

Isabel Bojanini

New York University

Climate change and increasing agricultural demand caused by population growth are risks for food security globally. To meet this growing demand, it is crucial to understand the mechanisms of organismal adaptation and develop climate-ready crops that are resilient to more arid conditions. Due to its broad geographical distribution, as well as its phenotypic diversity, Indica rice (Oryza sativa ssp. Indica) is a resource for elucidating how environmental variables have influenced genetic variation in this organism. In the present study, we investigate the whole genome data of 804 landraces distributed throughout Southeast Asia. We conducted a genotype-by-environment association study using SNP data and the first principal components from two climate categories: temperature (7 variables) and precipitation (6 variables), along with evapotranspiration (1 variable). We identified six genetic clusters in our dataset and performed an analysis of variance and Linear Discriminant Analysis on environmental data for each population. Results confirm that genetic clusters are coming from different environments and observed differences are not driven by spatial data. The association study was performed using Latent Factor Mixed Models (LFMM), a mixed-effects model that incorporates genetic structure as a random effect and genetic markers as fixed effects. Our results show 3754 SNPs associated with temperature; 479 with precipitation; and 562 with potential evapotranspiration. Gene analysis of significant SNPs show an enrichment for genes in pathways related to cold, salt, and drought tolerance, as well osmotic sensitivity response and bacterial disease resistance. These findings suggest how environmental factors could shape the genome of Indica rice.

S1-P05-MACHINE LEARNING FOR THE DETECTION OF NEXT-GENERATION GENETICALLY MODIFIED ORGANISMS THROUGH ANCESTRY INFERENCE

Sarai Reyes

UNINE

The production and trade of genetically modified (GM) crops are regulated by governments globally with a high degree of variation how rules affect the planting, marketing, labelling, and trading of GMs. To ensure compliance, accurate detection and screening methods for GMs are necessary. Traditional GMs are efficiently screened by existing molecular methods, however the advance of New Genomic Techniques (NGTs) including CRISPR/Cas9 introduce only minimal modifications (labelled NGT1 in the EU). NGT1 organisms are challenging to screen for if the modifications are not previously known. Here, we explore whether such screenings are still be feasible if the species context is taken into account. We apply an ancestry inference method for screening NGT1-like modifications as a proof-of-principle approach with an algorithm called low-depth copy model (LOCO). LOCO does not require a reference panel and uses genotype likelihoods as input suitable for low-input sequencing data. We focused on a diverse sample of unmodified rice cultivars and artificially created NGT1-like modifications by swapping chromosomal segments. We discuss the detection limits dependent on the scope of modifications, the genomic locations and the sequencing coverage. Future applications will include real-world NGT1 modifications and assess the suitability to implement the procedure in NGT1 screening efforts.

S1-P06-Genome-environment association evidence edaphic adaptation in rice landraces (O. sativa subsp. indica) from SouthEast Asia

Isaura Rosas Reinhold

New York University

Rice (Oryza sativa) is one of the most important food crops for humanity, providing nearly half of the global calories, is cultivated in more than 100 different countries under different environmental conditions, however as other crops and plant species, rice landraces are facing rapidly changing environments. Recent studies have shown the important role that soil components have in plant local adaptation. Soil is a key ecological factor from where plants obtain support, nutrients and water necessary for their survival. The aim of this study was to evaluate the effect of soil and radiation in the local adaptation of different rice landraces across geographical space using genome-environment association approach. We used 804 genomes of Oryza sativa subsp. indica landraces from SouthEast Asia corresponding to 6 genetic populations previously identified and latent factor mixed model (LFMM) to identify signatures of environmental selection. A total of 125 significant SNPs were detected in seven out of the 15 original variables included in the analysis. Variables that showed significant SNPs were volumetric fraction of coarse fragments, nitrogen, soil organic carbon content, soil water content capacity, phosphorus total, latitude and direct normal radiation. Gene ontology results revealed a set of genes with potentially key adaptive roles to the gradient of these edaphic and environmental variables. Specifically, the Fst test and minimum allele frequencies (MAF) shows that China's population is locally adapted to direct solar radiation, while Indonesia's population is adapted to organic carbon content and soil water content capacity variables.

S1-P07-Comparative Genomic and Functional Analysis of Olfactory Receptor Genes in Psylliodes Species to Determine Host Plant Choice

Zedi Gao

University of Hertfordshire

Olfaction plays a critical role in insect-host plant interactions, directly influencing behaviours such as feeding and mating. This study investigates the genetic and functional bases of host plant specificity in Psylliodes species, focusing on Psylliodes dulcamarae, a species closely related to Psylliodes chrysocephala, a major pest of UK oilseed rape crops. Using the sequenced genome of P. dulcamarae (Genome size: 1.03 Gb, N50: 18.0 Mb, 99.4% BUSCO completeness), along with genome and transcriptome data from other related Psylliodes and Phyllotreta species, we explore key genes influencing host plant preference.

Our evolutionary comparative genomics approach identifies genes crucial for host specificity, supporting the "Oscillation Hypothesis" model of diversification. Phylogenetic analyses, alongside selection pressure assessments on detoxification and olfactory gene families (including Odorant Receptors, ORs), highlight clades with higher omega (dN/dS) values, suggesting adaptive evolutionary forces shaping olfactory and detoxification mechanisms. Transcriptomic analysis of olfactory organs in brassica-feeding flea beetles further identifies candidate genes critical for host adaptation and pestification.

Additionally, protein modelling and docking studies on olfactory receptors interacting with volatile compounds (e.g., Isothiocyanates) emitted by host plants provide insights into potential targets for olfactory-based integrated pest management strategies. The findings from this study, led by Dr. Henrik Stotz and PhD student Zedi Gao, and partially funded by the Felix Cobbled Trust, have implications for the development of integrated pest control strategies targeting Psylliodes chrysocephala and other related species.

Symposium 2: Evolution and regulation of sex chromosomes and sex determination

S2-P01-Diversification of stickleback sex chromosomes driven by introgression

Dandan Wang

The University of Hong Kong

Recent studies on origin and processes driving sex chromosome evolution have highlighted introgression as a potential factor shaping sex chromosome evolution, including a study of one evolutionary lineage (EL) of nine-spined sticklebacks (Pungitius pungitius). We constructed haplotype-level assemblies for two additional P. pungitius lineages (WL and UK), as well as for their sister species, Pungitius laevis. Using whole genome resequencing data, we identified a short (~470 kb) sex determination region (SDR) in the WL Chr3, which shares the most recent common ancestor not with the X chromosome, but with the homologous region in P. laevis. This suggests that SDR in WL has been subject to introgression from P. laevis. While the SDR in the UK lineage remains unknown, the region homologous to WL SDR in UK lineage is divergent from that in WL and also subject introgression from P. laevis. This may suggest that sex determination mechanism in UK lineage is different from that in WL and EL. Our results highlight the extreme intraspecific diversity of young sex chromosomes and sex determination systems in sticklebacks and indicate important role for introgression in shaping the diversification of sex chromosomes in stickleback fishes.

S2-P02-Genetic Analysis of X Chromosome Suppression in the Male Germline of Drosophila melanogaster

Zhihui Bao

Faculty of Biology, Ludwig-Maximilians-Universität München, 82152 Planegg, Germany

Sex chromosome-specific processes, including dosage compensation and meiotic sex chromosome inactivation (MSCI), are crucial for regulating gene expression during germline development in various species possessing heteromorphic sex chromosomes. In somatic cells of Drosophila, the expression of X-chromosomal genes is upregulated in males to compensate for the dosage discrepancy. In germline cells, the canonical mechanism of dosage compensation is absent. Instead, a mechanism known as "X suppression" occurs in the male germline, which leads to reduced expression of X-linked genes and is thought to be analogous to the MSCI that occurs in mammals. However, the genes and regulatory mechanisms responsible for X suppression in the male germline of Drosophila remain poorly understood. To address this gap, we performed a forward mutagenesis screen and identified two mutant fly lines exhibiting elevated expression of a testis-specific, X-linked reporter gene. The males of both mutant lines are sterile and transcriptome analysis revealed widespread upregulation of both X- and Y-linked genes in the male germline, particularly in primary spermatocytes. By analyzing these mutant lines, we aim to shed light on this sex-, tissue-, and chromosome-specific mechanism of gene expression regulation.

S2-P03-Human single-cell RNA-sequencing data supports the hypothesis of X chromosome insensitivity but is ineffective in testing the dosage compensation model

Jiabi Chen

Sun Yat-sen University

A controversy in evolutionary genetics is whether active dosage compensation is necessary to resolve the gene dosage imbalance between the X chromosome and autosomes. ScRNA-seq data could provide insight into this issue. However, it ' s crucial to carefully evaluate whether inherent characteristics of scRNA-seq, such as the sparsity of detected genes, might bias the X:AA expression ratio in mammals. This study evaluated two common strategies for selecting genes in calculation of X:AA, namely, filter-by-expression and filter-by-fraction, with simulated scRNA-seq and bulk RNA-seq datasets. We found that both strategies produce an inflated X:AA, thus artifactually supporting dosage compensation. Analyzing an empirical human Smart-seq2 data, results from the filter-by-expression strategy suggested that X-linked genes were more highly expressed than autosomal genes, a pattern that is neither predicted by dosage compensation nor explained by genes escaping X-chromosome inactivation. However, the results of the filter-by-fraction strategy are consistent with the simulation. Furthermore, despite biasing for mean expression levels, we found that scRNA-seq data could be used to detect X-to-autosome expression noise differences as small as 10%, which enabled investigation into the distribution of genes that are more likely insensitive to gene dosage changes. Analysis of the empirical Smart-seq2 data revealed a 10-15% increase in expression noise for X chromosomes compared to autosomes, and a significant depletion of dosage-sensitive genes on X chromosomes. Overall, these results highlight the need to be cautious when interpreting scRNA-seq data, particularly when comparing the expression of different genes, and provide additional evidence for the hypothesis of X chromosome insensitivity.

S2-P04-Evolutionary dynamics of the Drosophila Y chromosomes

Xinning Dong

Peking University

Y chromosome plays a crucial role in sex determination and reproductive adaptation, however, it is highly heterochromatic and contains only a few genes. Due to the fact that most mammalian Y chromosomes are very ancient, it is difficult to study the degeneration of Y chromosomes. Fruit flies serve as an excellent model for research, as the highly variable Y chromosomes among Drosophila species exhibit different evolutionary states. Due to the recombination suppression, the Y chromosome exhibits a unique evolutionary pattern: it rapidly loses and gains genes, and reveals distinct genes composition and repetitive sequences in different Drosophila species. And in some Drosophila species like D. miranda, sex chromosomes fuse with autosomes to form a neo-sex chromosome system. Previous studies on the Y chromosome were limited to a few species and the constraints of genome assembly quality. Therefore, we assembled 14 high-quality Drosophila species genomes of with closely and distantly related relationships using single-molecule real-time sequencing. We then compared the DNA components of the chromosomes and found that one Drosophila species appears to have lost all Y-linked genes homologous to those of D. melanogaster. In a certain species strain, we discovered that 10% of females contained Y-linked gene fragments. These findings help us better understand the evolutionary dynamics of the Y chromosome.

S2-P05-Sex chromosome evolution in Australian burrowing frogs Neobatrachus

Jie Wang

Max Planck Institute for Plant Breeding Research

The evolution of the sex determination mechanism is one of the most fundamental questions in evolutionary biology, yet it is only well understood in a handful of vertebrate species. Polyploidization from whole genome duplication adds extra complexity to fully understand the genetic mechanism of sex determination in polyploid species, which remains largely unexplored, particularly in animals.

The Australian burrowing frogs of the genus Neobatrachus offer an excellent system to address the intriguing questions. It includes six diploid and three tetraploid species, yet all, possess homomorphic sex chromosomes and sexually reproduce.

We first assembled the genome of N. pelobatoides from a diploid female with PacBio and identified dmrt1 as part of its sex determination regions by conducting a scan of divergence between sexes on a population level. Subsequently, we assembled the genomes of one male and one female from each species within the genus using synthetic long reads (stLFR). Synteny analysis between these assemblies and the HiFi assembly of N. pelobatoides revealed that dmrt1 is consistently involved in their sex-determining region accordingly.

Interestingly, molecular evolution analysis across diverged vertebrate groups showed that dmrt1 exhibits a Neobatrachus-specific feature: at position 15 of DM domain, Neobatrachus has a tyrosine (Y), whereas the typical amino acid at this position is a serine (S, in dmrt1) or threonine (T, in dmy/dmw) in species like the diploid fish O. latipes or the allotetraploid frog X. laevis. Notably, the T-S distance is closer compared to T-Y or S-Y substitutions, highlighting a potential unique evolutionary trajectory for dmrt1 in the Neobatrachus

S2-P06-HLA inference, diversity and selection analysis in Mexico

Ruth Nanjala

Kennedy Institute of Rheumatology, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford

Immune-mediated diseases exert selective pressure on human evolution. Human leukocyte antigen (HLA) gene variations in the Major Histocompatibility Complex region (MHC) significantly influence the risk of immunological traits. Therefore, studying the diversity and selection within the MHC locus can improve our understanding of the human evolution process, particularly in Latin American populations, who remain underrepresented in genetics research.

We inferred HLA alleles from 10,000 whole-genome sequencing (WGS) samples from the Mexico City Prospective Study (MCPS) using a custom pipeline. We quantified and observed distinct diversity of HLA alleles and long-range haplotypes among individuals included in the MCPS cohort compared to other ancestral populations. We observed a total of 519 unique HLA alleles. As expected, for Class I, the B genes had more distinct alleles (153), compared to A (112), C (62) while the DRB1 had more unique alleles compared to (81), DQA1 (9), DQB1(29), DPA1(11) and DPB1(63) for Class II. Notably, we saw an overrepresentation of African ancestry in the MHC region among individuals with genetically admixed ancestry.Next, we constructed a HLA imputation reference panel based on classical HLA alleles inferred from WGS samples and 32748 genomic markers in the extended MHC region using a novel HLA-focused pipeline.

In this study, we present a comprehensive MHC map for a Latin American population and provide clarification of the dynamics that shaped the MHC locus. This work improves our understanding of how selection-driven alleles contributed to immune-mediated disease susceptibility today and will facilitate future MHC-disease association studies in Latino population.

S2-P07-Regulatory divergences in sex determination pathway cause hybrid male inviability in Caenorhabditis

Yimeng Gao

CNU

Sex determination is one of the critical developmental decisions undertaken by most biological organisms, showcasing a rich diversity across species. While the sex determination mechanisms in model organisms have been thoroughly studied, the variations in sex determination mechanisms among different species and their evolutionary implication remain obscure. Our previous work has identified that crossing Caenorhabditis nigoni males to C. briggsae females results in insufficient expression of Cbr-xol-1, an X-linked master switch responsible for intimately linked sex determination and dosage compensation pathways, consequently triggering aberrant X-chromosome repression in males, and ultimately leading to hybrid male inviability. Nevertheless, Cbr-xol-1 is expressed at normal level in C. briggsae males. Here, we identified a 1.8 Mb CbrII interval that can rescue lethality of hybrid males and is needed for normal Cbr-xol-1 expression in C. briggsae. Moreover, this locus harbors a pseudogenized repressor of xol-1. This autosomal xol-1 repressor is a paralog of the X Chromosome Signal Element (XSE), sex-1. The C. nigoni ortholog of this xol-1 repressor has intact open reading frame. Surprisingly, it upregulates xol-1 expression in native C. nigoni. In another word, this XSE paralog functions as an Autosomal Signal Element (ASE).

Our findings indicate that the evolutionary divergence in XSE and ASE within the sex determination pathway may be crucial drivers of reproductive isolation and species formation.

S2-P08-Extensive recombination suppression and genetic degeneration of a young ZW sex chromosome system in halfbeaks

Teng-Fei Xing

Institute of Oceanology, Chinese Academy of Sciences

Sex chromosome systems have evolved independently across the tree of life, at different times in the past, and the evolutionary consequences of lacking recombination in sex-linked regions have been characterized in many old-established systems. However, empirical studies of young sex chromosomes are still scarce, especially in vertebrates. Integrating whole-genome sequencing data of two species of halfbeak fish, Hyporhamphus sajori and H. intermedius, we identified the sex determining system in H. sajori as female heterogamety, involving a large fully sex-linked ZW region (~26 Mb) on chromosome 5. The closest relative, H. intermedius, has a small sex-linked region on a different chromosome, and shows male heterogamety, suggesting at least one turnover in this fish genus. The H. sajori sex-linked region includes two evolutionary strata, but the estimated Z-W divergence times are small, less than 3 million years for the older stratum, which is less than between the two species. Nevertheless, this evolutionarily young W-linked region is enriched with repetitive sequences, differs from the ancestral state by five inversions, and about one-third of its protein-coding genes have already become non-functional. Transcriptomic analysis suggests that some form of dosage compensation may already be evolving for some sex-linked genes.

S2-P09-Repeated evolution of transcript dosage compensation of independently formed nematode neo-X chromosomes

Jiachen Li

University of Michigan

Ohno proposed that, during the origin of X/Y sex chromosomes from a pair of autosomes, X-linked genes must double their per-allele expressions to compensate for the degeneration of their Y homologs. Whether Ohno's hypothesis holds in the nematode Caenorhabditis elegans remains unresolved despite that C. elegans is a model for studying between-sex X chromosome dosage compensation. Genome sequencing revealed independent fusions of an ancestrally autosomal linkage group to the X chromosome in C. elegans and Brugia malayi, two species belonging to different suborders of the order Rhabditida, allowing testing Ohno's hypothesis in repeated origins of neo-X chromosomes from the same autosomal linkage group. For each C. elegans X-linked gene and its autosomal ortholog in Pristionchus pacificus, we computed the X:AA ratio in transcript level and observed a median of approximately 1. The same is true for B. malayi X-linked genes when compared with their autosomal orthologs in Dirofilaria immitis. We find a significant enrichment of presumably dosage-sensitive transcription factor genes among the autosomal genes of P. pacificus (or D. immitis) that become X-linked in C. elegans (or B. malayi), but the results are mixed for other groups of presumably dosage-sensitive genes, providing a partial support to the hypothesis that X upregulation depends on the prevalence of dosage-sensitive genes in the proto-X. We conclude that, unlike the virtual absence of X upregulation at the transcript level in eutherian mammals, Ohno's hypothesis is strongly supported in both nematode lineages investigated.

S2-P10-Implications of Sex-Specific Genomic Differentiation in the Great Tit (Parus major) Species Complex

Toni Gossmann

TU Dortmund University

The generation and exposure of genetic novelties to natural selection are central to evolutionary processes and sex-specific variation may function as a source of genomic adaptation. In birds — characterized by highly stable karyotypes and recombination landscapes — rapid evolution appears to occur under seemingly constrained genomic conditions. In this study, we performed comprehensive genomic analyses of more than 200 wild great tits (Parus major)—based on WGS and SNPchip data—from five Asian and seven European populations to uncover substantial cryptic and sex-specific genomic variation. We identified more than 50 autosomal genomic regions spanning over 70 Mb that harbor cryptic genetic elements. These gene-dense regions, which harbor genetic variation that is either female-specific or shared between sexes, show a significant overrepresentation of reproduction-associated genes. Using a long-range deep sequencing approach, we discovered distinct haplotypes enriched in repetitive sequences with high GC and CpG content, which may stem from the sex chromosomes. These haplotypes contain hundreds of novel and divergent homologous genes under purifying selection, indicating functional conservation — particularly in pathways linked to female reproduction and development. We also address the technical implications of genome-wide sex differentiation, examining how sampling composition bias and SNP-based analytical limitations can skew inferences related to local adaptation, gene flow, and population structure. Our findings underscore the importance of rigorous quality control and sex-aware filtering strategies to account for these regions in population genetic analyses. At the same time, they demonstrate how sex-specific genomics can serve as a valuable tool for pinpointing genomic origins of rapid adaptation.

S2-P11-Uncovering the Genetic Basis of Heterodichogamy in Pterocarya Using a Low-Input Pan-Genomic Approach

Weining Bai

Beijing Normal University

The heterodichogamous mating system, characterized by two distinct mating types (protogyny and protandry), is rare among flowering plants, but it is present in nearly all species in Juglandaceae (the walnut family). While evolutionary parsimony suggests a single origin for this trait, recent studies reveal distinct genetic mechanisms in Juglans and Carya, indicating multiple independent origins. To further clarify the evolutionary origins of heterodichogamy in Juglandaceae, we investigated its genetic basis in Pterocarya, a genus closely related to Juglans. Using a pan-genome graph approach, we identified a structural variation region associated with mating types across the Pterocarya genus. This region includes 30 kb tandem repeats in the dominant allele and an insertion in the recessive allele, with shared polymorphisms spanning 67 kb from the 3′UTR of S12e to a Gypsy transposable element. Genomic synteny analysis confirms this region is non-homologous to those in Juglans or Carya. Downstream analyses suggest that small RNAs uniquely expressed from the tandem repeats regulate heterodichogamy. These findings provide novel evidences for the multiple independent originations of convergent genetic basis in regulating heterodichogamy in Juglandaceae and highlight the utility of pan-genome approaches in deciphering structural variation-associated traits.

S2-P12-The molecular evolution of ampliconic genes in primates

Emma Diepeveen

Aarhus University

Mammalian sex chromosomes arose from a pair of autosomes 170–180 million years ago and have since undergone episodic recombination suppression. This progressive loss of recombination has generated distinct evolutionary strata characterized by varying levels of X−Y divergence and has facilitated the accumulation of repetitive sequences. Ampliconic genes, located on the sex chromosomes, have multiple nearly identical copies and are largely expressed in the testes, suggesting their involvement in spermatogenesis and male fertility. Presence/absence and copy numbers of these ampliconic gene families vary widely across different species, even among closely related ones, indicating the rapid evolution of these gene families, and many are polymorphic in copy number within species. Their amplification may be driven by several evolutionary mechanisms, including dosage compensation to mitigate the effects of Y degeneration, selection for enhanced gene expression, and meiotic drive arising from a genetic arms race between the sex chromosomes. However, the precise mechanisms underlying their amplification remain unknown. Recent advancements in telomere-to-telomere and long-read sequencing technologies have overcome challenges associated with the repetitive nature of these regions, enabling detailed analyses of sex chromosome architecture in humans and other ape species. We present an evolutionary analysis of ampliconic genes on the X and on the Y chromosomes in humans, chimpanzees, bonobos, gorillas, orangutans and gibbons. We infer by phylogenetic analysis their duplication history and the pattern of non-synonymous and synonymous substitutions among species to infer their rates of adaptive evolution.

S2-P13-Evolutionary Genomics of Mamiellophyceae: Genomic insights into the lifecycle

Gaokang Liu

Sorbonne Université, CNRS, Integrative Biology of Marine Organisms Lab, BIOM UMR7232, F-66650 Banyuls-sur-Mer, France

Recombination, the exchange of DNA, is a hallmark of sexual reproduction and has been observed in most eukaryotic species. However, the molecular mechanisms underlying mating, such as the differentiation into distinct mating types, remain largely unknown in many lineages. In this study, I present my project on the evolutionary history of mating types in Mamiellophyceae, a class of photosynthetic green microalgae that includes the smallest photosynthetic eukaryotes, picoeukarotes (cell size <2 μm) which can be particularly abundant in sunlit coastal oceans. We sequenced an updated version of the Ostreococcus tauri genome using PacBio sequencing, incorporating iterative updates from previous versions, resulting in a Telomere-to-Telomere genome assembly. This updated sequence has been reannotated with improved resolution of complex repetitive regions. To explore mating type evolution in Mamiellophyceae, we completed genome projects for 9 additional species, including those with larger genomes up to 300 Mb. These efforts provide comprehensive insights into the evolution of mating types over more than 600 million years of evolutionary history.

S2-P14-Y Chromosome Expansion in Poecilia wingei

Yuying Lin

University of British Columbia

Sex chromosomes originate from a pair of autosomes. When one chromosome acquires a sex-determining locus, recombination between the proto-X and proto-Y chromosomes becomes suppressed. The non-recombining regions of the Y chromosome is prone to accumulating transposable elements. Poecilia wingei, shares a common origin of sex chromosomes with its closely related sister species, Poecilia reticulata, from which it diverged approximately 3 to 5 million years ago. Interestingly, cytogenetic evidence suggests that P. wingei Y chromosome is larger in size than the X, and is in fact the largest chromosome in the genome. This is in contrast to the P. reticulata Y, which is far more similar in size to the X., suggesting the P. wingei Y has undergone a rapid expansion. To investigate underlying genetic mechanism of Y chromosome expansion, we performed haplotype-resolved assembly of the genome of P. wingei using PacBio HiFi long-read sequencing in combination with Omni-C technology. Our results demonstrate that the expansion of the Y chromosome is predominantly driven by the accumulation of DNA transposons, followed by long terminal repeats (LTRs) and long interspersed nuclear elements (LINEs).

S2-P15-Epigenetic contribution to the evolution of separate sexes and sex chromosomes in the plant Silene latifolia

Aline Muyle

CNRS

Dioecy is a breeding system with separate females and males, where sex can be determined by sex chromosomes (for instance XY males and XX females). Dioecy is generally associated with the evolution of sexual dimorphism. In this study, we tested whether epigenetics plays a role in sexual dimorphism and sex chromosome evolution in Silene latifolia. We generated new data for female and male small RNAs and small RNA-mediated DNA methylation. We showed that small RNAs are globally sex-biased in S. latifolia. In particular, we identified a male-biased transcription factor which may potentially be regulated by sex-biased small RNAs and DNA methylation deposition. This transcription factor might contribute to the regulation of other male-biased genes, including key factors in sex-determination and phenotypic sexual dimorphism. Finally, we compared female and male epigenetic marks along S. latifolia sex chromosomes. The X and the Y have stopped recombining at different time points over portions of their lengths, called strata. We found that Y genes are significantly enriched in silencing promoter DNA methylation in the oldest stratum, suggesting epigenetics may be responsible for lower Y expression. In S.latifolia, reduced Y expression is partially dosage compensated by increased maternal X expression. DNA methylation of the female two Xs was often higher than the male single X. These differences may reflect the regulation of dosage compensation by DNA methylation.

Symposium 4: Gene trees and beyond: genealogical methods and their applications in evolutionary genetics

S4-P01-Powerful inference of bacterial evolution from genome-wide genealogies

Hongjin Wu

Shanghai Institute Immunity and Infection, Chinese Academy of Science

To make sense of the enormous diversity found in prokaryotes, it is first necessary to describe the processes that generated it. A particular challenge has been to account for the combined effects of clonal descent and homologous recombination. Recently, the technology of inferring genetic relationships from sequence data has advanced rapidly but the resulting algorithms have been designed with human genetic data in mind. We developed a software named BacRelate which can infer genome-wide genealogies and the clonal frame from bacterial genome data. BacRelate is adapted from the software Relate (Speidel, et al., 2019) and infers the trees in 4 steps: 1. construct local tree topologies using the standard Relate algorithm. 2. infer the clonal frame based on the local trees. 3. re-infer local tree topologies by using the clonal frame as a prior. 4. using the MCMC method to estimate the branches of trees. The resulting genealogical trees can be used for example to infer the recombination landscape along the genome, natural selection on specific loci and patterns of gene flow between bacterial clades. The wide utility of the method is demonstrated on data from Campylobacter jejuni, Escherichia coli, Streptococcus pneumoniae, Klebsiella pneumoniae and Vibrio parahaemolyticus. By allowing diverse features to be investigated within a single framework, BacRelate provides unprecedented insight into evolutionary processes in bacteria.

S4-P02-ScisTree2: An Improved Method for Large-scale Inference of Cell Lineage Trees and Genotype Calling from Noisy Single Cell Data

Yufeng Wu

University of Connecticut

Population admixture is an important genetic process that greatly affects the population demographic history of populations from many species. Admixture graph (AG) has been used extensively as a model for population admixture in the study of the evolution of populations using modern and ancient genomes. In the literature, AGs with ten or more populations and ten or more admixture events are being used for the understanding of complex population evolution. However, reconstructing AG of such complexity is challenging partly due to the astronomical size of possible AG topologies for this type of AGs. In 2020, I developed GTmix, a maximum likelihood AG inference method. GTmix uses inferred gene genealogies and aims at finding the optimal AG under the multispecies coalescent (MSC) model. One drawback of GTmix is that it cannot work with very complex AGs due to its computational inefficiency. In this talk, I will first describe how GTmix works. Then I will present GTmix2, an updated version of GTmix. I will show that GTmix2 significantly outperforms existing AG inference methods including TreeMix and AdmixtureTools2 in terms of inference accuracy. GTmix2 is much faster than the original GTmix. The current implementation of GTmix2 can reconstruct AGs with ten or more populations and up to six admixture events. Future releases of GTmix2 can potentially work for AGs with ten admixture events. Research partly supported by US National Science Foundation under grant IIS-1909425.

S4-P03-Phylogenomic insights into the determinants of bacterial growth rate variation

Chunhui Hao

East China Normal University

Bacterial growth rates exhibit substantial variation across species, ranging from doubling times of approximately 20 minutes to several years. Various genomic traits — such as genome size and rRNA operon copy number — and ecological and environmental factors, including niche breadth, pathogenicity, temperature, and pH, have been proposed to shape this variation. However, empirical evidence remains mixed, with some studies supporting these associations while others contradict them. Leveraging genomic data from 27,684 bacterial species, we applied various phylogenetic approaches, such as kernel ridge regression and path analysis, to identify three key patterns that reconcile these conflicting findings: (1) the relationships between growth rate and influencing factors are non-linear; (2) these relationships vary across evolutionary time scales, with stronger associations observed over relatively short time scales; and (3) the association between any factor and bacterial growth rate can be confounded by the presence of additional factors. Our findings highlight the limitations of previous studies that relied solely on pairwise associations, single time scales, or incomplete sampling, which may lead to biased conclusions on the determinants of bacterial growth rate variation.

S4-P04-The molecular phylogeny of Nassariidae (Mollusca: Gastropoda)

Mingjie Liu

Hainan University

The Nassariidae family, widely distributed in marine environments, comprises approximately 19 genera and over 300 species globally. The classification of Nassariidae has traditionally been based on morphological characteristics of the shell and radula. However, these traits often exhibit "homoplasy," making it difficult to objectively and accurately establish a robust classification system for Nassariidae based solely on morphology. This study involved a global collection of Nassariidae specimens and reconstructed the phylogenetic relationships of the family at the genus level using mitochondrial genome data. The results revealed that several genera previously classified within Nassariidae (including Cyllene, Anentome, Tomlinia, Engoniophos, Phos, and Antillophos) show more distant phylogenetic relationships within Buccinoidea, while only the subfamily Nassariinae and the genus Bullia accurately represent Nassariidae. The phylogenetic tree of Nassariinae received strong support and identified seven independent evolutionary clades, corresponding to different geographic groups. Biogeographic analysis, combined with the divergence time estimation, suggests that Nassariinae likely originated in the Indo-Pacific region during the late Paleocene. Subsequent paleoclimatic changes and continental drift events promoted diversification within the subfamily, leading to the current global distribution of its genera.

S4-P05-The molecular phylogeny of Caenogastropoda (Mollusca, Gastropoda)based on mitochondrial genomes and nuclear genes

Fengping Li

Hainan University

Comprising about 60% of gastropod diversity, caenogastropods display almost all kinds of shell forms and include many commercially important marine groups. Although the monophyly of Caenogastropoda has been widely accepted, thier internal phylogenetic relationships remain unclear. In the present study, a total of 27 caenogastropods belonging to eight superfamilies were sequenced and used for phylogenetic reconstruction. All newly sequenced mitogenomes adhered to the consensus gene order of caenogastropods, except for those of Vanikoroidea, Vermetoidea and Cerithioidea, which involved protein-coding genes. The reconstructed mitogenomic phylogeny suggested the monophylies of Architaenioglossa, Sorbeoconcha, Hypsogastropoda and the siphonate clade. The present study also identified a close affinity among Cypraeoidea, Ficoidea, Tonnoidea, and Neogastropoda, supported by the presence of a pleurembolic proboscis. The monophyly of Neogastropoda was not supported, as Cancellariidae was found to be sister to the limpet-shaped group Calyptraeoidea, and (Tonooidea + Ficoidea) were sister to the remaining neogastropods. This study provides important information for better understanding the evolution of caenogastropods, as well as for the protection and utilization of these diverse and economically significant marine resources.

S4-P06-Dating LUCA and early evolutionary events using gene duplications

Lina Uematsu

University of Bristol

The Last Universal Common Ancestor (LUCA) existed over four billion years ago and gave rise to the primary domains of life: Bacteria and Archaea. However, crown Bacteria and Archaea emerged nearly a billion years after LUCA. Combined with the incomplete nature of the prokaryotic fossil record, this long gap limits direct insights into early evolutionary processes. Here, we attempt leveraging ancient gene duplication events to infer evolutionary events that characterised the early evolution of life, focusing in particular on the three, key, stem lineages: the stem to LUCA, and those leading from LUCA to Archaea and Bacteria. We use a novel computational pipeline and assemble a new dataset of protein domain families containing pre-LUCA and post-LUCA duplications. We then apply gene-tree species-tree reconciliation methods to estimate whether domains are pre or post LUCA, linking duplication events to biological novelties. This approach improves the precision of LUCA ' s genome reconstruction and provide a first glimpse into the evolutionary events that led to the emergence of Bacteria and Archaea.

S4-P07-Incorporating punctuated evolution into Bayesian divergence time estimation

Robert Haobo Yuan

ETH Zürich

Determining the absolute timescale of lineage divergences is crucial to phylogenetic and macroevolutionary research. However, divergence time estimates derived from molecular clock dating analyses are often at odds with those suggested by the fossil record, especially for clades which are now highly diverse, such as placental mammals, crown birds, and angiosperms. One hypothesis for this discrepancy is that our current clock models assume gradual accumulation of character change over time and may overestimate node ages when this assumption is violated. Recently, a new clock model was implemented in BEAST2, the Gamma Spike model, to facilitate punctuated evolution. The model allows abrupt bursts of molecular or morphological change to take place at speciation events, in addition to the continual evolution permitted by standard clock models. The strength of bursts can vary between alignment partitions, allowing comparison of punctuation dynamics between genes. Here we assess the effects of accommodating punctuated evolution in Bayesian timetree inference, using a well-researched Eutherian total-evidence dataset encompassing major clades of placental mammals. We compare topologies and divergence times inferred using the Gamma Spike model with analyses applying commonly used gradual relaxed clock models, as well previous clock dating studies. In particular, we consider the results of our Eutherian analyses in the context of the Cretaceous-Palaeogene (K – Pg) boundary, and therefore potential drivers of mammalian diversification. Our findings suggest that neglecting punctuated evolution may lead to biases in divergence time estimates, potentially obscuring the timing and origins of major clades.

S4-P08-Fusang: a framework for phylogenetic tree inference via deep learning

Siwen Gao

Xi'an jiaotong University

Phylogenetic tree inference, a cornerstone of evolutionary biology, involves reconstructing evolutionary relationships from multiple sequence alignments (MSA). While classical methods such as Maximum Likelihood (ML) and Bayesian Inference (BI) remain widely adopted, their computational inefficiency poses significant limitations for large-scale phylogenomic analyses. To address this challenge, we present Fusang (http://fusang.cibr.ac.cn), a deep learning (DL) framework optimized for rapid and accurate phylogenetic reconstruction. Leveraging simulated sequence data, Fusang integrates a convolutional neural network (CNN) and bidirectional long short-term memory (Bi-LSTM) architecture to infer four-taxon tree topologies (Fusang_4taxa model), achieving 99.2% accuracy when incorporating InDel (insertion-deletion) polymorphisms. To scale to larger datasets, we implement a sliding window algorithm for variable-length sequences and a stepwise added beam search strategy for multi-node tree assembly. Our simulation protocol defines 12 evolutionary scenarios that mirror real-world biological complexity by accounting for sequence coding properties, length variation, and nucleotide base distribution. Benchmarking against traditional ML methods on both simulated and empirical datasets demonstrates Fusang ' s competitive performance in topological accuracy, computational efficiency, and generalization capacity. Crucially, Fusang ' s modular design enables future enhancements, such as scenario-specific fine-tuning using real-world sequence data, positioning it as a scalable alternative to conventional phylogenetics tools.

S4-P09-Evolutionary genomics approach to decipher genome stability in long-lived animals

Erik Widen

Tohoku University

Species lifespan varies by orders of magnitude across the animal kingdom, yet the fundamental mechanisms driving these differences in aging remain incompletely understood. Recent studies have established that somatic mutation rates inversely correlate with mammalian species longevity and that similar mutational processes are causing a majority of mutations across those species, suggesting a mechanistic link between genome stability and lifespan. To investigate the evolutionary basis of this relationship, we analyzed 55 mammalian genomes with CAFE5 and RERconverge to identify genetic patterns associated with enhanced longevity. Our analysis revealed significant enrichment of over 200 genes in long-lived species, with a subset involved in genomic stability through telomere maintenance,DNA repair mechanisms, and oxidative stress response, while others were related to longevity through diverse processes including immune function and cancer suppression. These findings demonstrate consistent evolutionary signatures across mammals that connect lifespan extension with enhanced genome maintenance, as well as other theorised mechanisms of longevity. This analysis was limited to mammals with available amino acid sequences on NCBI, which drastically limited the available species for analysis. A modified pipeline that includes all mammals with sequenced genomes would further help illuminate this question. Additionally, expanding the analysis beyond mammals to other taxonomic groups could provide broader insights into longevity mechanisms. While these additional studies could further demonstrate consistent evolutionary signatures across animals connecting lifespan with genome stability-related pathways, experimental validation would be needed to directly test the causal relationship between these identified genetic pathways and lifespan extension.

S4-P10-Coalescent theory of the ψ directionality index

Egor Lappo

Stanford University

The ψ directionality index was introduced by Peter and Slatkin (2013) to infer the direction of range expansions from single-nucleotide-polymorphism variation. Computed from the joint site frequency spectrum for a pair of populations, ψ uses shared genetic variants to measure the difference in the amount of genetic drift experienced by the two populations, associating excess drift with greater distance from the origin of the range expansion. Despite successful applications of ψ to a variety of natural populations, statistical properties of the index have not been well understood. Here we define ψ as a random variable originating from a coalescent process in a fixed two-population demography. For samples consisting of a pair of diploid genomes, one each from two populations, we derive expressions for the moments $E\psi^k$ for standard parameterizations of bottlenecks during a founder event. For the expectation of ψ, we identify parameter combinations that represent distinct demographic scenarios yet yield the same value of $E\psi$. We also show that the variance $V\psi$ increases with the time since the bottleneck, the bottleneck severity, and the size of the ancestral population. Finally, we analyze the values of $E\psi$ computed from existing demographic models of Drosophila melanogaster and compare them with empirically computed ψ values. Our work builds the foundation for a theoretical treatment of the ψ directionality index and can help in evaluating results from empirical applications.

S4-P11-Demographic Inference with Inferred ARGs: An EM-Based Approach Incorporating Coalescent Time Uncertainty

Jacky Kaiyuan Li

UC Berkeley

New methods for estimating Ancestral Recombination Graphs (ARGs) are promising to transform population genetics, and the inference of demographic parameters, by providing access to full likelihood functions for large whole-genome datasets.

However, fully harnessing this potential requires the development of statistical methods that can effectively utilize ARGs while addressing the inherent uncertainty in coalescent time estimation - a critical factor for accurate population size inference.

Simulation studies reveal that existing methods often are biased due to the misplacement of coalescence events in adjacent time intervals. To overcome these limitations, we introduce a novel method for inferring effective population sizes by integrating ARGs with an Expectation-Maximization (EM) algorithm, which explicitly incorporates coalescent time uncertainty into the estimation framework. Using msprime to simulate large sequencing datasets and SINGER and POLYGON to infer ARGs, we demonstrate that our new method provides accurate estimates of effective population size over time, effectively correcting biases observed in existing methods.

This approach exhibits superior statistical properties in terms of Mean Square Error inthe estimation of effective population size, establishing the method as a powerful andreliable tool.

S4-P12-Brachiopod phylogenomics and the Cambrian evolution of biomineralization

Mingzhu Yang

University of Bristol

Brachiopods first appear in the Cambrian and dominated Paleozoic benthic marine environments. Brachiopods have a shell that can either be calcitic or phosphatic and has high fossilisation potential. Accordingly, they play a key role in understand biomineralization. However, fundamental questions remain about brachiopod monophyly and the relationships of phosphatic versus calcitic brachiopods. We assembled a phylogenomic dataset (130,000 AA positions) using the BUSCO database and performed Maximum Likelihood analyses using a novel across-site compositional heterogeneous model (CAT-PMSF).

Here we present the first phylogenomic tree of Brachiopoda and their relatives, encompassing all five extant orders (Lingulida, Craniida, Rhynchonellida, Thecideida, and Terebratulida). We found strong support for Brachiopoda as a monophyletic group within a clade of " true lophophorates " including also Bryozoa and Phoronida. Brachiopoda is resolved as the sister group of Bryozoa plus Phoronida, in agreement with recent phylogenies of Spiralia. Within Brachiopoda, we found support for the clade of inarticulated Linguliformea (Linguloidea, Discinoidea) and Craniiformea (Craniida), sister to the articulated Rhynchonelliformea (Rhynchonellida, Thecideida, and Terebratulida), consistent with previous morphological phylogenies. At the ordinal level, we resolved thecideides (Pajaudina) and the " long-looped " terebratulides, as sister to the " short-looped " terebratulides. Thecideida is therefore unexpectedly nested within Terebratulida, implying reduction in their morphology as body size decreased. We found that the phosphatic linguloids and discinoids are nested within the calcitic brachiopods (Linguliformea sister to Craniiformea; e.g., monophyly of inarticulated brachiopods). This result implies that there must have been at the least one transition in primary skeletal mineralogy within Brachiopoda.

S4-P13-Rooting the tree of eukaryotes using pre-LECA gene duplications

Chenxue Lin

University of Bristol

Eukaryotes constitute a vast and diverse biological group, ranging from unicellular organisms to multicellular plants and animals, with a complex evolutionary history. However, the evolutionary history of eukaryotes is obscured by uncertainty of the position of the root within eukaryote phylogeny. Several competing root hypotheses have been proposed in recent phylogenomic studies, with differing predictions for the nature of the last eukaryotic common ancestor (LECA) and the sequence of eukaryote innovations. A particular challenge is the substantial evolutionary distance between eukaryotes and their closest bacterial and archaeal relatives, which can lead to phylogenetic artifacts that impede accurate placement of the root.

To address these difficulties and provide a complementary source of information on the nature of LECA, we investigated the evolutionary history of gene families that underwent gene duplication prior to the last eukaryotic common ancestor (LECA), pooling information across gene families using gene tree-species tree reconciliation. The logic of this approach is that each paralogue can be used to root its sibling, providing multiple independent tests of the position of root within eukaryote phylogeny that does not require inclusion of distant archaeal and bacterial homologues of eukaryotic genes. Our analyses support a root between unikonts (including animals, fungi, and amoebozoans) and bikonts (including plants, SAR, discobans and metamonads), one of the current leading hypotheses for the position of the eukaryotic root, and suggest that root hypotheses on, or within, metamonads may be a result of long branch attraction to prokaryotic outgroups.

S4-P14-Tip-calibration approaches for deriving evolutionary timescales using genetic data

Sidney Davies

University of Bristol

Tip-calibration approaches represent viable alternatives to more traditional node-based methods for deriving evolutionary timescales using molecular sequence data. Despite key differences in the way these methods are implemented, little has been done to investigate the impact of calibration strategy on the derivation of time-scaled phylogenetic trees. Using a dataset of living and extinct crocodilians, we systematically explored the performance of the most popular molecular clock methods in the context of a well-researched and oftimes phylogentically contentious empirical group. We find that for key nodes, tip-calibration methods using non-mechanistic methods of tree generation produce estimates that are much older than both the fossil record and more traditional node-based approaches. The fossilized birth-death process (incorporating information on lineage speciation, extinction, and fossil preservation) produces estimates more in line with previous research. We discuss the scope of methods available for deriving evolutionary timescales and their applicability for unravelling the dynamics of evolution using genetic data.

S4-P15-Inferring cell lineage trees with the Neighbor-Joining method using transformer-based metric learning

Junhyong Kim

University of Pennsylvania

Phylogeny depicts species lineage relationships, but lineages are an important generative process at many scales. In particular, organismal development involves a cell lineage tree. The knowledge of cell-lineage tree can help dissect developmental regulation of fate decisions and have applications in domains like tissue damage and repair.

Single cell technologies allow leveraging the molecular states of individual cell as inputs to phylogeny algorithms to reconstruct cell lineages. Single cell DNA data with phylogenetically informative replicative mutation rate is difficult to collect. Engineered systems that generate lineage-dependent bar-codes have limited capacity and cannot be applied to natural samples. Single cell transcriptome can be readily obtained but current data suggests that the overall transcriptome does not reflect lineages due to noise and parallel or convergent gene regulation. Here, we ask whether a latent transcriptome information exists that can be leveraged for cell lineage reconstruction.

We introduce CellTreeQM, a deep learning method based on a transformer architecture that learns an embedding space for single cell transcriptome with geometric properties optimized for tree-graph inference using the Neighbor-joining method. By formulating the lineage reconstruction problem as tree-metric learning, we systematically explore supervised, weakly supervised, and unsupervised training settings. We benchmark our method on lineage-resolved empirical datasets and simulations. Our results show that CellTreeQM significantly improves tree inference with minimal supervision and limited data, offering a scalable framework for uncovering cell lineage relationships. Furthermore, our approach can be applied to standard phylogeny problems to obtain optimal data transformation for tree reconstruction.

S4-P16-Gene concordance factor and site concordance factor for the resolution of relationships of major groups of Neoaves

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, "the magnificent seven" (I - VII) and three orphan groups (X, X1, X2). 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, conserved non-exonic elements (CNEEs), ultra-conserved elements (UCEs), and untranslated regions (UTRs)] from previous studies using gene concordance factor (GCF) (proportion that the relationship or the grouping appeared in gene trees) and site concordance factor (SCF) (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 clustering patterns of the major groups were supported by less than few percent of GCFs and 30% - 40% of SCFs, many of which were close to 33.3%, consistently for 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.

S4-P17-A Phylogenomic Pipeline for Mining and Integrating Genomic Data: A Case Study Using Ecdysozoa

Morgan Jones

University of Bristol

The rapid expansion of genomic data available in online repositories such as NCBI Genome provides an opportunity to increase taxonomic and genomic coverage in animal phylogenomics. However, the sheer volume and pace of data generation presents a challenge when building phylogenomic matrices. Genome superalignments containing thousands of genes are computationally intensive and often suboptimal; instead, single-copy ortholog alignments are a promising approach for accurate species trees (Pisani et al., 2022). The BUSCO pipeline (Simão et al., 2015), which evaluates genome completeness based on the presence or absence of universally conserved single-copy orthologs, has become a widely used tool for measuring assembly quality. However, recent studies show limited overlap in results between BUSCO and other orthology pipelines such as Orthofinder (Emms and Kelly, 2015; Liu et al., 2024). Here we introduce an automated system that continuously monitors newly uploaded genomes on NCBI databases and generates phylogenetic datasets. In the first instance, we focus our efforts on mining data for Ecdysozoa, an ecologically important superphylum which represents over 80% of described animal biodiversity. To enhance the reliability of capturing orthologous genes for our dataset, we developed a new conservative 'core BUSCO' set. This set comprises the intersection of single-copy orthologs identified using BUSCO and Orthofinder when applied to a set of high-quality genomes spanning the major phyla of Ecdysozoa. The pipeline automates dataset construction and ensures at any time the matrix produced is inclusive of recently published data. We demonstrate the system by building a new phylogeny of the Ecdysozoa.

S4-P18-Beyond Perfect Gene Trees: An Efficient Full-Likelihood Framework for Species Tree Inference

Wenjie Zhu

Ludwig-Maximilians-Universität München (LMU)

Gene tree-species tree discordance is widely recognized, and the multispecies coalescent (MSC) model has emerged as the standard framework for species tree inference from multilocus data. To address limitations of existing MSC methods, we present a novel methodological framework for species tree inference. Our approach bridges the efficiency of summary methods with the statistical rigor of full-likelihood approaches. We implement a two-step full-likelihood approach with sequential Bayesian inference, avoiding approximations of the MSC likelihood. The method accepts either estimated gene trees or posterior samples of gene trees to compute species tree likelihood under the multispecies coalescent process. Through simulations under varying levels of incomplete lineage sorting and gene tree estimation error, we evaluated our method against existing approaches. For methods that ignore gene tree uncertainty, we found a dual role of branch lengths: they provide valuable information for improving accuracy when gene trees are known, but become a source of noise when gene trees are estimated from sequence data. In the latter case, methods using only topological information demonstrate greater robustness. In contrast, our approach effectively propagates gene tree uncertainty in both topological and branch length inference, performing robustly across scenarios.

S4-P19-Inferring population structure using genome-wide short tandem repeats

Feifei Xia

Zurich university of appiled science & University of Zurich

Understanding population structure is essential for uncovering evolutionary processes. Short tandem repeats (STRs) have long been used as genetic markers in evolutionary and population genetics, yet their full potential for population structure inference at a genome-wide scale remains unexplored. Despite the increasing availability of large-scale genome-wide STR data, most existing genealogical approaches focus on single-nucleotide polymorphisms (SNPs), which may fail to capture recent evolutionary events due to their lower mutation rates.

Here, we present a novel STR-based phylogenetic and genealogical framework to reconstruct fine-scale population structure. By leveraging stepwise mutation models and efficient distance-based tree-building methods, we infer high-resolution evolutionary relationships from genome-wide STRs. Applied to the 1000 Genomes Project, our approach successfully resolves population differentiation and selection signals, providing more information than traditional SNP-based genealogy methods. Furthermore, we validated our framework for STRs in tumor genomes, demonstrating that distinct STR markers capture population structure and tumor evolutionary trajectories. These findings highlight the unique advantages of STRs in population genetic studies, motivating the integration of diverse genetic variations to enhance population structure inference.

S4-P20-Polymorphism-aware Framework to Analyse RNA-Seq Data for Trait Evolution Studies

Manel Ait El Hadj

University of St. Andrews

Phylogenetic inference is complicated by processes such as incomplete lineage sorting (ILS), where gene tree discordance arises from the stochastic retention of ancestral polymorphisms across speciation events. While integrating polymorphisms into phylogenetic models can improve resolution of ILS, current approaches; despite advances in DNA-based methods like the Polymorphism-aware phylogenetic Model (PoMo), remain underdeveloped for RNA data. classical differential expression tools such as DESeq2 and edgeR are designed for single populations and single genes and they ignore the phylogenetic context. To address this gap, we present a computational framework integrating RNA-Seq simulations, phylogenetic comparative methods, and machine learning to model polymorphisms and trait evolution. Our approach begins with in silico RNA-Seq data generation using the compcodeR package (Bioconductor), which simulates gene expression matrices under Brownian motion (BM) or Ornstein-Uhlenbeck (OU) models, conditioned on preset phylogenies and evolutionary scenarios. Simulated datasets are then combined with trait evolution analyses in RevBayes to infer parameters under BM/OU processes. Machine learning models, trained on simulated data, are then applied to empirical RNA-Seq datasets to detect subtle evolutionary patterns. We apply this approach using Mexican Goodeinae, a clade of viviparous fish exhibiting divergent reproductive strategies (including a rare reversion to oviparity) and sexual selection dynamics. By simulating RNA-Seq data reflective of ancestral polymorphisms, we reconstruct trait evolution while accounting for genetic variation typically excluded in traditional phylogenies. Our approach advances scalable methods for disentangling the roles of selection, standing genetic variation, and speciation drivers in adaptive diversification.

S4-P21-Representation Learning on Ancestral Recombination Graphs with Transformers

David Bonet Sole

University of California, Santa Cruz

Advances in ancestral recombination graph (ARG) inference have recently enabled the generation of unprecedented amounts of genealogical data, unlocking new avenues for leveraging deep learning in population genetics. In this work, we introduce ARGformer, a novel deep learning framework for ARG representation that leverages transformer architectures for genealogical inference. By encoding both inferred and extant haplotypes, our model learns rich latent representations of haplotypes across the genome through time. Our framework employs a self-supervised pre-training objective designed to induce learning of relatedness and reconstruction of genealogy at the node level. Furthermore, we integrate state-of-the-art transformer techniques to efficiently handle the immense scale of ARGs. We train ARGformer on simulated ARGs as well as on inferred ARGs incorporating genomes from ancient and modern populations. Our results demonstrate that ARGformer effectively learns context-dependent embeddings that capture both local genealogy and population structure. We explore how the learned embeddings can be applied for unsupervised ancestry clustering, and analysis of the learned self-attention maps reveals interpretable genealogical patterns. Finally, we present scalable training protocols that enable application to biobank-scale ARGs now available, demonstrating the framework ' s broad applicability to challenges in evolutionary genetics and genealogical inference.

Symposium 5: Human history and past social organisation in the light of palaeogenomes: new methods, new findings session A

S5-P01-Population history and cold adaptation in Upper Paleolithic hunter-gatherers of the Japanese archipelago

Yusuke Watanabe

University of Tokyo

Modern East Asians trace their ancestry to ancient migrations originating from Southeast Asia. Nonetheless, the biological adaptations that facilitated their survival during the south-to-north migration across eastern Eurasia during the Last Glacial Maximum (LGM) are not yet fully understood. Our analysis of 42 genomes (up to 30x coverage) from prehistoric Jomon hunter-gatherers in the Japanese archipelago reveals that they descended from Upper Paleolithic (UP) foragers who became geographically and genetically isolated in this region during the Late Pleistocene. Phylogenetic analyses indicate that the Jomon lineage diverged from continental East Asians between 27,000 and 19,000 years ago. This long-term isolation not only shaped the Jomon as a distinct lineage but also contributed to the preservation of unique genetic characteristics, providing critical insights into the evolutionary history of East Eurasians. Building on this foundation, we identified strong evidence of polygenic adaptations to extreme cold environments in the Jomon lineage. Mechanisms such as non-shivering thermogenesis (NST), mediated by positively selected alleles in genes like UCP1, APOA5, and FTO, likely enabled survival during the Last Glacial Maximum (LGM). Furthermore, selection signals in FADS genes indicate enhanced metabolism of omega-3 polyunsaturated fatty acids (PUFAs), derived from high-fat marine diets. These findings, combined with archaeological evidence of marine resource utilization, suggest a dietary adaptation that complemented thermogenic strategies for cold environments. By examining the Jomon as a key lineage, our findings enhance understanding of ancient human survival strategies and provide valuable context for exploring the evolutionary basis of metabolic traits in modern populations.

S5-P02-Ancient genomes reveal coexistence of demic diffusion and cultural assimilation in the development of Bronze Age in southwest China

Jianxin Guo

Kunming Institute of Zoology. CAS

Southwest China is vital for understanding the dispersal and development of Bronze because of the highly developed bronze civilization of Sanxingdui culture in this region ranging from 4000 to 2600 years before present (BP). However, the process of the Bronze Age transition in southwest China is largely unknown due to the lack of ancient DNA data from the Bronze Age period. Here we report genome-wide data from 3 human samples from the Lengjiayan and Wutuo sites dating from 3200 to 2000 BP in Sichuan Basin, southwest China. In the middle Bronze Age, the Lengjiayan ancients derived about 90% of their ancestry from millet farmers in Central Plain, suggesting a demic diffusion from the Yellow River region to southwest China. During the late Bronze Age to Early Iron Age, the dominant ancestry of the Wutuo ancients was likely derived from the ancient North Asia (ANA)-related nomad lineage. In addition, we have detected the extra rice farmer-related ancestry in the Lengjiayan ancients, indicating the rice agriculture-related human migration from southern China to southwest China in the middle Bronze Age. Our results suggest the coexistence of demic diffusion and cultural assimilation in the development of the bronze culture in southwest China.

S5-P03-Mass violence, age and gender in the Early Iron Age of the Carpathian Basin

Miren Iraeta-Orbegozo

University of Lausanne

This study presents the re-analysis of the mass grave at Gomolava, a 9th-century BCE burial site in the Carpathian Basin of Southeast Europe, through an array of cutting-edge bioarchaeological and biomolecular techniques applied to the remains of 77 individuals. The site is located at the crossroad of complex socio-spatial relations, divergent cultural traditions, and competing ideologies of landscape use. Lethal violence enacted mostly on women and children suggests a selective demographic bias. Employing a multidisciplinary framework, we suggest a model for the social context in which the violent events surrounding the death of those buried in Gomolava took place. We explore the cause of death through osteological analysis, discern mobility patterns and diet habits through isotope analyses, reconstruct their genetic ancestry through genomic analyses, ascertain the age and sex distribution across the complete assemblage using a combination of osteological examination, genetic sexing and enamel peptide analysis and establish the age and time frame of the burial event using radiocarbon dating and micro-CT scanning. These findings shed new light on the socioeconomic roles and significance of women and young individuals in later European prehistory and aim to serve as a model for integrating genetic data across disciplines to construct meaningful, context-rich narratives.

S5-P04-Delineating multi-layered origins of human populations in Beringia using ancient and present-day genomes

Sara Jang

School of Biological Sciences, Seoul National University

Neolithic populations in Yakutia, including the Early/Middle Neolithic ones associated with Syalakh and Belkachi cultures and the Late Neolithic ones associated with Ymyyakhtakh culture, played a pivotal role in the spread of Siberian ancestry from northeastern Europe to Greenland. While previous studies investigated the genetic legacy of ancient human populations from Yakutia in central/southern Siberians, their relationship with indigenous populations around the Bering Strait (hereafter " Beringian " populations), including the Yukaghir, Chukotko-Kamchatkans, Asian Inuits, and Aleuts, remain elusive. Here, we searched for plausible historical admixture models connecting present-day Beringian populations with ancient ones from Beringia and broader Siberia. We show that Asian Inuits, Aleuts, and Chukotko-Kamchatkans are all adequately modeled as a mixture of the Belkachi-related populations from the Middle Neolithic Yakutia with the First Native American gene pool: for Chukotko-Kamchatkans the Native American ancestry is likely relayed via their mixture with neighboring Asian Inuits. In contrast, Yukaghirs, who live in northern Siberia further away from the Bering Strait, are better modeled as a mixture of the Ymyyakhtakh-related populations from the Late Neolithic Yakutia and neighboring Chukotko-Kamchatkans, delineating the northeastern margin of the Middle-to-Late Neolithic genetic change in Yakutia. This study sheds light on the fine-resolution genetic history of the indigenous populations in Beringia, linking Siberia and North America.

S5-P05-Slab Grave expansion disrupted long co-existence of distinct Bronze Age herders in central Mongolia

Juhyeon Lee

Seoul National University

Dairy pastoralism reached Mongolia during the Early Bronze Age (EBA) and flourished in the Late Bronze Age (LBA) alongside the emergence of diverse mortuary practices including the Deer Stone-Khirgisuur Complex (DSKC) and figure-shaped/Ulaanzuukh burials. While the spread of pastoralism has been widely studied, interactions between culturally distinct pastoralists remain underexplored due to challenges in obtaining genomic and cultural data. In this study, we analyzed genome-wide and cultural data from 30 ancient individuals in central Mongolia, a key region where pastoralists with distinct mortuary practices converged. We observe the enduring presence of two distinct genetic clusters during the LBA, corresponding to DSKC and figure-shaped burials, suggesting limited genetic mixing and a maintenance of cultural separation despite their coexistence. These groups were eventually replaced during the Early Iron Age by the expansion of the Slab Grave population and the establishment of a new burial tradition. Finally, we refine the genetic origin of the LBA DSKC populations, tracing their minor western Eurasian ancestry back to the Eneolithic/EBA Afanasievo and EBA Khemtseg (Chemurchek) populations. This study provides fine-scaled genetic tracking of major cultural transitions in prehistoric Mongolia, offering new insights into the complex and divergent processes that shaped the ancient pastoralist societies of Asia.

S5-P06-Selection on Semen Proteins: Human Genomic Clues to Neanderthal Mating Systems

Bahar Patlar

Martin Luther University

Seminal fluid (i.e., semen) proteins play critical roles in reproduction and are known to evolve rapidly, often under sexual selection. In this study, I investigate selection patterns acting on human seminal fluid proteins (Sfps) in comparison to other reproductive and non-reproductive proteins, emphasizing evolutionary changes in relation to hominin mating structures. Using divergence and polymorphism data from humans and chimpanzees, I examine signs of positive selection on Sfps derived from various reproductive tissues. Preliminary results have confirmed previous findings of significant divergence in seminal vesicle proteins -a major source of Sfps- between humans and chimpanzees, with prior studies identifying positive selection in some proteins, while preliminary results reveal widespread relaxed constraints, often accompanied by high polymorphism in humans while low in chimpanzees. Furthermore, at least one highly diverged seminal vesicle protein (dN/dS >> pN/pS) shows a polymorphism pattern in the Neanderthal genome that aligns more closely with chimpanzees than modern humans, suggesting similarities in mating systems, including sperm competition. This study provides new insights into how sexual selection has shaped reproductive proteins in hominins, offering the first molecular evidence for a potentially polygamous Neanderthal mating system.

S5-P07-Mapping Fine-scale Genetic Connectivity in Large Ancient Biobanks

Zhe Xue

University of Cambridge

Ancient DNA has become essential in rewriting human history. With recent exponential increases in ancient genome data production, it is now possible to reconstruct an increasingly detailed picture of global human genetic history with an expanding, multi-modal ancient reference panel. Current methods, however, have not yet been scaled with the size and complexity of large datasets. To address this challenge, we present a new graph-based framework, NORNE, to map the individual-level connections driven by genetic proximity in large datasets with exceptional scalability, enabling joint analysis of thousands of unphased genetic samples. On simulated and modern datasets, the framework captures fine-scale genetic structure at a resolution similar to haplotype-sharing methods. By applying this method to AADR dataset comprising more than 6,000 ancient samples across the world, we show that the density of connectivity naturally reflects fine-scale genetic groups that align with known archeological contexts. When integrated with spatial and temporal information, we found the detected connectivity between individuals reflects global migration events including spread of Farming, Yamnaya expansion, Avar and Viking expansions, and demonstrated sharp mobility increases during the Neolithic and Bronze Age, comparable to previous results based on Identity-By-Descent sharing.We further showcased individual-level connectivity analysis using cases of putative sources of Early European Neolithic Farmer-like ancestry in Bronze Age Britain, the Asian-influence change in Hungary, and the possible origins of the Roopkund lake skeletons. Our method provides an alternative way to study individual-level genetic relationships, and enables fast and robust analysis with large-scale biobank level references.

Symposium 6: Mutation rate rules: on the origin of germline mutations and their evolutionary fate

S6-P01-The evolution of mutation rates in the light of development and cell-lineage selection

Paco Majic

EMBL Heidelberg

In multicellular organisms, all mutations first arise in a single cell in a developmental context. Whether a mutation enters a population's gene pool can therefore depend on developmental events that affect the likelihood of mutant cell lineages of producing gametes. For this reason, the evolution of mutation rates in populations is governed not only by changes in the rates at which mutations occur at the molecular level, but also by changes in developmental features of organisms and how mutations impact cellular fitness during development. We present a theoretical framweork that, supported by empirical data from mammals and experiments on the fruit fly, demonstrate how generational mutation rates can be shaped by changes in developmental parameters and intraorganismal selective processes even when molecular mutation rates are presumed constant. This perspective has important conceptual implications. Unlike adaptationist explanations like the drift-barrier hypothesis, our model highlights how the diversity of mutation rates observed across animals may be the byproduct of organismal development rather than the result of direct selection against mutator alleles. Furthermore, this perspective broadens the role of organismal development in population genetics. Development not only introduces phenotypic biases but it also shapes the rates and trajectories of genetic diversity and is thus at the core of evolutionary theory.

S6-P02-A deep learning method for building INDEL mutation rate maps

Cai Li

Sun Yat-sen University

Short insertions/deletions (INDELs) are a common type of genetic variants and an important source of genetic diversity. DNA mutation rate is an essential parameter for many analyses in genetics, genomics and evolutionary biology. Mutation rates vary substantially across the genome, but there is a lack of high-precision models for predicting local INDEL mutation rates, which hinders the in-depth study of INDEL-related questions. Predicting INDEL mutation rates faces many challenges, such as relatively less mutation data and many mutation subtypes. We develop a deep learning-based method for predicting INDEL mutation rates, which learns signals from long-range sequence context of a focal site to predict its mutation rate. We leverage a large number of published genetic variants to evaluate model performance and select the best model. We apply our method to build high-precision INDEL mutation rate maps for several representative species and use the predicted mutation rates for downstream evolutionary analyses. We believe our method can facilitate many INDEL related studies in the future.

S6-P03-Exploring variations in rates of genome structural evolution among vertebrates

Yawako Kawaguchi

National Institute of Genetics

Genome evolution is influenced by both small-scale mutations, such as nucleotide substitutions, and large-scale structural variations, including genome size changes, gene order rearrangements, and karyotype alterations. These structural variations significantly affect phenotypic diversity, adaptation, and speciation. Despite their importance, quantitative analyses of these dynamics across diverse taxa have been limited. This study aimed to understand the constraints and evolutionary dynamics of genome structures among vertebrates. We utilized whole-genome data from over 400 vertebrate species, encompassing a range from small genomes (~300 Mb) in bony fish to large genomes (~40 Gb) in lungfish. To grasp the variability of large-scale structures, we focused on characterizing genome diversity by examining genome size, karyotypic organization, non-coding landscape scaling, and their intragenomic heterogeneity. Results indicated that the genomic scaling and heterogeneity are conserved within each class. Next, to understand the evolutionary rates of small-scale structures, we used phylogenetic trees based on micro-synteny blocks and compared them with trees based on amino acid sequences. Notable differences in evolutionary rates among species and groups were found, with increased rates observed in branches with whole-genome duplications, reflecting deletions post-duplication. In addition, the evolutionary rate showed a tendency to maintain that rate once it increased. Our findings suggest that lineage-specific factors and whole-genome duplications significantly impact structural changes, providing insights into the mechanisms of genome evolution and their effects on phenotypic diversity. Future research will explore the features and scales of structural variations and their associations with biological traits.

S6-P04-The mutation variation map for plant organellar genome

Zhiqiang Wu

AGIS

Since Palmer and Herbon noticed the divergences in the mutation mode of mtDNA evolution, the longstanding mystery that plant mitochondrial DNA (mtDNA) evolves slowly in sequence (mutation rate) but rapidly in structure has persisted for over 40 years. Subsequent analyses also confirmed and extended this paradox. Although mtDNA and ptDNA mutation rates have not been directly estimated in land plants, phylogenetic analyses indicate that mitochondria experience less than one substitution per site per billion years in many lineages. During the past years, more than one hypothesis had been proposed to explain the underlying mechanisms for those two-organelle genomes. Particularly, the processes of organellar genome replication, recombination and repair (RRR for short) are mediated by the nuclear-encoded proteins would be one of other hand to explore the mutation for them. Here, we used Arabidopsis thaliana mutation accumulation (MA) lines to detect and measure mutation rates for plant mtDNA and ptDNA.

S6-P05-The evolutionary dynamics of organellar pan-genomes in Arabidopsis thaliana

Yi Zou

AGIS

In plants, comparative analyses of organellar genomes often depend on draft assemblies, with large-scale investigations into the complex structural rearrangements of mitochondrial genomes remaining scarce. Here, we conducted a comprehensive analysis of the dominant conformations and dynamic heteroplasmic variants of organellar genomes in the model plant Arabidopsis thaliana, utilizing high-quality long-read assemblies validated at read-level resolution from 149 samples. We found that mitochondrial and plastid genomes share common types of structural and small-scale variants driven by similar DNA sequence features. However, rearrangements mediated by repetitive sequences in mitochondrial genomes evolved so rapidly that they often became decoupled from other types of variants. Rare complex events involving elongation and fusion were also observed, contributing to the unalignable regions commonly found at the interspecies level. Additionally, we demonstrated that disrupting and rescuing organellar DNA maintenance could drive the rapid evolution of dominant mitochondrial genome conformations. Our study provides an unprecedentedly detailed view of the dynamics of organellar genomes at pan-genome scale in A. thaliana, paving the way to unlock the full potential of organellar genetic resources.

S6-P06-Evolutionary Dynamics and Maintenance of Epimutations Underpin Transcriptional Stability in Stress-Adapted Mangroves

Tian Tang

School of Life Sciences, Sun Yat-sen University

The evolutionary dynamics of spontaneous epimutation rates — critical yet understudied drivers of stress adaptation — remain unresolved in extremophile plants. By integrating whole-genome bisulfite sequencing with epigenetic population genetics, we investigate this gap using mangroves, a model system for extreme coastal adaptation. Although mangroves exhibit spontaneous epimutation rates mirroring Arabidopsis (where loss rates exceed gains), population-level dynamics paradoxically favor CG methylation gain bias, a signature of strong selection absent in non-mangrove species. Selection to retain CG methylation is most pronounced in promoter regions, surpassing coding ssequences — a pattern aligning with genomic signatures of environmental adaptation. Transcriptomic analyses reveal that CG methylation gains, regardless of genomic context, correlate with reduced gene expression variability under stress, stabilizing transcriptional homeostasis. This resolves the apparent contradiction between methylation instability and population-level persistence: natural selection actively maintains functionally advantageous methylation states despite their intrinsic mutational lability. Our findings establish DNA methylation as a target of selection in stress-adapted plants, enabling evolutionary resilience through transcriptional stability in dynamic environments.

Symposium 7: The Origin, Evolution, and Phenotypic Contributions of New Genes

S7-P01-Nsun7 is a conserved gene essential for spermatogenesis however is it classical methyltransferase?

Ekaterina Guseva

The Skolkovo Institute of Science and Technology

NSUN7 belongs to the NSUN protein family, where all members catalyse the methylation of the 5th position of cytosine. NSUN7 is found in all multicellular animals and it primarily exhibits its functional role in spermatogenic cells. Loss-of-function mutations in this gene are associated with a decline in male fertility in both humans and mice. A detailed investigation of the male mouse phenotype revealed that NSUN7 deficient sperm is less motile and exhibits abnormalities in spermatozoon ultrastructure, the most remarkable of which is a change in the binding pattern of the longitudinal columns of sperm flagella. Proteins of the NSUN family utilise a "two cysteine" mechanism of methyl transfer, in which cysteines of motifs IV and VI are involved. This mechanism has remained constant throughout evolution of NSUN methyltransferases. However, a significant reduction in the conservation of these cysteines was observed among NSUN7 orthologs. Furthermore, mice carrying an amino acid substitution of cysteine of motif IV to alanine exhibited a normal phenotype and fertility rate, suggesting that the methyltransferase activity of NSUN7 is not essential for spermatogenesis or that NSUN7 has undergone changes in its catalytic residues. In addition, the analysis of the NSUN7 structure model revealed the presence of an additional cysteine from motif VIII. It is highly conserved among NSUN7 orthologs, and it is only found in the NSUN7 structure. Taken together, these findings indicate that NSUN7 has undergone alterations to its catalytic residues during the course of evolution.

S7-P02-Neofunctionalization of tandem duplicate genes in Arabidopsis

Chuanzhu Fan

Wayne State University

Tandem duplication, one of the major types of duplication, provides the raw material for the evolution of divergent functions. Among hundreds tandem duplicate genes in the genome of Arabidopsis, we explored the evolutionary pattern and functionality of one pair of tandem duplicate genes (AT5G12950 and AT5G12960) that originated within the last 16 million years after the split of Arabidopsis from the Capsella-Boechera ancestor. We systematically used bioinformatic tools to redefine their putative biochemical function as β-L-arabinofuranosidases that release L-Arabinose from the β-L-Araf-containing molecules in Arabidopsis. Comprehensive transcriptomic and proteomic analyses using various datasets showed divergent expression patterns among tissues between the 2 duplicate genes. We further collected phenotypic data from 2 types of measurements to indicate that AT5G12950 and AT5G12960 have different roles resulting in divergent phenotypic effects. Overall, AT5G12950 and AT5G12960 represent putative β -L-arabinofuranosidase encoding genes in Arabidopsis. After duplication, one duplicate copy developed diverged biological functions and contributed to a different phenotypic evolution in Arabidopsis.

S7-P03-Effect of codon use bias on horizontal gene transfer

Xinran Fan

Sun Yat-Sen University

Horizontal gene transfer (HGT) is a crucial source of new genes in genomic evolution and a major driver of microbial evolution. Previous research has suggested that, due to the cis-regulatory effect of synonymous codon usage bias (CUB) on the exogenous gene itself, HGT is more likely to succeed when the CUB of the exogenous gene matches the tRNA pool of the recipient organism. However, the CUB's trans-regulatory effect on the host's exogenous genes is often overlooked. To investigate this, we examined over two hundred HGT events between 94 donor and 34 recipient species, combined with fitness assays of 36 yeast strains inserted with gradient CUB and ribosome profiling of seven of these strains. We found that when exogenous genes have CUB excessively similar to that of the recipient organism, the recipient's normal expression of endogenous genes will be significantly disrupted, leading to a suppressed HGT success rate. Therefore, HGT is facilitated only when the negative impact of the exogenous gene's CUB on the host is balanced with its own translational requirements. The results of this study will allow a deeper understanding of HGT between species or between cancerous cells, and facilitate the design of heterologous expressing/translating genetic elements in biomedical applications such as mRNA vaccines.

S7-P04-Cis-Regulatory Evolution of CCNB1IP1 Driving Gradual Increase of Cortical Size and Folding in Primates

Lei Shi

Kunming Institute of Zoology, CAS

Cortical expansion and folding have gradually increased during primate evolution, reaching the highest in the human brain. Indeed, cortical folding has a concerted relationship with neocortex expansion, underlying evolution of human cognitive functions. However, molecular mechanisms underlying concerted evolution remain largely unknown. Here, using Chinese tree shrew as an outgroup of primate, we identify the Cyclin B1 Interacting Protein 1 gene CCNB1IP1 (also named HEI10) as a key regulator of cortically concerted evolution. CCNB1IP1 has acquired brain expression in primates, and is highly expressed in human neuronal progenitor cells. Following the acquisition of cis-regulatory elements, CCNB1IP1 expression has gradually increased along primate brain evolution, coinciding with the increase of cortical size and folding. Mechanistically, we elucidated that overexpression of CCNB1IP1, leading to a decrease of G1 duration, and subsequently causes an increase in basal progenitors and neurons. Furthermore, we generated a novel CCNB1IP1 knock-in mouse model and identified that CCNB1IP1 can induce cortical folding in otherwise smooth mouse brain. Moreover, the CCNB1IP1 knock-in mouse model exhibited traits associated with enhanced learning and memory abilities. Together, our study reveals how the changes in CCNB1IP1 expression may have contributed to the increase of cortical size and folding during human brain evolution.

S7-P05-The role of histone modifications and transposable elements in the epigenetic regulation of dosage after gene duplication

Thomas Wolfe

UMass Lowell

The duplication of genes has long been recognized as a substrate for the emergence of new genes, but the factors that govern fixation of paralogs soon after duplication are only partially understood. Duplication often leads to an increase in gene dosage, and with genes for which an increased dosage is harmful (i.e., triplosensitive genes), a dosage balancing mechanism needs to be present immediately after duplication if it is to evade negative selection. Previous research in vertebrates has demonstrated a potential role for epigenetic factors in allowing triplosensitive genes to increase in copy number by regulating their expression post-duplication. We have expanded this research by investigating the epigenetic landscape of duplicate genes in humans and the slime mold D. discoideum, two lineages separated by over a billion years of evolution. We found that activating histone modifications are quickly lost in duplicate genes before gradually increasing in enrichment as paralogs age. For the repressive modification H3K9me3, we found it was enriched in the youngest paralogs, and that this enrichment was likely mediated by heterochromatin spread from transposable elements. We similarly found enrichment of H3K9me3 in young human duplicates, and again found transposable elements as a potential mediator. Interestingly, while we found no significant link between enrichment for the repressive mark H3K9me3 and triplosensitivity in human paralogs, we did find a significant association between triplosensitivity and transposon proximity. Our findings suggest that transposons may contribute to the epigenetic regulatory environment associated with dosage balancing of young duplicates across broadly diverged taxa.

S7-P06-Expression variability following gene duplication facilitates gene retention

Haoran Cai

Massachusetts Institute of Technology

Genome analyses reveal that gene duplication in eukaryotes is pervasive, providing a primary source for the emergence of new genes. The mechanisms governing the retention of duplicated genes, particularly in the early stages post-duplication, remain ambiguous. Patterns of divergence between duplicated genes vary, leading to biases in the functions of genes retained following duplication. For example, genes that arise from tandem duplication tend to be involved in environmental responses. However, the mechanisms that cause such functional biases remain elusive. Here, to better understand the mechanisms and factors determining gene retention, we study a frequently overlooked molecular feature ---expression variability---as measured by within-line expression variation, termed variability. We find that, on average, genes with duplicates exhibit higher expression variability than singletons. We further find that such elevated expression variability is likely driven by the immediate increase in expression variability following small-scale duplications (SSDs) and prolonged evolutionary processes after whole-genome duplications (WGDs). These findings suggest a potential overarching mechanism that facilitates gene expression divergence for SSDs and, consequently, duplicate retention.

S7-P07-The role of primate-specific genes in the phenotypic evolution of lorises

Chunyan Chen

Shaanxi University of Chinese Medicine

Lorises, endangered primates, exhibit unique physiological and behavioral traits, including a low metabolic rate, slow movement, and hibernation. As omnivores, they primarily consume wild fruits and can metabolize toxic insects and tree exudates, providing distinct survival advantages. Research highlights the role of lineage-specific genes in phenotypic evolution. By collecting high-quality genomic and transcriptomic data from pygmy loris and other species, we identified 1,161 primate-specific genes (new genes) in the pygmy loris genome. These genes, arising mainly from gene duplication, exhibit narrow expression patterns and lower abundance compared to conserved genes. Transcriptomic analyses show that these new genes are predominantly expressed in organs linked to metabolism, detoxification, and feeding, such as the liver, heart, stomach, tongue, and spleen. A detailed study of the KLRD1 gene family, which plays a key role in immune response and apoptosis, revealed its expansion in pygmy loris through DNA duplication. Younger copies of this gene primarily express in the spleen, followed by the liver. Given the critical roles of the spleen and liver in immune function, hematopoiesis, detoxification, and metabolic regulation, the expansion of the KLRD1 gene likely contributes to loris-specific phenotypic traits. In conclusion, this study, through an analysis of genomic and transcriptomic data, underscores the significant role of primate-specific genes in the phenotypic evolution of lorises, providing new insights into their unique characteristics.

S7-P08-The Impact of Protein Abundance Modifications on Fitness Following Gene Duplication

Frederic Bedard

Laval University

Gene duplication is a mechanism that plays an important role in evolution by creating new genetic material. Different models have been suggested to explain paralog retention following a duplication event, including the positive dosage model, which proposes that gene duplication provides an immediate selective advantage by increasing the dosage of the gene product. A previous study identified several genes for which an increase in fitness is observed immediately following gene duplication, implying that the dosage of these genes is sub-optimal. However, it remains impossible to identify if the increase in protein dosage associated with the duplication of these genes represents the optimal dosage, or if the optimum would be reached following a subsequent modification of the expression level reached by gene duplication. To answer this question, we study the relationship between protein dosage and fitness for genes for which duplication leads to a selective advantage using Saccharomyces cerevisiae as a model organism. We manipulated the expression of these genes individually using synthetic promoters of variable strength and measured protein abundance associated with each duplication. We then pooled these strains together and competed them to determine their relative fitness, allowing us to build fitness functions and identify dosage optima for each gene. Our results provide a better understanding of the mechanisms by which gene duplication can immediately provide a selective advantage and also allow to assess how dosage is most likely to evolve following duplication to reach the optimal expression.

S7-P09-Building a Placenta: Roles for Syncytin-Like Retroviral Envelopes of Livebearing Fish?

Amy Ó Brolcháin

Moyne Institute of Preventitive Medicine, Trinity College Dublin

Endogenous retroviruses (ERVs) have long been implicated in mammalian placentation, primarily through the fusogenic activity of co-opted retroviral envelopes also known as syncytins. Additionally, syncytins possess immunosuppressive properties potentially necessary for tolerance of the feto-placental unit. However, whether syncytins are necessary for placental evolution remains unclear. To further investigate the roles and necessity of syncytins, we focus on non-mammalian vertebrates with placentas or placenta-like structures. Livebearing fish of the family Poeciliidae are ideal models for investigating placental evolution, having independently evolved placentas an estimated 9 times over the past ~2 million years. Syngnathids (seahorses and pipefish) are known for their distinctive gestation where embryos are carried by males, sometimes with development of a fetopaternal interface known as a "pseudoplacenta". We screened 30 poeciliid and 24 syngnathid genomic assemblies using SERVE, a custom-built retroviral envelope identification tool (https://github.com/AmyOBro/SERVE) which uses similarity search methods to identify candidate syncytins and putative orthologues. Through our analysis we identified conserved syncytin-like candidates and other conserved ERV elements that may play a role in placentation or other biological processes of livebearing fishes, as observed in mammals. We are undertaking functional characterisation of identified envelopes using immunohistochemistry and RT-qPCR on placental and other organ tissues which will help to confirm whether identified envelopes are placenta-specific and actively transcribed. Through this research, we aim to address a critical gap in the literature on retroviruses and placentation in non-mammalian vertebrate species.

S7-P10-Laccase deficiency drives adaptive calcium metabolism in cockroaches

Olga Volovych

Hebrew University of Jerusalem

The enzyme Laccase 2 (Lac2) is indispensable for insect cuticle maturation, driving key postmoult processes, combined under the term tanning. From an evolutionary perspective, the fast and calcium-independent postmoult process is one of the factors in the success of terrestrial hexapods. Although most insects have lost cuticle calcification, orthologues of the crustacean calcification machinery are still present in insect genomes. In Drosophila melanogaster, genes such as vermiform and serpentine, have acquired novel functions unrelated to their original role in calcium metabolism or binding, while the roles of others remain unknown. In this work, we performed an elaborate study of Lac2-dependent tanning in Blattella germanica, a global pest and model organism, and explored its potential interplay with calcium metabolism. First, we monitored postmoult cuticle tanning times across developmental stages, finding that the tanning duration generally increases with development, except in hatchlings, which demonstrate the longest process. The expression of BgerLac2 was analysed across three stages to identify the expression peak. RNA-interference experiments resulted in a range of phenotypes from white to brown individuals. Calcium-rich diet improved the survival of BgerLac2-deficient animals. Using scanning electron microscopy with element-mapping we detected an increased percentage of calcium in the haemolymph and cuticular surface. Further analysis revealed the induced expression of two putative orthologues of crustacean calcification genes.

S7-P11-Horizontal gene transfers in the genomes of thermoacidophilic red algae

Yongsung Lee

Sungkyunkwan university

Cyanidiophyceae, the earliest diverged class in red algae, is clearly distinguished from its sister groups by its ability to thrive in extreme environments, such as thermoacidic volcanic area rich in heavy metals. Although horizontal gene transfer (HGT) has been highlighted as a genetic source for the adaptation, most attention has been given to bacterial genes, neglecting eukaryotic sources. In addition, misidentified strains and limitations in the taxon sampling obscure a systematic overview of how HGT has influenced these thermoacidophiles. Here, improving the past limitations, we surveyed HGT genes of Cyanidiophyceae and found significant contributions from eukaryotic sources, particularly Fungi and Metazoa. For example, transporters that facilitate nutrient uptake from the environment were derived from Fungi, helping to establish a unique ecological niche for Galdieriales, one of the two clades that divide Cyanidiophyceae. Surprisingly, the majority of HGT genes in Galdieriales were not shared with the other clade, potentially implying independent acquisition of such genes. We discuss how these HGT genes confer the ability to adapt to extreme environment.

S7-P12-Evolution of Sex Pheromone Synthesis in Drosophila

Guan Jiayin

Westlake University

The sex pheromone 11-cis-vaccenyl acetate (cVA) plays a crucial role in sexual and social behaviors across the melanogaster group species. However, the biosynthesis pathway of cVA has not been fully elucidated. In this study, we identified two key enzymes, a fatty acyl-CoA reductase and an acyltransferase, involved in the final two steps of cVA biosynthesis, thereby providing the last pieces of the puzzle for a complete understanding of cVA synthesis. RNAi-mediated knockdown of these two genes in the male reproductive system of Drosophlia melanogaster resulted in a significant reduction in cVA levels, as detected by GC-MS. Functional predictions suggested that the fatty acyl-CoA reductase catalyzes the reduction of (11Z)-octadecenoyl-CoA to cis-vaccenyl alcohol, while the acyltransferase converts cis-vaccenyl alcohol to cVA in the presence of coenzyme A. The enzymatic activity of the acyltransferase was further confirmed through in vitro assays. We compared the expression of acyltransferase in five Drosophila species, including two cVA-producing species and three non-cVA-producing species, and found that its expression in the male reproductive system was significantly lower in non-cVA-producing species compared to cVA-producing species. The acyltransferase was conserved in all cVA-producing Drosophila species but was lost in some non-cVA-producing species. The loss of cVA production in certain species was associated either with the deletion of the acyltransferase gene or a substantial reduction in its expression the male reproductive system.

S7-P13-A versatile pipeline to identify convergently lost ancestral conserved fragments associated with convergent evolution of vocal learning

Ying Zhen

Westlake University

Exploring molecular convergence in convergently evolved lineages provides valuable insights into the shared genetic basis of converged phenotypes. However, current studies have mainly focused on protein-coding regions, with limited exploration of noncoding regions, which contain regulatory elements and play an important role in gene expression and trait evolution. Our study focused on investigating molecular mechanism underlying convergent evolution of avian vocal learning ability, a unique ability to modify and acquire new sounds through imitation. We developed a whole-genome-alignment-free method for detecting Convergent Lost Ancestral Conserved fragments (CLACs) across multiple avian lineages. Our analysis uncovered 2,711 such CLACs, predominantly located in noncoding genomic regions. Genes associated to CLACs exhibit significant enrichment patterns in neurological pathways, such as glutamate receptor signaling pathway and axon guidance pathway. Notably, several CLAC-associated genes, including ROBO family, SLIT2, GRIN1 and GRIN2B, have known functions in speech and vocal learning ability. Additionally, several motifs, which assembled binding motifs of transcriptional factors involved in neurogenesis and gene expression regulation in brain, are found significantly enriched in noncoding CLACs. Furthermore, we pinpointed 19 candidate genes that harbor CLACs in both human and multiple avian vocal learning lineages, suggesting their potential contribution to the independent evolution of vocal learning in both birds and humans.

S7-P14-RNA-seq reveals genetic basis of structural coloration in the Neotropical butterfly Parides sesostris (Papilionidae)

Leo Camino Smithsonian

Tropical Research Institute

In the last two decades, research in Lepidoptera has largely focused on identifying key switch genes involved in wing-patterning and coloration. However, a rarer form of color production, known as structural coloration, arises from the interaction between light and nanostructures. Despite its deep evolutionary origins in butterfly wing patterning, the genes underlying this process remain unknown. The Neotropical Parides genus is ideal for exploring the ecological and genetic basis of structural coloration due to two key traits: aposematic mimicry, present across the genus in both sexes due to sexual dimorphism, and structural green coloration restricted to males. In Parides sesostris, this hue is produced through gyroid-like photonic crystals. The genetic mechanisms behind gyroid formation remain poorly understood despite its role in mate-choice interactions in this species. To identify candidate genes involved, we reared 25 males and 25 females at 25 °C and 55% relative humidity, tracking pupation timing with time-lapse cameras. Forewings were dissected at five post-pupation timepoints (2, 4, 6, 8, and 10 days), with five biological replicates per sex per stage (n = 50). Total mRNA was extracted and sequenced via bulk RNA-seq, with females serving as non-iridescent controls. Differential expression analysis (DESeq2, FDR < 0.05, |log2FC| > 2) revealed male-biased expression of genes associated with chitin formation, with no significant expression in females. Future studies using CRISPR, RNA interference (RNAi), single-cell genomics, and imaging techniques will further clarify their role in gyroid development and the evolution of structural coloration.

S7-P15-Genome-wide association study reveals the genetic basis for altered body conformation of game chickens

Zheng Li

1. Uppsala University, Sweden 2. Beijing Normal University, China

Cockfighting has a history that spans centuries across diverse cultures, playing a significant role in religious, social, and political contexts. Game chickens, characterized by their unique traits such as large body size, altered body conformation, enhanced muscularity, and aggressive behavior, exemplify the remarkable phenotypic and behavioral diversity resulting from selective breeding, making them an interesting model to study the genotype-phenotype relationships. To investigate the genetic basis underlying these complex traits, we conducted a genome-wide association study (GWAS) using a cohort of 371 individuals from 37 different game chicken breeds and 1,230 individuals from 210 non-game chicken breeds. Our analysis identified a major 600-kb association block, with most single nucleotide polymorphisms (SNPs) showing consistently significant p-values below 10^{-25} across the interval. The most significant signal in this region had a p-value of 3.01×10^{-150}, suggesting a potential hotspot for association. Population genetic analysis further supported the findings, revealing an elevated fixation index (FST) and reduced nucleotide diversity across the block. Given the size and distinct genetic characteristics of this block, it is possible that it harbors a large-scale structural variation or multiple causal mutations. Future investigations leveraging PacBio long-read sequencing data will aim to identify and characterize such structural variants, shedding further light on their roles in shaping the unique traits of game chickens. These results shed light on the genetic underpinnings of key traits in game chickens, providing promising targets for future functional studies and breeding strategies.

Symposium 9: Origin and evolution of cell types

S9-P01-Generation of bat iPSC cell lines across multiple species reveals high differentiation potential

Xiaohan Lyu

University of Edinburgh

Bats are highly diverse mammals, more genetically similar to human than rodents, and exhibiting extended longevity. Yet they also appear naturally resistant to neurodegenerative diseases such as Alzheimer's disease. Any protective mechanisms have not been studied so far due to limitations in obtaining both live and post-mortem samples. We optimized the general protocols for reprogramming somatic cells to a pluripotent stem cell-like state, induced pluripotent stem cell (iPSCs), for several bat cell lines, including Rhinolophus sinicus (Chinese horseshoe bat) and Pteropus alecto. The generated bat iPSCs were positive of alkaline phosphatase activity and expression of marker genes and maintained in the iPS state over time. Sequencing data from ATAC-seq revealed high accessibility of nuclear DNA, suggesting a high differentiation potential. Using our novel protocol of secreted bat cytokines, our biPSCs were differentiated into macrophage and neuron-like cells. Immunofluorescence indicates positive expression of cell type specific proteins and through RNA sequencing, we were able to characterize the expression profiles of iPSC-derived cells and compared across phylogenies. This overcomes most challenges and provides valuable tools for in vitro cellular research, identifying key differences associated different evolutionary paths. This work provides insight into the cellular functions of bats, enabling further studies on mechanisms to potentially translate back to human biology.

S9-P02-3D single-cell transcriptome of lamprey identifies cellular and molecular innovations during the origin of the vertebrate brain

Haixu Wu

Kunming Institute of Zoology, Chinese Academy of Sciences

The emergence of brain structures is one of the significant innovations in the origin of vertebrates. Lampreys, as the representative lineage of agnathans, possess relatively primitive brain structures and are considered an ideal model for studying the origin of the vertebrate brain. We have combined single-cell transcriptomics with spatial transcriptomics to construct a 3D transcriptomic map of the lamprey brain (Lampetra reissneri). We identified 209 cellular clusters across 14 major brain regions, including different neurons, glial cells, ependymal cells and so on. We observed conserved gene expression patterns across the anterior-posterior axis in the lamprey bain, as well as a process of neuronal specialization towards distinct excitatory and inhibitory functions. Through cross-species comparisons of nine representative vertebrate species from different classes, we further constructed the evolutionary changes of different cell types, especially neurons, and for the first time comprehensively revealed the systematic evolutionary process of neurons in vertebrates. Furthermore, we have attempted to elucidate the origins of the cerebellum and hippocampal structures, and we found that in lampreys, only some conserved cell types and structural features are present, representing a proto-type of the vertebrate brain

S9-P03-Unveiling cell-type-specific mode of evolution in comparative single-cell expression data

Tian Qin

IOZ

While methodology for determining the mode of evolution in coding sequences has been well established, evaluation of adaptation events in emerging types of phenotype data needs further development. Here we propose an analysis framework (expression variance decomposition, EVaDe) for comparative single-cell expression data based on phenotypic evolution theory. In EVaDe, we used two strategies to identify genes exhibiting large between-taxon expression divergence and small within-cell-type expression noise in certain cell types, attributing this pattern to putative adaptive evolution. In primate prefrontal cortex, we found such human-specific key genes with neurodevelopment-related functions, associated with specific neuron cell types, while most other genes exhibited neutral evolution patterns. Reassuringly, the key genes were significantly associated with the rapidly evolving conserved non-coding elements. Additional case analysis on the naked mole-rat revealed innate-immunity-related genes and cell types with putative expression adaptation. Overall, the EVaDe framework may effectively probe adaptive evolution mode in single-cell expression data.

S9-P04-A Model for the Evolution and Deployment of Functional Gene Expression Programs Across Tissues

Mark Kim

Cornell University

The complexity of multicellular life arises from the coordinated roles of specialized cells. Advances in single-cell RNA sequencing (scRNA-seq) and dimensionality-reduction methods have revealed that quasi-discrete gene expression programs (GEPs) underlie tissue-specific cellular functions, linking sub-cellular gene regulation and organismal functions. Evolutionary novelty often arises through lineage-specific subfunctionalization of GEPs, where ancestral multifunctional cells partition roles into specialized counterparts. Novel tissue or organ-level functions emerge as pre-existing GEPs are redeployed spatiotemporally, accompanied by new cooperative interactions.

While these advances have made it possible to measure and annotate GEPs in unprecedented detail, a major gap exists in integrating these insights into evolutionary theories. Specifically, how evolutionary forces shape GEP partitioning and deployment across functional tissues is poorly understood. To address this, we developed a hierarchical model of GEP evolution and deployment, built on a model of the evolution of dynamic modularity and trade-offs, to study the evolution of GEPs and their usage in organisms with multiple functional tissues.

Using this model, we study how the usage and deployment of GEPs across functional tissues evolve over time. Our results suggest stronger partitioning of the GEP usage is favored when the number of GEPs is limited, while higher numbers of GEPs allow greater fitness gains through the expansion of the accessible phenotypic space. By systematically exploring evolutionary parameters, we investigate how selection at the organismal level — mediated through tissue-level performance — shapes the partitioning and overlap of sub-cellular gene expression programs, which serve as the raw material for functional adaptation and innovation.

S9-P05-Single cell sequencing reveals novel bacterial-induced pathways of eukaryotic differentiation

Shelbi Russell

University of California Santa Cruz

Animals associate with diverse bacteria that manipulate host reproduction to promote bacterial survival and transmission. The intracellular alphaproteobacterium Wolbachia pipientis is a master of host manipulation, influencing offspring viability, host development, and viral resistance. Field applications exploit these traits to suppress mosquito populations and limit viral pathogen spread. However, the molecular mechanisms underlying Wolbachia's cell- and tissue-specific effects remain poorly understood. Most gene expression studies rely on bulk RNA, averaging signals across diverse cell types, which obscures infection-driven transcriptional changes. Our lab previously augmented the Wolbachia-induced transcriptomic signal by knocking down host genes that interact with Wolbachia in vivo and by concentrating single host cell types in culture. These studies revealed that Wolbachia has significant impacts on host gene expression, which are underestimated in bulk datasets. Here, we use single-cell RNA sequencing to resolve Wolbachia-host interactions with high resolution. We generate single-cell datasets for an immortal Drosophila melanogaster cell line with and without wMel infection using the 10x Genomics Chromium platform and the new, low-cost PIP-seq protocol. Then, we apply the PIP-seq approach to mutant D. melanogaster ovaries with and without wMel infection. Through unsupervised clustering, pseudotime network inference, and marker gene analysis, we uncover how Wolbachia reprograms host somatic cells in vitro to evade immunity, establish infection, and acquire nutrients. In vivo, we reveal how Wolbachia rewires dysfunctional developmental pathways to restore fertility. These findings expose Wolbachia's hidden role in eukaryotic cell and developmental biology and will inform targeted strategies to harness Wolbachia's phenotypic effects for applied vector control.

S9-P06-Exploring Cell Types in Brown Algae Using Single-Cell RNA Sequencing

Liping Wang

MPI for Biology Tuebingen

The multiplicity of cell types comprising extant multicellular eukaryotes begs the question as to how this diversity emerged from a single cell (unicellular) ancestor, and how cell type identities evolve over time. The brown algae independently evolved multicellularity becoming the third most complex multicellular lineage after animals and land plants, and are therefore important comparative models to understand evolution of development. We employed single-cell and tissue-specific RNA sequencing across multiple brown algal species to investigate cell type and tissue origin and evolution. The comprehensive gene expression atlas is enabling the molecular characterization of cell and tissue types and allowing us to gain insights into how morphological diversity in the brown algae has evolved from a unicellular ancestor. Our study establishes the first large-scale molecular framework to explore cell type origin and evolution in brown algae, highlighting their unique evolutionary trajectory as a non-model eukaryotic lineage.

S9-P07-The molecular and cellular evolution of mammalian livers

Xuefei Yuan

ZMBH, Heidelberg University

The liver, essential for metabolism, detoxification, pathogen clearance, and homeostasis, has played a key role in mammals' adaptation to diverse and extreme environments. Previous studies suggest rapid molecular evolution in the mammalian liver, marked by increased use of evolutionarily young or positively selected genes. However, the molecular, cellular, and regulatory mechanisms underlying these adaptive changes remain largely unexplored. To address this, we generated single-nuclei gene expression and chromatin accessibility data for 16 mammals and one bird (chicken) as an evolutionary outgroup. Combined with spatial transcriptomics, we found that liver zonation, gradient hepatic gene expression within the lobule, is conserved across all major mammalian lineages but absent in many other vertebrates. Liver zonation is associated with greater hepatocyte heterogeneity, suggesting increased cellular complexity. Despite its conserved structure, zonated genes exhibit rapid evolutionary turnover across mammalian species. Consistently, we found hepatocytes to be among the fastest-evolving liver cell types, likely due to low pleiotropic constraints on their transcriptome. Given these broader evolutionary patterns, we next examined how liver biology has diverged specifically in humans. Using our dataset from four other great apes, we identified human-specific cis-regulatory elements and genes with altered cell-type expression in the human lineage. These human-specific changes, particularly associated with lipid metabolism genes, align with known differences in energy expenditure and fat storage. Overall, our study uncovers key cellular mechanisms and regulatory changes underlying mammalian liver evolution and human-specific metabolic traits.

S9-P08-Evolution of Transcription Factor Regulatory Networks in Vertebrate Immune Cells Using Single-Cell Multi-Omics

Wenxu Ren

Sun Yat-sen University, School of Life Sciences.

Vertebrate immune cells exhibit remarkable diversity and play an essential role in responding to external pathogens and endogenous diseases. Transcription factor regulatory networks (TFRNs) are key determinants of immune cell function and identity by regulating gene expression. However, TFRNs evolution in vertebrate immune cells remain underexplored. In this study, we systematically investigate the evolution of TFRNs in vertebrate immune cells using single-cell multi-omics approaches, including single-cell RNA sequencing and single-cell chromatin accessibility sequencing. We identified vertebrate immune cell types at key phylogenetic nodes and characterized the essential TFRNs within vertebrate immune cells. Preliminary results indicate that immune cell types exhibit conserved TFRNs across species, suggesting that these networks may serve as core components of immune cell transcriptional regulation. Additionally, we identified species-specific changes in TFRNs, which will help reveal the relationship between species-specific regulatory network variations and immune functional phenotypes. These findings provide new insights into immune cell function and evolution.

S9-P09-Comparative Spatiotemporal Transcriptomics Unveils Conserved Lignification Trajectories and Evolutionary Signals During Nut Shell Development in Juglans

Zhenzhen Liu

Beijing Normal University

Nuts are often hypothesized to have evolved from the selection towards seeds protection and dispersal. However, the molecular mechanisms underlying the nut shell development is unclear, and hence the evolutionary process of the nut shell remain poorly understood. In this study, we utilized the spatial transcriptome (ST) and Single Nuclei RNA Sequencing (snRNA-seq) techniques to comprehensively characterizing the molecular pathways driving the nut shell development from a comparative analysis of two congeneric walnut species, Juglans mandshurica and Juglans regia. Anatomical analysis suggests that shell lignification commences from the outer layer of the endocarp and gradually extends inwards as signal molecules permeate. Differentially expressed genes identified between endocarp and outer-tissue cells across two species demonstrated a conserved progress of the shell lignification. Genesets that were involved in the initiation and regulation of lignin-synthesis pathways were further revealed using pseudotime trajectory and intercellular communication analysis. Finally, the population genetics study across the identified gene sets demonstrated rapid evolutionary changes in these related pathways. Our study draw a comprehensive transcriptome atlas and lignification trajectories during the nut shell development in walnuts and provided a novel insight into the underneath cell evolution.

S9-P10-Unicellular, Filamentous and Multicellular Cyanobacterial Genome Evolution Through Geological Time

Lingyun Yang

University of Bristol

Cyanobacteria, which are one of the oldest known life forms on Earth, were also the first organisms to perform oxygenic photosynthesis. Throughout the past 3.0 Ga years, cyanobacteria have undergone major evolutionary transitions, including the shift from oceanic benthic to planktonic habitats, as well as the origins of multicellularity, amongst others. These transitions have had profound implications for Earth's biogeochemical cycles and have been critical drivers of biological innovation that shaped the early Earth environment. By studying the evolution of cyanobacterial genomes, we have gained insights into the origin and early diversification of cyanobacteria over geological time. Multicellularity offered considerable advantages for exploring new ecological niches, thus facilitating the diversification of new lineages. Phylogenetic studies have shown that multicellularity evolved early within Cyanobacteria, and there have been several instances where multicellular linages revert to unicellular. With the increasing quality and quantity of genome-wide data, it is now possible to use comparative genomics techniques to gain insights into the origins and evolution of different morphological cyanobacterial taxa. In this study, we implement phylogenomics and evolutionary genomics pipelines to study hundred complete genomes. Our aim was to identify homologous genes that emerged and/or were lost at key times in cyanobacteria evolution, with a specific focus on transitions from multicellular to unicellular events. Our findings highlight processes behind genetic mechanisms during cyanobacterial morphological transitions and emphasise the pivotal role of conserved novel genes and lost gene families in the process of cyanobacterial morphological diversification.

S9-P11-Evolution of von Economo Neurons in the Anterior Cingulate Cortex Across Primates

Levin Alleweldt

Freie Universität Berlin

Dissecting the unique molecular features of neuronal cell types across species provides fundamental insights into brain evolution. Von Economo neurons represent a particularly compelling target for investigation, given their association with human-specific cognitive and behavioural disorders, such as the behavioural variant of frontotemporal dementia, schizophrenia, and autism spectrum disorder. These neurons account for most of the layer V extratelencephalic (L5 ET) projection neuron population within the anterior cingulate cortex (ACC). We aimed to characterize the molecular evolution of L5 ET neurons across primates. Utilizing single-nucleus RNA sequencing data from the ACC of marmosets, macaques, chimpanzees, and humans, we observed gene expression signatures that distinguish L5 ET neurons from other neurons within each species. We identified genes specifically upregulated in human L5 ET neurons, which have been previously implicated in cognitive and behavioural disorders. In agreement with recent studies, this analysis revealed a notable enrichment of genes involved in extracellular organization. However, inclusion of more primates allowed us to further resolve great ape- and human-specific upregulated genes in L5 ET neurons. Several collagen genes were identified as great ape specific, while genes involved in cell-cell-adhesion (some cadherin and catenin genes) exhibited human specific upregulation. Our findings underscore the importance of membrane and extracellular matrix constituents in these morphologically distinct neurons. Moreover, they elucidate the molecular divergence of these neurons during primate evolution. Further investigation into the functional roles of the identified genes is critical for understanding the evolution of complex brain functions and their dysregulation in human neurological disorders.

Symposium 10: Pangenome graphs and their applications in biodiversity genomics

S10-P01-White lupin′s genomic diversity and domestication history through pangenome graphs

Matteo Tommaso Ungaro

The University of Ferrara

Domestication is a human-mediated process which has been promoted by the transition from a hunter-gatherer subsistence economy to farmer-oriented one. In the white lupin (Lupinus albus) such process was potentially started in the Old World by ancient Greeks; from then onwards, gradually, different varieties for this species have been locally selected across the Mediterranean basin. However, it is still unclear where and when exactly the domestication process begun but also whether it has been characterized by a single or multiple events. Additionally, it is well-known that genomic structural variants might play a major role in domestication contributing to local adaptations to diverse environmental conditions in many animals and plants. Yet, traditional methodologies using short reads and linear reference genomes systematically failed to capture the full extent of this variants' class and, consequently, their impact on domestication is often neglected.

Recent methodological and algorithmic advances allow to thread phased genome assemblies into graph-based data structures, referred to as pangenome graphs. Graph space is a better model for representing the genetic diversity of a species or a population. Therefore, to address the former aspects related to the observed genomic diversity of the white lupin we built a pangenome graph including 16 chromosome-scale wild and landrace accessions. Such data structure was finally combined with more than 300 short reads re-sequencing whole genome experiments to describe the observed genomic diversity and to obtain insights into the domestication history for the species as well as critical regions of the genome targeted by artificial selection.

S10-P02-Pangenome applications with de novo diploid genome assemblies in longitudinal cancer genome evolution

Yue Hao

Translational Genomics Research Institute

Glioma is an aggressive cancer of the central nervous system characterized by extensive aneuploidy and structural variation (SV). We engineered normal human astrocytes (NHA) to express HPV18 E6 and E7 oncoproteins which trigger telomere dysfunction (TD) via suppression of the p53/pRb pathways. For multiple timepoints over a period of 6 months following transduction, TD as well as accumulation of SVs and aneuploidy in NHA-E6E7 cells were measured using long read and Hi-C sequencing. Diploid genome assembly was performed using Verkko2. In primary pre-transduction NHA cells, we achieved a complete, fully-phased assembly size on par in quality with HG002. Conversely, NHA-E6E7 cells at later time points exhibited chromosomal instability resulting in more fragmented assemblies. Additional challenges emerged as current SV callers do not support a diploid reference, limiting the phased inference of SVs from our aneuploid genomes. To address these limitations in traditional variant calling approaches, we used the PanGenome Graph Builder (pggb) to construct a pangenome reference from our three diploid assemblies, including the CHM13v2.0 reference to provide a unified coordinate system. We then used the vg toolkit to correctly infer haplotype-specific SVs at each time point. We found an elevated number of SVs in NHA-E6E7 cells compared to the primary NHA cells. Our results demonstrate that personalized pangenome graph assembly followed by graph-based variant detection is a feasible and highly appropriate approach for the somatic analysis of cancer genomes.

S10-P03-A universal and constant rate of gene content change traces pangenome flux to LUCA

Katharina Trost

Heinrich Heine University Düsseldorf

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 5655 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. The long-term average rate of gene content flux is remarkably constant across all higher prokaryotic taxa sampled, whereby the size of the accessory genome—the proportion of the genome harboring gene content difference for genome pairs — varies across taxa. The proportion of species-level accessory genes per genome, varies from 0% (Chlamydia) to 30% – 33% (Alphaproteobacteria, Gammaproteobacteria, and Clostridia). A clock-like rate of gene content change across all prokaryotic taxa sampled 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.

S10-P04-From pangenomes to pangenome graphs: resource curation, methods development, and biological discovery

Jia-Xing Yue

State Key Laboratory of Oncology in South China, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, China

With the increasing availability of high-quality genome assemblies, pangenomes and pangenome graphs are emerging as the new paradigm in the genomics field. High-quality resources and novel methodologies are needed to better explore the power of pangenomics. Recently, we assembled and expanded the Saccharomyces cerevisiae reference assembly panel (ScRAP), a telomere-to-telomere (T2T) pangenome resource for the classic budding yeast model, which serves an invaluable testing ground for novel methodology development and biological discovery in both linear- and graph-based pangenomic contexts. With the pangenome and pangenome graph resources from ScRAP and the Human Pangenome Reference Consortium (HPRC), we further developed an interpretative visualization framework, named VRPG, for pangenome graphs by linear-reference-based coordinate projection and annotation integration, which enables a highly synchronized and intuitive visualization between pangenome graphs and the conventional linear-reference-based functional annotations. Building upon this work, we are currently developing a novel methodology for accurately characterizing structurally divergent genomic regions based on both linear- and graph-based pangenome datasets. By further applying these newly developed methods to the ScRAP and HPRC pangenome graphs, we are approaching an enhanced understanding of the structural genomic diversity of global yeast and human populations in the context of pangenome graphs.

S10-P05-Accurate and scalable detection of structural variants from linear- and graph-based pangenomes

Zepu Miao

Sun Yat-sen University Cancer Center

Structural variants (SVs) can lead to significant phenotypic changes, with important implications in biodiversity and disease susceptibility. The accurate and characterization detection of SVs , however, can be challenging, especially regarding their exact breakpoints. It is even more challenging when it comes to analyzing the SV landscape at a population-scale, during which the conventional SV merging algorithms tend to ignore the haplotype-specific differences. The recently emerged pangenome and pangenome graphs can theoretically encapsulate all genomic variants from the entire population or species into a compact representation, which offers a promising way to tackle this challenge. Here we present a pair of tools, pairwiseSV and graphSV, for accurate characterization of SVs from linear-based and graph-based pangenomes respectively. We validated their sensitivity and specificity for SV identification and breakpoint resolution using high-quality pangenome and pangenome graph datasets from both yeast and human. We anticipate pairwiseSV and graphSV to empower the pangenome-based studies in the field, helping us to better understand the population-wide SV diversity and their functional implications.

Symposium 29: Open symposium session A

S29-P01-Exploring the drivers of gut microbiome diversification in Anolis lizards

Yanileth Lopez Tacoaman

Smithsonian Tropical Research Institute

Animals maintain complex associations with diverse microbiota residing in their guts. The composition of the gut microbiome can be influenced by various factors, including the environment and the host's genetic background. In several taxonomic groups, studies have shown a positive correlation between host genetic and microbiome divergence among species, a phenomenon called phylosymbiosis. While our understanding of the relative roles of host species and environments in microbial community assembly is growing, it remains limited in reptiles. Anolis lizards are an excellent model for studying the factors that drive microbiome diversification as anoles often live-in diverse communities of related species that occupy different microhabitats and consume different prey items. To achieve this, we studied a community of anoles from a lowland tropical rainforest site in the mainland and islands in central Panama to determine if microhabitat or species identity has a stronger effect on gut microbiome composition. We collected fecal samples from six species in the mainland and from two species on an experimental island in the Panama Canal across 2 years and used 16S rRNA gene sequencing to inventory their gut bacterial communities. We use these data to understand the drivers of diversification in Anolis gut microbiomes, and their stability across years. Our work will give insight into how the microbiome can influence resilience to rapid environmental change in vulnerable species.

S29-P02-The evolution of Kallima butterflies

Shuting Wang

Peking University

Kallima butterflies are famous for their leaf-mimic wing pattern. However, the evolution scenario of Kallima butterflies remains elusive. Here, based on advanced sequencing and analysis methods, we explored the evolution of Kallima butterflies, including clarifying the phylogeny, analyzing demography, and detecting introgression signals between species. Our results suggest that Kallima butterflies likely differentiated in the eastern Himalayas and subsequently dispersed to areas that are now islands. Moreover, we find that Kallima butterflies hybridized in the early stage of speciation and an introgressed gene may have contributed to the adaptation of some populations. In summary, our study provides insights into the evolution of Kallima butterflies with genomic analysis.

S29-P03-Phylogenomics and Population Genetics of a Widespread Polytypic Bird Species Reveals Complex Evolutionary and Demographic History

Huy Chi Truong

Cornell University

Large-scale landscape changes and demographic processes can have significant impacts on population structure and the evolutionary history of a species. The song sparrow, Melospiza melodia, is a widespread North American breeding bird with significant plumage and morphological diversity in environments gradients across North America. However, despite them being very polytypic and ubiquitous, patterns of differentiation are difficult to compare as previous molecular phylogenomic work has focused on individual mitochondrial markers with incomplete sampling. Here we present the first subspecies-level phylogeny for the species using variable sites sampled from single-nucleotide polymorphisms (SNPs) in whole-genome resequencing data in 354 individuals. We also constructed a time-calibrated mitochondrial tree for the species to corroborate estimated divergence times and detect any mitonuclear discordance. Based on this data, we also report the population structure of each subspecies and individual admixture using hierarchical clustering algorithms, as well as reconstructing hypothesized demographic histories for the identified major song sparrow clades. Our results indicate a complex and dynamic relationship between the demographic history of song sparrows and the present landscape of geographic structuring in the species related to both past and present environmental barriers to gene flow, as explored through effective migration surfaces and demographic modelling. Furthermore, the homogeneity of the mitochondrial gene trees compared to the structure of the nuclear loci supports the hypothesis of rapid contraction and expansion in the population as a result of Pleistocene glacial refugia, providing crucial insight into the processes that generated currently observed patterns of biodiversity and genetics.

S29-P04-Covariance and punctuation characterize the macroevolution of chemodiversity in gymnosperm volatile organic compound system

Ziyue Xu

Department of Ocean Science, The Hong Kong University of Science and Technology

Organisms have evolved complex metabolic systems capable of producing diverse metabolites, resulting in remarkable chemodiversity. However, the structure and the evolutionary trajectories of chemodiversity remain poorly understood. In this study we conduct a macroevolutionary analysis by analyzing a comprehensive dataset of gymnosperm volatile organic compounds (VOC) within a phylogenetic comparative framework. We find that the principal components of the gymnosperm VOC system exhibit significant covariances, with its chemodiversity represented by a phylomorphospace with four main axes, indicating the chemodiversity is convergence towards a limited set of major trait combinations. Additionally, each major single VOC's evolutionary tempo and mode follow a pattern of "punctuated gradualism". Furthermore, the formulations of the gymnosperm VOC system were already established in ancestral lineages prior to the Carboniferous period. Our findings shed light on the macroevolution of the gymnosperm VOC system at the phenotypic level and provide a framework for exploring the macroevolution of chemodiversity.

S29-P05-Single-cell multiome analyses of lungfish brain uncover cellular and molecular adaptation during water-to-land transition

Xiaoyu Meng

Kunming Insititute of Zoology, CAS

The water-to-land transition represents a pivotal evolutionary shift that facilitated the radiation of vertebrates and the subsequent flourishing of tetrapods. The brain is a crucial anatomical innovation that has enabled vertebrates to adapt to new habitats, particularly the terrestrial environment. In this study, we selected the African lungfish (Protopterus annectens), a member of the sarcopterygians and the closest living relative to tetrapods, as a model to uncover clues regarding brain evolution. We first conducted a high-quality genomic annotation based on long-read transcripts of the lungfish brain. We collected over 80,000 cells and constructed a comprehensive cell-type atlas of the entire lungfish brain using single-nucleus RNA sequencing (snRNA-seq) and single-nucleus ATAC-seq (snATAC-seq). Through analysis and integration with other datasets, we characterized the cytoarchitecture and molecular features within sarcopterygians. Moreover, by combining our findings with high-definition spatial-transcriptomic data, we identified a conserved lungfish pallium structure comparable to that of tetrapods. Collectively, these results provide valuable insights into the homology and diversity of brain evolution during animal terrestrialization.

S29-P06-Novel random forest-based machine learning method for detecting subtle nuclear chromatin changes during programmed cell death

Igor Pantić

University of Belgrade, Faculty of Medicine, Serbia

In recent years, significant advances have been made in the field of Artificial Intelligence (AI) and its use for detecting physiological and pathological phenomena in cells. Here, we present our work on the utilization of supervised machine learning for detecting discrete alterations in chromatin organization and distribution associated with cell damage and the early stages of programmed cell death. A model based on the random forest algorithm was developed using nuclear regions of interest (ROIs) from digital micrographs of Saccharomyces cerevisiae cells previously exposed to sublethal doses of proapoptotic chemical agents. For quantitative analysis of chromatin structure during conventional microscopy, we calculated textural features of the Gray Level Run Length Matrix (GLRM), such as short-run emphasis, long-run emphasis, and run-length non-uniformity. For each nuclear ROI, we also calculated coefficient energies of the discrete wavelet transform. These parameters were used as inputs for training the random forest model to distinguish damaged from intact cells. The model demonstrated solid performance, achieving a classification accuracy above 80%. The model, along with freely accessible Python code, could be used to further enhance contemporary AI-based sensors in cell biology.

Acknowledgements: The work was supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia grant 451-03-66/2024-03/200110.

S29-P07-Reconstruction the evolutionary history of the European snow vole based on ancient DNA and radiocarbon dating in the context of climate changes in the late Pleistocene and Holocene

Aleksandra Zeromska

Laboratory of Paleogenetics and Conservation Genetics, Centre of New Technologies, University of Warsaw

A great help in determining the impact of future climate and environmental changes on populations can be provided by studies of past analogues, so ancient DNA (aDNA) isolated from fossil specimens is an invaluable source of genetic information about the past, which is wiped out in present DNA. Particular species can response differently to climate changes. Thus, generalisations of the climate influence can be incorrect and gathering data on individual species is essential. One of interesting species in this context is the European snow vole Chionomys nivalis, a rodent, which is a glacial relict with a fragmented distribution in mountainous parts of south-eastern Europe as well as south-western Asia. In order to reconstruct the most likely phylogenetic relations in Chionomys nivalis populations since the Pleistocene we sequenced 96 contemporaries and 60 ancient mtDNA. Radiocarbon dates to calibrate the molecular clock were obtained from 10 fossil specimens. Based on fossil records, we also applied an approach using the divergence time of Chionomys nivalis from its closely related species Chionomys gut, Chionomys roberty. Our research shows that the earliest diverged Ch. Nivalis lineages include samples from Central Asia. The Asiatic populations gave rise to the European lineage and the Middle East lineage. In the European lineage we can distinguish 4 distinct genetic groups correlated with geographical location. Intraspecific morphological variation in Ch. nivalis coincides with its phylogeographic structure based on genetics. It indicates a concordance between the morphological and molecular evolution, which is not common in other species.

S29-P08-The mystery of Prdm9-independent meiotic recombination pathways in Canids

Jasmin Cichy

Max-Planck-Institute for Evolutionary Biology

Successful recombination is an essential driving force of evolution, being crucial for sexually reproducing organisms to generate functional germ cells and ensure genetic variation in the offspring. In this fine-tuned, yet enigmatic process, one highly conserved gene in particular was found indispensable in most metazoan taxa; Prdm9.

Lack of functional Prdm9 in almost all mammalian species typically leads to early meiotic arrest and infertility, including humans. However, Canidae evolved to lack functional Prdm9 millions of years ago without suffering these cytotoxic consequences.

What mechanism saves canine germ cells from meiotic arrest?

Rare findings of Prdm9-dependent organisms being able to cope with Prdm9-deficiency, including one human female, indicate the presence of an even more ancient and conserved pathway than Prdm9, which had an evolutionary advantage in Canidae. Consequently, this pathway could be an answer to infertility cases originating in Prdm9-deficiency even in humans.

In this project, I analyze the Prdm9-independent recombination landscape of the domestic dog at the fine-scale. Combining high-resolution sperm-typing techniques with cutting-edge Nanopore sequencing tools, I identify de novo hotspot regions and their characteristics, including epigenetic modifications, at the nucleotide level. By applying Fluorescence-Activated-Cell-Sorting (FACS) and immunostaining, I provide the first knowledge of PRDM9-independent stage-specific meiotic gene expression and intragenetic consequences of PRDM9-loss visually and quantitatively.

With these results, we can start understanding what mechanism took over in canines that is able to overcome cytotoxic consequences of lacking Prdm9 - potentially even laying the cornerstone in understanding Prdm9-absence induced meiotic arrest and infertility in humans.

S29-P09-Support vector machine model utilizing fractal features for detecting ethanol-induced alterations in cell nuclei

Igor Pantić

University of Belgrade, Faculty of Medicine, Serbia

Fractal analysis is a contemporary method that can in some cases be used for quantification of structural complexity in biological systems. As demonstrated by recent studies, fractal features such as fractal dimension and lacunarity can be applied for detection of discrete alterations of nuclear texture after induction of cell damage by various toxins. In this work we present our recent findings related to the development of innovative support vector machine model that uses nuclear fractal dimension and lacunarity as inputs for prediction of existance of ethanol-induced damage in yeast cells. The model was developped in freely-available libraries and platforms for Python programming language using a dataset of nuclear fractal data for treated and control cells. It included 5-fold cross-validation, hyperparameter optimzation and rigorous performance evaluation. The model showd solid performance in separation of damaged from intact cells with adequate classification accuracy, area under the receiver operating characteristic curve and F1 score. The model shows potential for future integration with other computer-aided AI systems for identification of ethanol-induced damage in in-vitro conditions.

Ancknowledgements: This work was supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia grant 451-03-66/2024-03/200110.

S29-P10-Detecting cell damage with AI: Gradient boosting tree algorithm combined with gray-level co-occurrence matrix

Jovana Paunović Pantić

University of Belgrade, Faculty of Medicine

In this work, we employ gray-level co-occurrence matrix (GLCM) and other textural features of cell nuclei to create a sophisticated machine learning system for detecting subtle structural changes associated with cell damage. Experiments were conducted on Saccharomyces cerevisiae cells treated with sublethal doses of hydrogen peroxide. Digital micrographs of cells and nuclear regions of interest were analyzed by calculating GLCM features, including angular second moment, inverse difference moment, textural correlation, and sum variance. These GLCM features, along with other textural parameters, served as inputs for a gradient boosting supervised machine learning model created using contemporary programming methods. The model achieved an accuracy of over 75% in identifying damaged cells, which were otherwise undetectable through conventional microscopic examination. The advantages of gradient boosting over alternative models include its ability to handle non-linear relationships between input features and the target variable (cell status) and its resilience to overfitting. Furthermore, the gradient boosting model is well-suited for imbalanced datasets and provides granular control over complexity and performance through hyperparameter tuning, including adjustments to learning rate, number of trees, maximum tree depth, and subsampling. This model lays the foundation for future research in AI applications in molecular and cell biology.

S29-P11-Compensatory mutations and the evolution of reduced costs of resistance in Pseudomonas aeruginosa

Nicholas Chow

McGill University, Department of Biology

The effectiveness of modern antibiotic treatments is threatened by the evolution of antimicrobial resistance (AMR) among human pathogens. In the absence of new drug development, to control infection, alternative treatment strategies using existing drugs are needed. Previous work in the Kassen Lab has shown that in the opportunistic human pathogen Pseudomonas aeruginosa, spatially variable drug treatment can delay, but not prevent, the spread of resistance. Although resistance typically comes with fitness trade-offs, under a spatially variable regime, generalist resistant strains eventually evolve to have high fitness even in the absence of drugs, thereby displacing sensitive strains. In line with studies such as Levin et al. (2000), here, we explore the role of compensatory mutations, which are second-site mutations that mask the costs of a resistance mutation, as an evolutionary mechanism for generalist strains. We use high-throughput amplicon sequencing of evolved P. aeruginosa populations under constant and spatially variable ciprofloxacin treatment at different time points to identify putative compensatory mutations. We identified mutations in two response regulator genes, gacA and gacS, that fixed repeatedly at later time points in tandem with mutations in known resistance genes gyrA and nfxB, signifying a compensatory relationship. To further confirm these genes as compensatory, resistance was measured by minimum inhibitory concentration (MIC) and fitness in the presence and absence of ciprofloxacin. Our results provide more detailed insight into how problematic resistant strains with reduced costs evolve and spread under a spatially variable drug regime and how these regimes may be used to improve treatment.

S29-P12-A Genetic and Histological Investigation of Testicular Abnormalities in North African Catfish

Worapong Singchat

Animal Genomics and Bioresource Research Unit (AGB Research Unit), Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

Thailand's hybrid catfish decline is due to poor-quality North African catfish (Clarias gariepinus) parent stock, with 50% of testes unsuitable for breeding, causing fertilization problems and low hatching rates. This study evaluated spermatogenesis in two testis forms—normal white (CGA_A) and black-brown (CGA_B)—via genetic and histological analyses, including microsatellite genotyping, CASA, meiotic chromosome analysis, gene expression profiling, histology, and TUNEL assays. Results revealed similar gene pool patterns despite differences in cytological processes, sperm quality, and fertility. The study revealed low sperm motility in African catfish (11.15%) and differences in spermatocyte stages between testis types. In both CGA_A and CGA_B testes, the interphase stage was predominant, with CGA_A exhibiting a higher percentage of pachytene stages compared to CGA_B, indicating differences in spermatogenic progression. CGA_B testes showed reduced expression of testicular development genes, while CGA_A testes demonstrated significantly higher expression of GtH I, GtH II, Dazl, and Cxcr4b, along with lower levels of Caspase-3 and p53 ($p < 0.05$), reflecting superior testicular function. Histological analysis showed that normal testes in the mature gonad phase were characterized by seminiferous tubules filled primarily with spermatozoa, whereas abnormal testes predominantly contained spermatogonia and spermatocytes. Apoptotic observations revealed that cell death mainly occurred in secondary spermatogonia within abnormal testes. The study reveals notable disparities in testicular semen quality and gene expression between normal and abnormal African catfish testes, providing crucial insights into the decline of Thailand's hybrid catfish production. Immediate research and intervention are vital to ensure sustainable aquaculture productivity.

S29-P13-Exploring Human Adaptation across Europe through a Polygenic Perspective

Xavier Martí-Pérez

Institute of Evolutionary Biology (UPF-CSIC)

Most of the phenotypic spectrum consists of complex traits, which are influenced by many loci throughout the genome. Though many minor allele frequency changes, polygenic selection acting on such multiple loci of small effect can produce remarkably rapid phenotypic shifts. As a result of pleiotropy, selection acting on complex traits can increase the prevalence of certain complex diseases as a direct evolutionary trade-off. Here, we infer whether selection has acted in recent and very recent time scales in immune responses to bacterial and viral stimuli as well as in polygenic traits of different nature from the UK Biobank and GWAS Catalog. To this aim, we use the statistics "trait integrated haplotype score" (tiHS), "trait singleton density score" (tSDS), and "polygenic adaptation likelihood method" (PALM) test. Additionally, we perform the joint PALM (J-PALM) analysis to detect signals of correlated selection and antagonistic selection. We applied these polygenic selection tests on the GCAT cohort –a proxy for the Iberian population– and on the European populations of the 1000 Genomes Project. In addition to the well-known directional selection for lighter skin pigmentation, our results show a generalized adaptive trend towards lighter hair colors and higher nervousness, along with an adaptive decrease in pulse rate and the prevalence of black hair color in most European populations. We also appreciate highly complex correlated selective responses between pigmentation and hair color, which suggest that there might have been synergic responses in these traits, rather than one being the driver of the others.

S29-P14-Pervasive sex-dependent effects in the genetic architecture of starvation resistance in Drosophila melanogaster

Junhao Chen

Peking University

Genetic variants often have sex-specific, sex-biased, or sexually antagonistic fitness effects, but their role in environmental adaptation related traits is not fully understood. Using pooled phenotype sorting and sequencing of male Drosophila melanogaster from natural populations, we identified genetic variants associated with starvation resistance. Many variants were located in regulatory regions or altered protein sequences. RNAi experiments showed that ~83% (83/100) of candidate genes with non-lethal knockdown effects exhibited sex-specific, sex-biased, or sexually antagonistic effects. 16 genes displayed sexually antagonistic effects, with 14 increasing female resistance while decreasing male resistance and two showing the opposite pattern. These effects did not correlate with sex-biased gene expression patterns. Further analysis of the Lnk gene indicated that non-synonymous mutations and gene expression changes have sex-dependent effects. Some candidate genes also showed trade-offs across different stress types. Our findings highlight the pervasive nature of sexual conflict in environmental adaptation related traits and suggest that the interplay among polygenic architecture, sex-dependent effects, and pleiotropic effects may maintain genetic variation and drive rapid adaptation.

S29-P15-The fission yeast ortholog of Coilin, Mug174, forms Cajal body-like nuclear condensates and is essential for cellular quiescence

Xiaoling Deng

Shanghai Jiao Tong University

The Cajal body, one of the biological condensates in the nucleus, was discovered over a century ago and is crucial for ribonucleoprotein (RNP) formation, including snRNPs and snoRNPs. Cajal bodies (CBs) also play an essential role in nonsense-mediated mRNA decay, RNAi-based silencing, and stress responses, with mutations in their components linked to several human diseases. Therefore, CBs are nuclear structures essential for various nuclear transactions in eukaryotes.

One of the most representative constituents of CBs is Coilin. Its precise functions remain unknown. Additionally, no Coilin ortholog has been identified in unicellular organisms so far. Therefore, characterizing Coilin in lower eukaryotic model organisms like yeast would be a critical step forward to better understanding Coilin.

This study reveals that Mug174 displays characteristics of Coilin: (1) Mug174 has an intrinsically disordered domain and can form phase-separated condensates; (2) Mug174 forms nuclear foci that overlap with the nucleolus and cleavage body; (3) Mug174 interacts with Tgs1 and U snRNAs; (4) Mug174 is required for TMG capping of U snRNAs and mRNA splicing; (5) Weak homology exists between Mug174 and human Coilin, and human Coilin expressed in fission yeast colocalizes with Mug174. According to these results, Mug174 is identified as the fission yeast ortholog of Coilin, and fission yeast possesses Cajal body-like nuclear condensates.

Critically, Mug174 is indispensable for the maintenance of and transition from cellular quiescence. These findings highlight the Coilin ortholog in fission yeast and suggest that the Cajal body is implicated in cellular quiescence.

S29-P16-Alcaligenes symbiont enhance parasitoid wasp Nasonia vitripennis tolerance to alcohol stress

Guan-Hong Wang

Chinese Academy of Sciences

Environmental alcohol, a byproduct of organic matter fermentation, represents a pervasive environmental stressor that impacts all organisms within alcohol-rich habitats. In response to this challenge, the bacterium Alcaligenes has evolved a remarkable capacity to metabolize alcohol, acquiring the essential adh2 and aldh2 genes through horizontal gene transfer (HGT). These genes, responsible for efficient alcohol metabolism, enable Alcaligenes to thrive in alcohol-laden environments. In this study, We demonstrate that Af colonizing the host Nasonia vitripennis can enhance its alcohol tolerance. This adaptation is driven by both the selective enrichment of Af in alcohol-exposed hosts and its metabolic contributions, specifically through the degradation of alcohol. Our findings reveal a novel, universal mechanism where symbiotic microbes, facilitated by HGT, bolster host adaptation to environmental stressors. Notably, using metagenomic data from diverse vertebrate and invertebrate species with chronic alcohol exposure, we observe a consistent increase in Alcaligenes genus in their microbiomes, reinforcing the role of HGT in microbial-driven host adaptation across species. This study offers the first experimental and data-driven evidence of how HGT-mediated adaptation via symbiotic microbes plays a critical role in enhancing host resilience to environmental pressures, with profound implications for understanding microbial evolution and host-microbe interactions.

S29-P17-Phenotypic Plasticity Facilitates Cellular Adaptation to Serum-Free Conditions via Transcriptomic Reprogramming

Jiayu Guo

Sun Yat-Sen University School of Life Sciences

Tumor cells can adopt diverse phenotypes to adapt to varying environments, and this phenotypic plasticity contributes to tumor heterogeneity, enabling tumor cell migration and drug resistance. However, the underlie molecular mechanisms remain poorly understood. For instance, it remains debated whether phenotypic changes in response to environmental shifts are driven by genomic mutations or epigenetic factors, and whether the phenotypic transformation process is driven by selection. Moreover, previous studies have largely focused on the stable states of cells after complete adaptation to new environments, with limited attention given to the adaptation process itself. In this study, we used the Chinese hamster ovary cell suspension line (CHO-S) adapting to serum-free medium as a model to systematically explore these issues at both the population and single-cell levels. The results revealed that during adaptation to serum-free medium, CHO-S cells underwent a series of phenotypic changes, including alterations in cell division and survival rates, without experiencing a significant bottleneck. Throughout this process, cellular phenotypic heterogeneity initially increased and then gradually decreased. These findings suggest that adaptation to serum-free conditions is not driven by selection and highlight the critical role of phenotypic plasticity in facilitating adaptation. Further bulk and single-cell transcriptomic sequencing across this process revealed significant transcriptomic reprogramming during adaptation, with distinct gene expression profiles and subpopulation compositions at different stages. Our study provides an in-depth analysis of phenotypic and transcriptomic changes during environmental adaptation, offering new insights into the role of phenotypic plasticity in tumor cell evolution.

Symposium 31: Open symposium session C

S31-P01-The natural and cultural history of brown bears in the Central Balkans as reconstructed through archaeological data and ancient DNA

Danijela Popovic

University of Warsaw

The brown bear (Ursus arctos L.) is an iconic megafauna species that survived the Late Pleistocene extinction. Historically, it has been widespread throughout Europe; however, human activities have led to its extirpation in many regions. In the Central Balkans, present-day Serbia, the brown bear is strictly protected. Extensive faunal collections from Holocene archaeological sites in the region revealed that brown bears were regularly hunted throughout (pre)history, suggesting that they were long-term residents of the local forests. However, the spread of agriculture, large-scale deforestation, and human expansion have significantly altered the natural habitats of bears. To better understand how these changes affected Holocene brown bears from Serbia, we conducted a mitochondrial DNA (mtDNA) analysis and obtained 34 complete mitogenomes. We detected not only all three known European mtDNA lineages but also a new haplogroup that has not yet been described. This suggests that the genetic diversity was much higher in the past. The greatest diversity was found in Viminacium, an important Roman site with a large amphitheatre where bears participated in spectacles, suggesting that this diversity may have resulted from intensive trading and the translocation of brown bears by humans. Our results showed that changes in mtDNA diversity over time were not only due to environmental changes but also by human expansion and activities. Future analyses of ancient and modern nuclear DNA will improve the current conservation efforts in the management of brown bears in the Balkans.

S31-P02-Obligate aerobic bacteria show higher A>G mutation rate comparing to anaerobes

Aleksei Skudnov

Immanuel Kant Baltic Federal University (IK BFU)

Understanding the mutational spectrum (MutSpec) provides insights into the factors driving mutations. While the processes underlying specific mutational signatures have been extensively studied in cancer and normal tissues, comparatively less attention has been given to bacterial species.

Here, we introduce BacNeMu, a bioinformatics pipeline designed to reconstruct neutral bacterial mutational spectra. BacNeMu is adapted from the recently published mitochondria-focused pipeline NeMu (doi:10.1093/nar/gkae438). BacNeMu is unique in its ability to analyze all available bacterial data with precision and ease due to its deep integration with leading procaryotic resources Annotree (doi:10.1093/nar/gkz246) and GTDB (doi:10.1093/nar/gkab776).

The mutational spectra inferred using BacNeMu show strong concordance with those measured directly through mutation accumulation experiments across eight bacterial species spanning the phylogenetic tree. This advancement enables large-scale investigations into the factors shaping bacterial MutSpec.

As an initial application, we tested the hypothesis that aerobic and anaerobic bacterial species exhibit distinct mutational spectra. Our analysis confirmed that the A>G mutation, previously implicated in oxidative damage responses in mitochondria (doi:10.1093/nar/gkac779), is significantly overrepresented in obligate aerobes compared to obligate anaerobes. Additionally, we analyzed four clades of closely related aerobic and anaerobic bacteria and found that the mutational spectrum is generally conserved due to phylogenetic inertia. However, at larger phylogenetic distances, it is primarily determined by the aerobic status of the bacterial clade.

Moving forward, we will leverage the BacNeMu pipeline across all sequenced/annotated bacterial species to uncover the profound impact of ecological factors (e.g., temperature) and intracellular mechanisms (e.g., DNA repair, replication properties) on mutation spectra evolution.

S31-P03-Evolution of duplicated elements regulating myogenesis following the salmonid whole genome duplication

Linjing Lan

The Roslin Institute, University of Edinburgh

Several eukaryotic lineages have experienced whole-genome duplication (WGD), providing raw genetic material that drives evolutionary innovation. Investigating WGD thus provides a deeper understanding of genetic, developmental and physiological evolution. Following the two rounds of WGD (2R) that shaped early vertebrate evolution, teleost fish underwent a third round of WGD (Ts3R) approximately 275 million years ago. Family Salmonidae later experienced a fourth round (Ss4R), making them an ideal model for studying evolution after WGD.

Skeletal muscle of farmed teleosts is of significant economic and nutritional importance, while the evolution of myogenesis in fish offers valuable comparative insights into vertebrate muscle development. The AQUA-FAANG (https://www.aqua-faang.eu) project produced comprehensive epigenomic annotations for a range of economically important farmed fish. Leveraging these resources, we focus on the myogenic regulatory factor (MRF) family of Atlantic salmon, the most farmed salmonid globally. MRFs comprise the vertebrate paralogues MyoD, Myf5, Mrf4, and Myog, representing the master transcription factors for myogenesis, providing an interesting system to investigate the evolutionary dynamics of duplicated genes following WGD.

After resolving the phylogenetic relationships of novel paralogous MRF genes in salmonids, we annotated and compared regulatory elements (enhancers, promoters, and conserved non-coding elements) in light of gene expression data, to trace the origins and explore functional divergence and conservation among MRF paralogues. Such comparative analyses provide insights into the specialised functional roles of these master muscle genes, ultimately contributing to a more comprehensive understanding of how myogenesis is regulated in vertebrates.

S31-P04-Inferring the Evolutionary History of Lycalopex Foxes in South America

Damián Zepeda

California State University, Northridge

South American canids (subtribe Cerdocyonina) are an extraordinary example of adaptive radiation. They comprise 10 extant species of canids distributed across diverse habitats, including six species of Lycalopex foxes. Many aspects of their evolution are poorly understood, such as their migration patterns, speciation, and adaptive evolution, as well as the extent and impact of hybridization among species. Using low coverage genomic ancient DNA (aDNA) data for 15 ancient Lycalopex spp. specimens, we sought to characterize the evolutionary history of these foxes. The ancient specimens were originally collected in Patagonia and were dated with C14 to 365 to 3,200 years Before Present. Endogenous DNA content ranged from ~0.05-59%, and the depth of coverage for the mitochondrial DNA (mtDNA) ranged from ~0.005-100x. Using these data, combined with publicly available data from other canid species, we characterize the genetic diversity of Lycalopex, detect population structure, identify potential signals of hybridization/introgression and natural selection, and construct a phylogeny for the mtDNA revealing the evolutionary relationships between these foxes and other canids. Our results shed light on the evolution of this understudied clade and will inform conservation strategies for canids.

S31-P05-Chromosome-level reference genome for the medically important Arabian horned viper (Cerastes gasperettii)

Gabriel Mochales-Riaño

NYUAD

Venoms have traditionally been studied from a proteomic and/or transcriptomic perspective, often overlooking the true genetic complexity underlying venom production. The recent surge in genome-based venom research (sometimes called "venomics") has proven to be instrumental in deepening our molecular understanding of venom evolution, particularly through the identification and mapping of toxin-coding loci across the broader chromosomal architecture. Although venomous snakes are a model system in venom research, the number of high-quality reference genomes in the group remains limited. In this study, we present a chromosome-resolution reference genome for the Arabian horned viper (Cerastes gasperettii), a venomous snake native to the Arabian Peninsula. Our highly-contiguous genome allowed us to explore macrochromosomal rearrangements within the Viperidae family, as well as across squamates. We identified the main highly-expressed toxin genes compousing the venom ' s core, in line with our proteomic results. We also compared microsyntenic changes in the main toxin gene clusters with those of other venomous snake species, highlighting the pivotal role of gene duplication and loss in the emergence and diversification of Snake Venom Metalloproteinases (SVMPs) and Snake Venom Serine Proteases (SVSPs) for Cerastes gasperettii. Using Illumina short-read sequencing data, we reconstructed the demographic history and genome-wide diversity of the species, revealing how historical aridity likely drove population expansions. Finally, this study highlights the importance of using long-read sequencing as well as chromosome-level reference genomes to disentangle the origin and diversification of toxin gene families in venomous species.

S31-P06-Comparative Analysis of Single- vs. Double-Stranded Library Preparation Methods for Ancient DNA Recovery

Motahare Feizabadi Farahani

University of Stockholm

Ancient DNA (aDNA) studies rely on efficient library preparation methods to maximize the recovery of degraded genetic material. In this study, we compare the efficiency of the Meyer & Kircher double-stranded library preparation method (Meyer & Kircher, 2010) with two single-stranded approaches: the Santa Cruz method (Kapp et al., 2021) and the single-stranded library method published by Gansauge et al. (2017, 2020). Our dataset consists of ancient human and mammoth bones spanning different time periods and exhibiting varying degrees of preservation.

In addition to comparing library preparation efficiency, we assess the impact of uracil-DNA glycosylase (UDG) treatment, which removes characteristic aDNA damage, by analyzing both UDG-treated and non-UDG-treated libraries for each method. The primary aim is to determine which approach recovers the highest proportion of authentic aDNA fragments while also considering factors such as time and resource efficiency. We will compare the library complexity, endogenous content, and fragment length across methods to evaluate their performance in different preservation conditions. The findings of this study will contribute to refining best practices in aDNA research, particularly for highly degraded specimens, and provide insights into method selection based on sample condition and available resources.

S31-P07-Evidence for corrin biosynthesis in the last universal common ancestor

Luca David Modjewski

Heinrich Heine University Düsseldorf

Corrinoids are cobalt-containing tetrapyrroles. They include adenosylcobalamin and cobamides that function as cofactors and coenzymes for methyl transfer, radical-dependent and redox reactions. Though cobamides are the most complex cofactors in nature, they are essential in the acetyl-CoA pathway, thought to be the most ancient CO2-fixation pathway, where they perform a pterin-to-cobalt-to-nickel methyl transfer reaction catalyzed by the corrinoid iron – sulphur protein (CoFeS). CoFeS occurs in H2-dependent archaeal methanogens, the oldest microbial lineage by measure of physiology and carbon isotope data, dating corrinoids to ca. 3.5 billion years. However, CoFeS and cobamides are also essential in the acetyl-CoA pathway of H2-dependent bacterial acetogens. To determine whether corrin biosynthesis was established before archaea and bacteria diverged, whether the pathways arose independently or whether cobamide biosynthesis was transferred between archaeal and bacterial lineages during evolution, we investigated phylogenies and structural data for 26 enzymes of corrin ring and lower ligand biosynthesis. The data trace cobamide synthesis to the last universal common ancestor (LUCA), while pterin-dependent methyl synthesis pathways likely arose independently post-LUCA in the lineages leading to bacteria and archaea. Enzymes of corrin biosynthesis were recruited from preexisting ancient pathways. Evolutionary forerunners of CoFeS function were likely Fe-, Ni- and Co-containing solid-state surfaces, which, in the laboratory, catalyze the reactions of the acetyl-CoA pathway from CO2 to pyruvate under serpentinizing hydrothermal conditions. The data suggest that enzymatic corrin biosynthesis replaced insoluble solid-state catalysts that tethered primordial CO2 assimilation to the Earth's crust, suggesting a role for corrin synthesis in the origin of free-living cells.

S31-P08-The evolutionary fate of laterally acquired genes depends on their utility and functional novelty

Catherine Collins

University of Sheffield

Lateral gene transfer (LGT) is widespread in eukaryotes, including plants, where it can drive adaptation and innovation. However, the adaptive significance and persistence of LGT remain unclear. Using Alloteropsis as a model, we examined grass-to-grass transfers through transcriptome and genome analyses. We found that most laterally acquired genes are expressed at lower levels and often truncated when compared to vertically inherited orthologs that co-exist in the genome. For laterally acquired genes that do maintain their expression, the level is more similar to the ortholog in the donor species than to the vertically inherited ortholog. Furthermore, laterally acquired genes exhibit higher methylation levels than vertically inherited genes. Overall, the results indicate that most laterally acquired genes detected in a genome are likely to be neutral or deleterious, and the actual background rate is likely to be many times higher as most LGT are rapidly silenced and lost. Those that are retained may encode their own regulation (e.g. presence of adjacent cis-regulatory element) and/or have a modified protein function. For a laterally acquired gene to drive adaptive evolution it needs to be readily usable by the recipient and to offer an evolutionary advantage.

S31-P09-A Time-Scaled Phylogeny of Brachyura: Insights into Divergence, Radiation, and Drivers of Crab Evolution

Jiayou Shao

Sun Yat-sen University

Brachyura comprises over 7,600 species, exhibiting extremely rich morphological and ecological diversity. Current research is constrained by insufficient molecular markers preventing a complete understanding of the crab phylogeny. To construct a reliable time-scaled phylogeny for all crab species, we used a "backbone-and-patch" approach by combining transcriptomic data from 96 species with 13 fossil calibration points to infer the relationships and divergence times of major lineages as the backbone tree. We then used 10 gene sequences from 2450 species to infer the patches. For species without DNA data, we estimated divergence time confidence intervals using birth-death models and imputed them under taxonomic constraints. These time-scaled patches were then joined to the backbone, resulting in a species-level phylogenetic tree of all Brachyura species. The results confirmed the polyphyly of the currently defined four superfamilies: Grapsoidea, Ocypodoidea, Xanthoidea, and Eriphioidea. Divergence time estimates suggest that crabs likely diverged as early as the Early Jurassic, with family-level divergences occurring primarily in the Late Cretaceous and early Paleogene. Ancestral state reconstruction analyses show that terrestrialization may be one of the main drivers of Thoracotremata diversity, with multiple semi-terrestrial and terrestrial lineages originating and radiating in the Early Eocene, coinciding with the global warming event during the Paleocene-Eocene Thermal Maximum (PETM). Additionally, species formation rate estimates show that several processes, such as the parasitic lifestyle in the Pinnotheridae family and the unique courtship behavior of Ocypodidae fiddler crabs, promote a higher speciation rate.

S31-P10-Population genomic evidence for MATa-MAT α and MAT α -MAT α sexual reproductions in Cryptococcus deneoformans

Jianping Xu

McMaster University

The Cryptococcus neoformans species complex is a critical priority fungal pathogen capable of both MATa-MATα and MATα-MATα sexual reproduction in laboratory settings. However, the genomic signature of MATa-MAT α and MAT α -MAT α sexual reproduction in natural populations remains largely unexplored. Here we analyzed the whole-genome sequences of 24 environmental strains of C. deneoformans from western Saudi Arabia, including one MATa and 23 MAT α isolates, with 15 MAT α isolates belonging to genotype ST160 as defined by their combined DNA sequences at seven loci. To identify signatures for MATa-MATα and MATα-MATα reproduction, three samples were analyzed: total, MATα, and ST160. For each subpopulation, single nucleotide polymorphisms (SNPs) were identified for both the nuclear and mitochondrial genomes and subjected to four-gametes test. In the total and the MATα subpopulations, variable proportions of SNP pairs within as well as between the nuclear and the mitochondrial genomes showed evidence for recombination. Though no mitogenome SNP was found among the 15 strains of ST160, the nuclear genome showed clear evidence for recombination, including among SNPs within the mating type region. In addition, the nuclear genome SNP pairs located further apart on the same chromosome showed greater evidence of recombination in all three sample types. In contrast, mitogenome recombination breakpoints were mainly located in two genomic regions. Together, these results provide unambiguous evidence for both MATa-MAT α and MAT α -MAT α sexual reproduction within this environmental population of C. deneoformans.

S31-P11-A novel data filtering method resolves the controversy in the phylogeny of the Chondrichthyes

Chenhong Li

Shanghai Ocean University

Phylogenomics, which uses genome-scale data for phylogenetic inference, has clarified many controversial nodes in the tree of life. This extensive data improve tree resolution and better reflect organismal history compared to analyses based on single or a few gene loci. However, some relationships within the tree of life remain unresolved, as increased data can yield high node support without ensuring accuracy due to systematic errors. For example, the order-level relationships among chondrichthyans are still contentious despite the use of phylogenomic data. To address systematic errors, complex models have been developed, and filtering for less erroneous data shows great promise. Current filtering methods often rank loci based on overall tree statistics, but problematic nodes are local, and other approaches struggle with circular assumptions. In this study, we introduced two novel data filtering methods based on the ratio of branch length or GC content between problematic clades. We applied these methods to a dataset of 4,577 single-copy exons extracted from 98 chondrichthyan species. The results using all loci showed that the Hexanchiformes was positioned at the root of Elasmobranchii, pulling other squalomorphs to the basal position and forming a paraphyletic Squalomorphii. Contrastingly, filtering for loci with more even branch length or GC content between focal clades strongly supported the monophyly of all orders of chondrichthyans as well as their higher classification grouping, such as Selachii and Batoidea. By concentrating on problematic nodes, our assumption-free filtering methods have great potentials in resolving contentious relationships in the tree of life.

S31-P12-Spatiotemporal evolution of AMR in Streptococcus suis: Gene variants and transmission dynamics.

Saron Tekie

University College London

Antimicrobial resistance (AMR) poses a critical global health threat, contributing to 4.71 million deaths in 2021 (Naghavi et al.), with this burden projected to rise sharply by 2050. Here, we explore the evolution of AMR in Streptococcus suis, a zoonotic livestock pathogen, in a global dataset of 3070 isolates dated 1960 to 2022. We capture gene diversity within TetO (tetracycline resistance gene) and ErmB and MefA (macrolide resistance genes). Gene variants show different transmission patterns, some with high rates of HGT and others vertically transmitted. We also see regional specificity of some variants, e.g., in the UK, Canada and China.

Additionally, we investigate the effect of antimicrobial use (AMU) on AMR. AMU significantly influences resistance prevalence and evolution, with notable disparities in resistance patterns across countries with differing AMU strategies. For instance, countries with higher AMU (e.g., USA) exhibit up to 30% and 48% higher prevalence of macrolide and tetracycline resistance, respectively, compared to countries with lower AMU (e.g., Denmark). Despite a decline in antimicrobial use over time, resistance persists in certain countries, likely driven by multiple factors. Interestingly, commensal pathogens exhibit lower sensitivity to AMU variations, raising questions about their role as reservoirs of AMR.

These findings deepen our understanding of the evolutionary dynamics of AMR, highlighting the complex interplay between gene transmission, antimicrobial use, and other ecological factors. Simultaneously, they emphasise the need for context-specific AMR mitigation strategies.

S31-P13-CAT-PMSF improves phylogenetic modeling and resolves Martialinae as the sister group to all remaining ants except Leptanillinae

Wentao Tao

University of Bristol

Formicidae, also known as ants, is a group of small social insects that inhabit nearly all terrestrial environments on Earth. They belong to Aculeata, a member of Hymenoptera, including 17 subfamilies. The phylogeny of ants is debated and three alternative hypotheses, differing because of the placement of the Martialinae have been proposed. To test these different scenarios, we used BUSCO genes and the novel CAT-posterior mean site frequencies (CAT-PMSF) method, which allow using dataset specific CAT-based models, that seem to outperform other models (e.g. LG and LG[C10-C60]) in describing across-site compositional dataset in maximum likelihood analyses. We show that when using standard models (LG or LG[C10-C60]), three topologies are consistently recovered that differ in the placement of Martialinae, as expected from the literature. The strength of the signals for these topologies, estimated as the number of genes trees, supporting the three topologies are comparable. We show that the newly developed CAT-PMSF model better accounts for across-site compositional heterogeneity in the data, and decrease the strength of two out of the three competing topologies. CAT-PMSF identify Martialinae as the most likely sister group of all the other ants but Leptanillinae. Our results suggest that while incomplete lineage sorting or introgression seem to have played a role in early ant evolution, phylogenetic errors induced by poor modelling of the underpinning data, might lead to an overestimation of ILS, as it might be difficult to distinguish gene trees genuinely supporting alternative topologies from incorrectly resolved gene trees.

S31-P14-Scorpions in disguise: Unraveling the cryptic Hottentotta diversity in India

Vivek Suranse

Indian Institute of Science, Bengaluru, India

Scorpions are venomous terrestrial arthropods with a remarkably conserved body plan despite over 400 million years of evolution. The family Buthidae includes several medically significant species, notably those in the genus Hottentotta, which is comprised of three major lineages: African, Sahar-Sindian, and Indian. In India, Hottentotta is a cryptic group with eight described species, yet the actual diversity remains unclear due to historical reliance on morphology with limited molecular insights and an absence of exhaustive nationwide investigation. This study adopts an integrative approach, combining taxonomy, phylogenetics, phylogenomics, and population genomics to reassess species boundaries and evolutionary relationships. The findings indicate the need for systematic revisions within the genus, suggesting cryptic diversity and potential taxonomic realignments. By leveraging genomic tools alongside traditional taxonomy, this research aims to refine species delimitations, uncover biogeographic patterns, and enhance our understanding of Hottentotta diversity in India. These findings will provide a critical foundation for future ecological and evolutionary research while contributing to scorpion systematics in the region.

S31-P15-Ancient mitogenomes reveals two divergent caballine horse lineages present at the Basin of Mexico towards the end of the Late Pleistocene

Pablo Esteban Uribe Herrera

Facultad de Estudios Superiores Iztacala. Universidad Nacional Autónoma de México. Tlalnepantla de Baz , Edo. Mex. México

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 within their phylogeny is problematic due to the limitations of morphological-morphometric characterization methods. Paleogenomics has provided an opportunity to perform high resolution taxonomic identification, as well as a characterization of the evolutionary trajectory of ancient horses; although such approximation has yet to be performed incorporating Mexican samples. We obtained 86.4x and 3.19x mitogenomes from two prehistoric horse samples from the Basin of Mexico (Central Mexico), radiocarbon dated to circa 11,270 and 11,750, respectively. We performed mitochondrial phylogenetic analyses restricting to >10x ancient and present-day mitogenomes associated to the caballine clade, to characterize their topology and divergence time of nodes within this clade. Finally, we phylogenetically placed the 3.19x coverage sample into the caballine mitogenome phylogeny to assess its genetic affinity to other samples. We found two different lineages of caballine horses in Central Mexico: 1) The low coverage sample being related to Late Pleistocene Russian lineages. 2) The high coverage sample being a basal to all the caballine mitochondrial genetic variation reported to date. Our study provides a new insight to further understand the diversity and evolutionary trajectory of prehistoric horse lineages in Mexico and North America, and highlights the importance of increased genetic sampling from tropical latitudes.

Symposium 35: Open symposium session E

S35-P01-Hypoxia response is conserved between Cnidaria and Bilateria

Sen Hadife

Institute of Science and Technology

Oxygen availability is hypothesized as one of the critically defining elements that led to the diversification of Metazoa. Geological records indicate a rise in atmospheric O2 levels during the Neoproterozoic, and molecular phylogenies suggest that this increase in O2 availability coincided with the emergence of animal lineages. Extant metazoans rely on oxygen for metabolism, gene expression, and developmental processes, underscoring its integral role in early animal evolution. Despite its presumed importance, the molecular mechanisms underlying oxygen-dependent gene regulation mediating early animal evolution remain poorly characterized.

To explore these mechanisms, the cnidarian Nematostella vectensis, a representative of early-diverging metazoans, is used to comprehensively investigate the genetic responses to hypoxia stress. RNA-sequencing based transcriptomic profiling reveals that the response of major canonical hypoxia-inducible factor (HIF) signalling components cannot be detected, yet a suite of genes homologous to bilaterian HIF targets is upregulated under low-oxygen conditions. Furthermore, comparative analysis of the hypoxia response between N. vectensis and several representative bilaterian species uncovers a core hypoxia-responsive gene set and metabolic pathway repertoire shared between Cnidaria and Bilateria. Notably, both phyla exhibit transcriptional reprogramming toward anaerobic glycolysis and enhanced NADPH-producing pathways during embryogenesis under hypoxia.

Collectively, these findings demonstrate that the eumetazoan ancestor possessed an oxygen-sensitive transcriptional network predating the divergence of Cnidaria and Bilateria, and highlight the deep evolutionary roots of hypoxia adaptation in Metazoa.

Session 2

Symposium 3: Experimental evolution: studying evolution in action

S3-P01-Mutation rate assessment uncovers cooperation and conflict dynamics between mitochondrial and nuclear genomes in Daphnia pulex

Jigar Trivedi

McGill University

The uniparental inheritance and hierarchical population structure of mitochondrial DNA present a unique challenge in studying its evolution. In this study, we assessed essential parameters of the mitochondrial evolution of Daphnia pulex by analyzing 73 mitochondria genomes from a decade-long mutation-accumulation experiment spanning over 189 generations and sequenced over three time points. We estimated the mitochondrial mutation rate, spectrum, and heteroplasmy levels by implementing three key bioinformatic approaches: i) assembling and annotating the reference genome of the progenitor MA mother, ii) developing a bioinformatics pipeline to segregate nuclear-mitochondrial sequences (NumtS), iii) assessing the fitness effect of the selfish nonsynonymous mutation on the gene function. The observed mutational spectrum was dominated by low-frequency heteroplasmic deletions with rare single nucleotide mutations increasing in frequency. Additionally, we tracked the selfish propagation of five mutations, resulting in their fixation in the mitochondrial genomes, two of which had a negative fitness effect. The average mitochondrial DNA copy number was estimated to be 111, with a 33% variance between MA and controls lines (CC) propagated under the influence of selection. Notably, the mutation rate, spectrum, and heteroplasmy did not differ significantly, suggesting the role of germ-line bottleneck and multilevel purifying selection in preventing the fixation of deleterious mutations. Our findings underscores the importance of cooperation and conflict between mitochondrial genomes and its host nuclear genomes in the mito-nuclear coevolution of Daphnia pulex

S3-P02-The loci of environmental adaptation in a model eukaryote

Piaopiao Chen

Zhejiang Univerisity

While the underlying genetic changes have been uncovered in some cases of adaptive evolution, the lack of a systematic study prevents a general understanding of the genomic basis of adaptation. For example, it is unclear whether protein-coding or noncoding mutations are more important to adaptive evolution and whether adaptations to different environments are brought by genetic changes distributed in diverse genes and biological processes or concentrated in a core set. We here perform laboratory evolution of 3360Saccharomyces cerevisiae populations in 252 environments of varying levels of stress. We find the yeast adaptations to be primarily fueled by large-effect coding mutations overrepresented in a relatively small gene set, despite prevalent antagonistic pleiotropy across environments. Populations generally adapt faster in more stressful environments, partly because of greater benefits of the same mutations in more stressful environments. These and other findings from this model eukaryote help unravel the genomic principles of environmental adaptation.

S3-P03-Just Keep Swimming - Using motility rescue as a model system to determine whether Sigma factor control or genetic background is a bigger driving force influencing bacterial evolution

Mitchell Reynolds

University of Bath

Gene regulatory networks exhibit remarkable adaptability, but the factors shaping their evolutionary rewiring remain poorly understood. Using flagellar motility rescue as a model, I investigate how differences in global regulatory network architecture and genomic context influence the predictability of rewiring pathways across diverse bacterial species. Previous work has shown that in Pseudomonas fluorescens, motility can be restored in the absence of the master regulator FleQ through the co-option of NtrC, a response regulator associated with nitrogen uptake. To explore the generality of this process, we engineered deletions of the flagellar master regulator in six bacterial species—Pseudomonas aeruginosa, Caulobacter crescentus, Vibrio cholerae, Escherichia coli, Bacillus subtilis, and Bordetella bronchiseptica—representing a range of regulatory network architectures. Under strong selection for motility, P. aeruginosa demonstrated rapid and repeatable motility restoration, similar to observations in P. fluorescens, with evolved mutants exhibiting slow motility within 42 hours. However, genome sequencing of motile isolates revealed distinct rewiring pathways, highlighting how variations in network architecture shape adaptive trajectories. By integrating these results across species, this study aims to uncover the principles governing the evolution of regulatory networks, providing broader insights into their adaptability and evolutionary dynamics.

S3-P04-Evolution temperature predicts lower, but not upper, temperature limits for growth in experimental populations of two bacteria

Ya-Nan Zhen

Beijing Normal University

Natural populations usually occupy thermal niche space that ranges between near optimal and mildly cold temperatures. The lower temperature limit for growth is likely determined by rate-limiting processes, and the upper temperature limit by thermal stability of proteins. We argue that the fitness effects of mutations to impact rate-limiting processes at low temperatures are likely consistent across a range of relatively lower temperatures; however, mutations that affect protein stability may have highly variable fitness consequences between optimal and even high temperatures. We therefore predict that the lower, but not upper, temperature limits can be predicable from populations' evolution temperatures. This prediction is consistent with our observations in a long-term experimental evolution study with Escherichia coli and Pseudomonas fluorescens. Moreover, we did not observe a trade-off between tolerance to cold and hot temperatures. Candidate mutations responsible for evolutionary changes in lower and upper temperature limits were identified. Our results help to understand the potential for rapid evolutionary adaptation under global warming.

S3-P05-Contribution of the SOS response and the DNA repair systems to norfloxacin induced mutations in E. coli

Hongan Long

Ocean university of China

Antibiotic-resistant bacteria pose a major threat to human health, with resistance often driven by the selection of pre-existing mutations or the induction of de novo mutations by antimicrobial agents like fluoroquinolones. Oure previous study found that these antibiotics elevate the genome-wide mutation rate through mechanisms such as activation of the SOS response pathway and the impairment of DNA repair systems. However, the relative contributions of these mutagenic mechanisms remain unclear. To address this, we conducted mutation accumulation experiments with wild-type and SOS-uninducible Escherichia coli strains, along with strains deficient in mismatch repair (MMR) and oxidative damage repair. Our results revealed that the SOS response accounts for approximately 30–50% of the increase in base-substitution mutation rate upon treatment with sublethal doses of norfloxacin. Moreover, DNA repair systems were found to significantly reduce the BPS rates and the transversion rates while increasing the structural variation SV rate, thus offering new insights into how these mechanisms contribute to antibiotic-induced mutagenesis. We also extended our research to explore the potential mutagenic effects and mechanisms of fluoroquinolones on non-target eukaryotic genomes. By contrast, no mutation rate elevation was observed in the fission yeast S. pombe, likely due to nucleotide excision repair counteracting potential mutagenic effects. Together, these findings provide a comprehensive understanding of antibiotic-induced mutagenesis in both prokaryotic and eukaryotic organisms, highlighting the differential risks to bacterial and eukaryotic genomes.

S3-P06-Gene expression plasticity in the antibiotic environment is predominantly adaptive in the budding yeast

Yang Qian

Zhejiang University

Phenotypic plasticity is a widespread phenomenon observed across species. Generally, it can benefit organisms in two primary ways: by increasing survival in novel environments or by shifting phenotypic traits closer to their optimal values. However, even when phenotypic plasticity has beneficial effects, it does not necessarily imply that it is an adaptive trait selectively favored and fixed by natural selection. Due to its biological significance and the feasibility of high-throughput measurement, gene expression plasticity has been extensively studied. Whether plastic gene expression is adaptive remains a topic of active debate. In this study, we directly addressed this debate by examining how plastic gene expression changes in the absence of natural selection. We quantified expression plasticity levels in 21 yeast mutation accumulation (MA) lines, and in 22 wildtype (WT) yeast strains, which represent natural selection is largely absent and strong respectively. Both the overall level of gene expression plasticity and the number of differentially expressed genes were reduced in the antibiotic environment in the MA lines, where the direction of plastic expression was more frequently reversed. These suggest that gene expression plasticity in the antibiotics is adaptive. In contrast, no significant signal of adaptiveness was detected in the salt environment, indicating that plastic expression in this condition may be neutral. Finally, we identified shared biological pathways that were overexpressed in WT yeast in the antibiotic environment, potentially contributing to increased survival. Our study provides a direct test of the adaptiveness of gene expression plasticity, offering insights into its underlying molecular mechanisms.

S3-P07-Hiking the Peaks and Valleys of GFP Fitness Landscapes for Data-Driven Protein Design

Anna Toidze

ETH Zürich

Protein fitness landscapes map genotype to phenotype, offering insights into the constraints underlying protein evolution and design. We studied the fitness landscapes of four highly divergent green fluorescent proteins (GFPs) by associating thousands of mutant genotypes with their corresponding fitness, fluorescence intensity. The four genes varied widely: some were mutationally robust with flatter fitness landscapes, while others had sharper fitness peaks and lost fitness rapidly due to negative epistasis between mutations. Machine learning (ML) models trained on this data were used to design novel, functional GFPs with over five times as many mutations as in the training set. Surprisingly, designs based on mutationally robust proteins were less successful, suggesting that highly epistatic proteins are more suitable for generating ML-compatible data. This led us to investigate whether mutational robustness itself could serve as a predictor for design-friendly proteins. We further analyzed functionally validated artificial GFPs derived from a single fragile GFP. Despite sharing high sequence identity, these variants displayed considerable variability in mutational robustness and epistasis, unrelated to their sequence divergence from the wild type. We found that multiple small landscapes can be as effective as a single large one for designing functional GFPs, as long as sequence divergence remains small. Since mutational robustness influences the ML-predictability of fitness landscapes, we asked whether it could, in turn, be predicted by a measurable property. We investigated computational folding predictions, thermostability, and sensitivity to denaturing agents as potential proxies. These findings highlight strategies to survey fitness landscapes and apply them to protein design.

S3-P08-The relative contributions of mutation types to adaptation in changing environments

Shichun Wang

College of Life Sciences, Zhejiang University, Hangzhou 310058, China

The relative contributions of different types of mutations (e.g., coding vs. noncoding mutations) to adaptation in molecular evolution remain a subject of debate. Evolution experiments under constant laboratory conditions typically show a higher proportion of mutations in coding regions compared to noncoding regions, particularly a higher frequency of loss-of-function mutations, suggesting that these mutations may be under stronger positive selection. In contrast, natural evolution inferred from population genetic analyses indicates that such mutations are seldom observed, implying they are more likely to experience stronger purifying selection. Here we hypothesize that while large-effect mutations, such as loss-of-function mutations in coding regions, confer stronger adaptive benefits in constant environments, they may become deleterious in natural changing environments due to antagonistic pleiotropy, impeding their fixation. To test this hypothesis, we conducted yeast experimental evolution across 100 different media, with a changing environment cycling through them at four distinct frequencies. In parallel, we performed evolution experiments in constant environments for comparison. Our findings demonstrate a lower ratio of coding region and loss-of-function mutations in changing environments compared to that in constant environments. Moreover, large-effect mutations are less unlikely to undergo loss of heterozygosity in changing environments. Additionally, natural populations exhibit more polymorphic sites corresponding to mutations from experimental evolution under changing environments than under constant conditions. Notably, these patterns extend beyond unicellular organisms to multicellular systems. These findings provide fundamental insights into the mechanisms of adaptation and highlight the differing roles of mutation types in adaptive evolution.

S3-P09-Mechanisms driving adaptive evolution in the cross-kingdom fungal pathogen Fusarium oxysporum

Li-Jun Ma

University of Massachusetts Amherst

Fungi constitute a vital component of our ecosystem, significantly impacting agricultural productivity and public health. Climate change has intensified the pressure caused by fungal diseases due to their adaptability in changing environments. Understanding mechanisms driving fungal adaptation, particularly in environments that are associated with disease and impacted by climate change, is vitally important. Fusarium oxysporum is a unique fungal pathogen with cross-kingdom virulence. Using multi-omics and short-term experimental evolution, we studied genome dynamic of a plant pathogenic strain Fol4287 and an eye infecting strain MRL8996. They share a core genome of about 46 Mb, but have distinct transposable element (TE) profiles and encode genes unique for host-specific virulence. Serial passaging of a clonal isolate resulted in populations displaying significantly increased fitness in the respective passaging condition. Sequencing of the evolved lines revealed changes both at the nucleotide and chromosome level, some are fixed in the population. The most detected variants in both evolving passages are TE insertions, as DNA transposons are highly active in Fol4287 and retrotransposons are active in MRL8996. Parallelism or the predictability was observed, as multiple independent lineages evolved under the same conditions carry events at the same locus. Epistasis was detected showing a stepwise optimization of mutations, one depending on the fixation of the other. Collectively, our results provide new insights into the genetic mechanisms underlying fungal adaptation and rectify the power of the experimental evolution in helping us to strategize novel means to control diverse fungal diseases under the rapidly changing environment conditions.

S3-P10-Evolutionary features of sexual reproduction loss in rice-derived Magnaporthe oryzae

Xifang Zong

School of Life Sciences, Nanjing University

The fungal disease caused by Magnaporthe oryzae is one of the most destructive threats to crops globally, making the study of its genetic evolution essential for effective control of rice blast. While sexual reproduction is thought to facilitate host jumping and enhance disease spread, the prevalence of clonal lineages in natural rice blast populations challenges this expectation. In this study, we collected 252 population samples and tracked the fixation and differentiation of repeat-induced point (RIP) mutations, evaluating the sexual reproduction status of M. oryzae in natural environments. Our results indicated that populations infecting non-rice plants exhibit stronger sexual reproduction capabilities, whereas some rice-derived lineages have lost this ability. To further examine whether RIP mutation events occur during sexual reproduction, we conducted cross-experiments across different populations and obtained 92 progenies. We observed significant differences in recombination and RIP rates between rice- and non-rice-derived strains. The weakened sexual characteristics in rice-derived strains are likely a result of relaxed selection pressures. These findings provide new insights into the adaptive evolution and spread of M. oryzae.

S3-P11-Phenotypic Plasticity and Distinct Adaptive Strategies of Cancer Cells Under r- and K-Selection

Lu Yang

Kunming Institute of Zoology

The r- and K-selection theory, originally developed in ecology, has proven useful in understanding the phenotypic diversity of tumor cells in response to environmental factors. Applying this framework, we previously identified two distinct tumor cell phenotypes: r-cells, characterized by rapid proliferation, and K-cells, which excel at adapting to high-density environments. The trade-off between growth and survival in r- and K-cells has been validated across several biological dimensions. In this study, we analyzed gene expression variation at the single-cell RNA level, revealing distinct transcriptional profiles between the two phenotypes. Specifically, r-cells exhibited a more stochastic gene expression pattern, reflecting their high plasticity, whereas K-cells segregated into two relatively stable sub-populations, indicative of their specialized and consistent state. To further investigate their adaptive capabilities, we evaluated the responses and survival rates of these phenotypes to two chemotherapeutic agents —harringtonine and cisplatin -- both known for their broad anti-tumor efficacy. The results revealed that r-cells exhibited significantly changes in gene expression when exposed to stress, highlighting their ability to adapt in unpredictable conditions. In contract, K-cells showed responded less gene expression changes in response to these stressors. Although the specialized traits of K cells make them highly competitive in high-density environments, this same specialization appear rapidly adapt when faced with sudden environmental shifts. These findings provide new insights into how phenotypic plasticity drives cancer cell survival and resistance, highlighting the role of r- and K-selection in tumor evolution and therapy resistance.

S3-P12-A Methyltransferase Homolog in Neurospora crassa Mediates RIP Mutation and Heterochromatin Formation in Repeat Sequences

Nannan Wu

Nanjing University

In most eukaryotes, repeat sequences - such as transposable elements - are compacted into heterochromatin bearing repressive epigenetics marks to safeguard genome stability. However, the mechanisms underlying repeat recognition and the initiation of heterochromatin formation remain poorly understood. The fungus Neurospora crassa has long served as a model for heterochromatin research. In its genome, constitutive heterochromatin - defined by H3K9me3 and DNA methylation - predominantly colocalizes with repeat sequences that bear signatures of extensive repeat-induced point mutations (RIP). Studies using artificially constructed Repeat-Linker-Repeat (R-L-R) structure suggest that heterochromatin assembly is mediated by a RIP-dependent mechanism involving two DNA methyltransferases homologs, RID and DIM-2. While externally introduced duplicates are typically unmethylated prior to RIP, endogenous duplicates are frequently organized into constitutive heterochromatin due to historical RIP events. By systematically comparing historical RIP-affected duplicates with newly formed, RIP-unaffected duplicates, we identified potential factors in heterochromatin formation and demonstrated that chromatin state can influence mutation mechanisms. Importantly, using a pair of newly formed native duplicates as a model, we uncovered a novel de novo heterochromatinization pathway that is largely RIP-independent, shedding light on the transition to and maintenance of constitutive heterochromatin. Overall, our experimental systems provide new insights into the interplay between chromatin state and genome stability, offering potential pathways for chromatin state transitions and the regulation of mutation rates.

S3-P13-Populational Evolutionary Dynamics of Fusion Recovery in Prm1-Deficient Schizosaccharomyces pombe

Lin Bi

Shenzhen Institute of Advanced Technology

In Schizosaccharomyces pombe, knockout of Prm1 in the wild-type strain results in a dramatic reduction in fusion efficiency to 10%.

In previous study, we conducted a bottleneck single-colony selection evolutionary experiment on the prm1-knockout strain, applying selective pressure through tetrad dissection to isolate successfully fused spores. We achieved a fusion-efficiency-recovered strain, termed E-ancestor with ~40% fusion efficiency. Compared to single-colony evolution, featured in genetic drift and rapid fixation of mutations, population-level evolution preserves multiple subpopulations evolving in parallel, increases genetic diversity and broadens adaptive potential, facilitating multiple independent evolutionary routes for fusion recovery.

To investigate the adaptive trajectory of fusion efficiency recovery following prm1 knockout, we subjected the wild-type, prm1Δ ancestor, and the single-spore evolved strain (prm1Δ E-ancestor) to a population-level evolution experiment involving 50 cycles of mitosis and meiosis alternation. Selective pressure was applied using snailase treatment, which selectively digests the cell walls of unfused cells. Throughout the experiment, we monitored fusion efficiency and growth rates through high-throughput automation platform of flow cytometer and spectrometer.

Our data revealed a progressive increase in fusion efficiency, characterized by notable fluctuations in performance across evolutionary time. We tracked 72 strains over 50 cycles of mitosis and meiosis, totaling ~650 generations of mitosis and 50 rounds of meiosis. To dissect the genetic basis of these adaptive responses, we sequenced ~3,500 samples, identifying genetic variants and tracking evolutionary dynamics across different populations. This extensive dataset allows us to explore distinct evolutionary trajectories and uncover potential mechanisms underlying fusion efficiency fluctuations.

S3-P14-Exploring Host Genome Responses to Newly Integrated Mobile Genetic Elements in Neurospora crassa

Gai Niu

Nanjing University

Mobile genetic elements (MGEs) drive genome evolution but also threaten genome stability. All cellular life forms face genetic parasitism from viruses and other MGEs and have evolved defense systems, such as DNA methylation, RNAi and cytosine deamination, to protect against these elements. The evolution of life is characterized by an ongoing conflict between hosts and parasites, with no ultimate victor. However, the dynamics of newly emerged MGEs in host genomes remain poorly understood.

In this study, we used the filamentous fungus Neurospora crassa as a model to explore the interaction between young MGEs and the host genome. N. crassa silences MGEs repeats through several mechanisms: DNA methylation, repeat-induced point mutations (RIP) during sexual reproduction, and quelling, an RNAi-like pathway during asexual reproduction. To investigate the response of N. crassa to newly integrated MGEs, we introduced two DNA viruses and a retrotransposon into its genome.

Our results show that RIP induces mutations in viral sequences, both in genomic and extrachromosomal repeats. Sequencing revealed that viral sequences form tandem repeats, leading to local amplification, likely through long terminal repeats. Furthermore, both viruses undergo de novo DNA methylation, although the mechanisms differ: one virus's methylation is associated with a quelling-related protein, while the other follows a distinct pathway. Additionally, the cytosine deaminase pathway appears to conflict with quelling-related DNA methylation in the defense response. In conclusion, N. crassa employs multiple defense strategies to combat MGEs, which can cooperate or conflict.

S3-P15-Research on the Evolutionary Dynamics of HLA in TCR-T Cell Therapy

Jiarui Weng

Guangzhou Medical University

Melanoma, an aggressive skin cancer, is difficult to treat in late stages. While TCR-T cell therapy shows promise, tumor escape remains a significant challenge, often due to the loss of MHC molecules. To investigate how cancer evolution and heterogeneity drive immune evasion, we first subcutaneously transplanted the A375 melanoma cell line into NCG mice, followed by multi-site puncture sampling at different stages of 1G4TCR-T therapy.This was followed by HLA gene-targeted sequencing and omics analysis.Growth curve analysis showed that tumor recurrence was not related to 1G4TCR-T dose, suggesting that factors beyond treatment intensity, such as cancer evolution and tumor heterogeneity, are driving resistance. Analysis of non-synonymous mutations in MHC molecules revealed that HLA-B had the most mutations, followed by HLA-A. These mutations highlight the dynamic nature of tumor evolution, where selective pressures contribute to the emergence of resistant mutations within the heterogeneous tumor population. Since 1G4TCR targets HLA-A-presented antigens, HLA-A mutations likely contribute to tumor recurrence. VAF (Variant Allele Frequency) analysis identified mutations associated with drug resistance, emphasizing the role of genetic heterogeneity and evolving immune escape mechanisms. These findings underscore the need for therapeutic strategies that target the evolving landscape of resistant mutations in melanoma.

S3-P16-High-Throughput Experimental Evolution Uncovers Universal Mechanisms of Enzyme Temperature Adaptation

Ke Li

Institute of Genetics and Developmental Biology, Chinese Academy of Sciences

Enzymes evolve distinct strategies to balance catalytic activity and structural stability under temperature extremes. While previous studies primarily employed sequence and phylogenetic analyses to examine temperature adaptation in natural enzymes, such approaches inherently conflate multifactorial selection pressures (e.g., pH fluctuations, substrate diversity) imposed by natural evolution. To identify temperature-specific adaptation mechanisms, we implemented an experimental evolution framework in yeast by constructing high-throughput mutant libraries of model enzymes (including dehydrogenases, oxidoreductases, decarboxylases) and subjecting them to adaptive evolution under gradient temperatures (10°C, 20°C, 30°C). Integrated high-resolution profiling of low- and high-temperature adapted mutants, combined with detailed structural analyses, revealed principles governing temperature adaptation: Low-temperature adaptation showed enrichment of mutations near the active site, which may contribute to maintaining flexibility for substrate binding under low thermal energy, whereas high-temperature adaptation frequently involved residues distal to the active site, possibly supporting structural stabilization through hydrogen bond reorganization. This work disentangles temperature-specific adaptation pathways from complex natural evolutionary constraints, providing actionable guidelines for engineering thermally adaptive biocatalysts in synthetic and industrial applications.

Symposium 8: Novel insights on genome architecture evolution

S8-P01-Architecture and adaptiveness of B chromosomes in the yellow-necked mouse genome

Vladimir Jovanovic

Freie Universität Berlin

B chromosomes (Bs) are supernumerary chromosomes present in some individuals within a population, representing a distinctive form of genomic polymorphism found in approximately 3% of eukaryotic species. The yellow-necked mouse (Apodemus flavicollis) is a widespread European rodent exhibiting B chromosomes that are, in part, derived from fusions involving the pericentromeric regions of sex chromosomes and sequences homologous to the standard chromosome complement. To investigate B chromosome architecture and evolution, we assembled de novo genomes of two male yellow-necked mice from the same litter, one possessing a B chromosome (B+) and one lacking it (B-). Comparative genomic analysis identified extended B chromosome-specific sequences, characterized their sequence similarity to autosomes and sex chromosomes, and pinpointed chromosomal breakpoint junctions. Crucially, we analyzed 5-methylcytosine (5mC) methylation patterns across the B chromosome and at breakpoint regions using SMRT sequencing data, revealing a distinct methylation profile of Bs. While Bs are often considered parasitic or neutral genomic elements, their potential adaptive roles remain debated. To address this, we assessed the transcriptional activity of B chromosome-derived sequences in yellow-necked mice inhabiting the Chernobyl Exclusion Zone, a chronically irradiated environment. A transcriptional activity originating from the B chromosome was detected and assigned to the specific genes. This study provides novel insights into the structural organization and epigenetic regulation of Bs. Furthermore, it assesses the significance of B chromosomes in a natural population under strong selective pressure.

S8-P03-Insights into Genome Reorganization in Phylum Mollusca

Zeyuan Chen

Senckenberg Nature Museum

Phylum Mollusca exhibits the widest diversity of body plans in metazoan evolution. Over the course of nearly 500 million years, some taxa, such as chitons (the taxonomic class Polyplacophora), have remained highly conserved morphologically, despite experiencing significant and dynamic genome rearrangements. In contrast, bivalves (Bivalvia), the second-largest molluscan class, show remarkable stability in chromosome numbers and structures. Among the class Gastropoda, a whole-genome duplication event in the ancestors of pulmonate land snails and slugs is believed to have facilitated their transition from aquatic to terrestrial environments. Cephalopods, on the other hand, have undergone extraordinary genomic reorganizations, displaying significant variation between species. For less studied and more enigmatic classes such as Aplacophora and Monoplacophora, our knowledge remains limited. Working with a large collaborative team, we recently completed the first comprehensive phylogeny of the whole phylum Mollusca based on multiple genomes from each of the 8 living classes, including the first two complete genomes of the enigmatic monoplacophorans. The new genomes assembled as part of this project enable new insights about the relationship of genome architecture and molluscan evolution, morphology, and adaptations. The flexibility of the molluscan genome likely underpins both historical challenges in genome research and the evolutionary success of this phylum.

S8-P04-Intron-exon dynamics and splicing mechanisms in Myxozoa: insights from long-read sequencing

Dorothee Huchon

Tel Aviv University

The function and evolution of spliceosomal introns remain enigmatic, with intron lengths and gain-loss dynamics reflecting diverse selective pressures across eukaryotes. Myxozoans are highly reduced parasitic cnidarians, and represent a unique system for exploring intron evolution in animals, due to their compact genomes. Spliceosomal introns shorter than 30 bp were thought to be absent in animals, plants, and fungi (except for the parasitic Microsporidia). While most animal introns range between 50–90 bp, we now show that myxozoans have introns shorter than 30 bp. Intron length distributions vary among myxozoan species, with peaks of ~20 bp in freshwater species but ~50 bp in marine species, suggesting lineage-specific splicing constraints. Unlike Microsporidia, where introns are rare, we found that the average intron content in Myxozoa ranges from 2.10 to 4.91 introns per gene, comparable to Caenorhabditis elegans (4.84) and Drosophila melanogaster (3.32). Moreover, we confirmed the absence of minor spliceosome components in myxozoan genomes, indicating that their short introns are processed exclusively by the major spliceosome. It has been suggested that short introns led to the loss of alternative splicing in the myxozoan Thelohanellus kitauei. Conversely, our PacBio sequencing of the Henneguya salminicola transcriptome revealed evidence of alternative splicing, including alternative 3′/5′ splice sites and cassette exons, and widespread intron retention. Notably, no strong splicing motifs were detected in the short introns of H. salminicola. These findings challenge assumptions about splicing in highly reduced genomes and highlight myxozoans as an invaluable model for studying intron evolution and splicing mechanisms in animals.

S8-P05-Chromosomal Synteny Provides Robust Evidence for the Phylogenetics of Fabales

Qiyi Yang

South China Botanical Garden, Chinese Academy of Sciences

The order Fabales, comprising the families Leguminosae, Polygalaceae, Quillajaceae, and Surianaceae, is widely recognized as a monophyletic group in angiosperm molecular phylogenies. Despite good support for the order, the interfamilial relationships remain unresolved. To address this, we conducted a macro-synteny oriented phylogenomic analysis of 7 Fabales genomes and identified 7 ancestral linkage groups (ALGs) that trace back to the common ancestor of all Fabales. We demonstrate the consistent fusion with mixing events of these ALGs throughout the evolutionary history of Fabales, despite extensive genomic exchange, chromosomal shuffling and gene fractionation driven by polyploidy. We also reconstructed phylogenetic frameworks for Fabales based on nuclear or organelle genes. The results show that Polygalaceae diverged first from the other families, while Leguminosae is sister clade to Quillajaceae and Surianaceae. This study highlights the utility of synteny-based phylogenomics in addressing the complexities of angiosperm evolution, particularly in the context of polyploidy and sheds a light on untangling unresolved key evolutionary episodes in plant. Additionally, it provides a comparative framework for characterizing newly sequenced Fabales genomes.

S8-P06-Recombinative DNA Deletion Dominates Gene Loss After Polyploidization in Angiosperms

Zhen Li

VIB-UGent Center for Plant Systems Biology

Large-scale structural reorganization is a hallmark of genome evolution in angiosperms, especially following whole-genome duplications (WGDs). Although extensive gene loss often accompanies rediploidization, the molecular mechanisms underlying this process remain unclear. Here, we compare the outcomes of pseudogenization versus recombinative DNA deletion in 12 angiosperms that underwent ancient polyploid events. Across these species, we find far fewer WGD-derived pseudogenes than expected if pseudogenization were the primary route of gene loss. To evaluate potential underestimation, we simulated neutrally evolving (pseudo)genes and examined three neo-autopolyploid genomes retaining much of their duplicated structure. Our simulations predict that pseudogenes should remain detectable over longer timescales in the absence of large-scale DNA deletion, yet these predictions do not match the scarcity of observed pseudogenes in real paleo-polyploid genomes. Moreover, significant gene loss has already occurred in the neo-polyploids, with only ~40% of these events attributed to pseudogenization. Thus, recombinative DNA deletion appears to predominate during the rapid rearrangement phase that follows WGDs. Interestingly, the pseudogenes we identified are enriched in low-recombination regions and sometimes show non-neutral signatures, suggesting that some may retain regulatory functions. Collectively, our results highlight a central role for recombination-driven deletion in shaping genome architecture and gene contents after WGD, underscoring how structural variation gives birth to adaptive potential in newly formed polyploids.

S8-P07-Temporal genomics in Hawaiian crickets reveals compensatory intragenomic coadaptation during adaptive evolution

Xiao Zhang

College of Life Sciences, Tianjin Normal University

Modern adaptation genomics almost exclusively focuses on characterizing genomic hotspots of adaptation, rather than compensatory genetic changes mitigating negative pleiotropy and other indirect effects of adaptive variants. We addressed this using temporal population genomics in a wild population of Hawaiian crickets undergoing contemporary adaptation. A mutation, flatwing, silences males and rapidly spread under selection from an acoustically-orienting parasitoid. Our sampling spanned a major social transition during which the last singing male died and the population went silent, which we predicted would cause genome-wide selection. By generating a high-quality chromosome-level cricket genome and resequencing data, we found that long-range linkage disequilibrium around the putative flatwing locus was maintained over time, and hitchhiking genes had functions related to known flatwing-associated effects. We developed a combinatorial enrichment approach using gene expression data to test for compensatory, intragenomic coevolution provoked by the spread of flatwing. Temporal changes in genomic selection were distributed genome-wide and functionally associated with the population's transition to silence. Selected regions were enriched for genes expressed during immune challenge by parasitoid infestation and behavioural plasticity to silent environments. Our results demonstrate how 'adaptation begets adaptation' and suggest that feedback from the sociogenetic environment is an overlooked component of rapid adaptive evolution.

S8-P08-Independent hybridization events, followed by genomic rearrangements, shape genome evolution of polyploid A.kamchatica

Maria Vasilarou

Max Plank for Plant Breeding Research

Polyploidy is an enduring force in plant evolution and speciation. In allotetraploids, the fusion of two divergent genomes, known as subgenomes, in a single nucleus can lead to dynamic interactions and genomic changes. Arabidopsis kamchatica is an established allotetraploid with multiple origins, resulting from independent hybridization events between A. lyrata and A. halleri. With a considerably wider distribution than its parental species, A. kamchatica offers a unique opportunity to examine the population genetic dynamics that enable its adaptive potential. This work delves into how A. kamchatica's complicated evolutionary history, along with independent hybridization events, has shaped its genome. We investigate the distribution, origins, and consequences of large genomic rearrangements by utilizing population-level resequencing data and de-novo genome assemblies from different populations, across the specie’s range. In particular, we identified a large-scale, non-homologous translocation between the subgenomes of the East-Asian chromosome-level assembly. Our findings suggest that this translocation may be unique to this assembly, with ongoing analyses aimed at determining its evolutionary significance. By comparing the distribution of structural rearrangements across different populations, we examine whether they are private to particular lineages and to what extent they are linked to the historical hybridization events. Moving forward, we are investigating, both computationally and experimentally, whether candidate rearrangements might contribute to reproductive isolation, particularly focusing on their meiotic behavior. By integrating population-level investigations, comparative genomics and cytological validation, we aim to elucidate the role of chromosomal rearrangements in shaping evolutionary trajectories within complex allopolyploid systems.

S8-P09-Comparative genomics show convergent evolution in the conquest of land by animals

Jialin Wei

University of Bristol

The ecological transition of life from water to land represents an iconic episode in the evolutionary history of a number of animal lineages, from the diversity of insects to the gamut of land vertebrates. Recent research has uncovered genomic changes underlying this transition in different animal lineages. However, the role of convergent evolution at genome level in these diverse events across all the Animal Kingdom remains unexplored. In this study, we compared 154 high-quality complete genomes to reconstruct the ancestral protein-coding content linked to multiple terrestrialisation events and produced a timescale of terrestrialisation. Here we show different patterns of gene gain and loss underlying each transition to land, but similar biological functions emerged recurrently pointing to specific traits as key adaptations to life on land. Additionally, semi-terrestrial species share convergent functional patterns, in contrast with fully terrestrial lineages which developed independent biological functions. Our timeline supports three episodes of land colonisation by animals during the history of Earth. These results strongly support the role of convergent genomic change as one of the key drivers in animal terrestrialisation, and draw an integrative picture of terrestrialisation in animals from genes to ecosystems.

S8-P10-Subgenome Asymmetry and Diversification in an Allotetraploid Cavefish: Structural Variation as a Catalyst for Extreme-Environment Adaptation

Santiago Montero-Mendieta

Institute of Zoology, Chinese Academy of Sciences

Cave environments impose extreme selective pressures that drive radical genetic and phenotypic adaptations in cave-dwelling organisms. In the allotetraploid cavefish Sinocyclocheilus microphthalmus (endemic to Guangxi, China), these pressures intersect with the genomic complexity of polyploidy, providing dual subgenomes that may respond asymmetrically to selection. Using whole-genome resequencing of 47 individuals across six isolated cave populations, we dissect how genome architecture, selection, and neutral processes interact to shape adaptation in this extreme environment. Our population structure analyses reveal pronounced genetic differentiation, consistent with allopatric divergence in fragmented caves. We identify seven single outlier clusters (SOCs) linked to local adaptation, with signatures of balancing selection and elevated linkage disequilibrium. Although we observe no substantial genome-wide divergence between subgenomes, asymmetric evolution within SOCs and among homoeologs suggest divergent evolutionary trajectories. Notably, one subgenome often retains ancestral functions (e.g., core cellular processes under purifying selection), while the other accumulates cave-associated innovations, particularly in vision-related genes due to relaxed selection. Altogether, these findings show how allotetraploidy has created a genomic "playground" for extreme-environment adaptation, where cave-driven selection leverages structural variation to balance redundancy with innovation.

S8-P11-Unraveling the DNA Methylation Landscape of New Gene Duplications

Michael Olufemi

University of Massachusetts, Lowell

Gene duplication can facilitate adaptation to changing environments. Most new duplicate genes are expected to disrupt stoichiometric relationships via gene dosage changes, often resulting in their loss over time through purifying selection. How some genes overcome these dosage effects and become preserved in the genome, remains unclear. It has been observed in several species that many new gene duplications initially display a decrease in expression levels. This decreased expression may be facilitated by epigenetic modifications like DNA methylation, which may rapidly balance dosage and allow new duplicates to persist as gene CNVs in a population. In this study, we leveraged high-quality genome assemblies and nanopore sequencing data from different stickleback fish ecotypes to investigate the interplay between gene copy number and DNA methylation. Additionally, we explored tissue-specific methylation patterns across multiple tissues to uncover potential epigenetic mechanisms underlying the regulation and persistence of gene duplications. By revealing differential methylation of gene duplications among individuals from different ecotypes that have adapted to different environments, our findings shed light on the role of epigenetic modifications in enabling duplicate genes to survive "hidden" from selection and contributing to local adaptation. These results expand our understanding of the evolutionary and tissue-specific dynamics that govern gene duplication and their potential to influence adaptation and speciation.

S8-P12-Polymorphic genomic inversions illuminate rapid adaptation in silent cricket populations

Shangzhe Zhang

University of St Andrews

Chromosomal inversions are structural variants associated with phenotypic variation and local adaptation across many organisms. However, their role in rapid adaptation remains poorly understood. To address this gap, we examined the population of Hawaiian field crickets (Teleogryllus oceanicus), a well-established system in which a male-silenced "flatwing" morphology arose and spread quickly due to intense selection pressure from a lethal parasitoid fly. Here, we leveraged diverse genomic resources—including a pan-genome and next-generation sequencing—to investigate the genetic basis of the flatwing adaptation, revealing a large (~150 Mb) polymorphic inversion on the X chromosome that widespread in Hawaiian crickets and involved in the flatwing trait. Population genomic analyses also shed light on the inversion's temporal dynamics and the selective forces shaping its trajectory during the adaptive process. Our findings provide a model for understanding the dynamics of genomic inversions in the context of rapid adaptation and their influence on demographic trajectories. Furthermore, they offer insights into how species respond to severe environmental pressures in natural settings.

S8-P13-Comparative Genomics of Arthropods: Insights into Genome Evolution across Diverse Lineages

Sean Chun-Chang Chen

Taipei Medical University

The phylum Arthropoda is the largest and most numerous on the planet, exhibiting remarkable diversity and inhabiting a wide range of environments. While insect genomes have been extensively studied, our understanding of patterns of genome evolution across the Arthropoda have remained limited. Here, we leveraged a comprehensive dataset of arthropod genomes from all subphyla, including a substantial sampling of crustaceans, to investigate patterns of arthropod genome architecture evolution. Our analyses revealed substantial variation in genome size, percentage of coding sequence, and nucleotide composition across arthropod lineages. We observed significant correlations between effective population size (Ne) and key genomic features, with smaller Ne associated with larger and less compact genomes. Furthermore, while most of the arthropod genomes were AT-rich, reduced AT-richness observed in coding regions, especially at the first codon positions, suggested increased selective constraints in these genomic regions. We also observed a positive correlation between chromosome number and GC content, for a given genome size, consistent with the role of recombination (and GC-biased gene conversion) in shaping nucleotide composition. Additionally, we identified lineage-specific gene family expansions that likely contribute to the ecological success of different arthropod groups. This study provides a crucial foundation for understanding the evolutionary forces shaping the diverse genomes of this ecologically dominant phylum.

S8-P14-Asian noodlefishes vs. Antarctic icefishes: alternative evolutionary trajectories to deterioration of the oxygen-transport system

Yu-Long Li

CAS Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences

Determination of evolutionary mechanisms underlying innovative traits is crucial for understanding the vast diversity of species and phenotypes across the tree of life. Given their diverse life histories and respiratory physiologies, fishes are compelling subjects for evolutionary analysis of the hemoprotein-based oxygen-transport systems, which have long been assumed to be essential to vertebrate life. Here we report the genome sequences of 11 species of the pedomorphic Asian noodlefishes (Osmeriformes: Salangidae), which generally do not express the respiratory pigments myoglobin and hemoglobin. By performing comparative genomic analysis, we document the complete loss of the myoglobin gene in Asian noodlefishes, which occurred by DNA transposon-mediated insertional mutagenesis prior to diversification of this clade. We also show that genes encoding the α- and β-chains of hemoglobin, the hemoglobin-clearance protein haptoglobin, and the erythrocyte transcription factor hemogen deteriorated independently in noodlefish genomes, possibly due to relaxed purifying selective pressure. By contrast, genes involved in angiogenesis and in heart development were under positive selection, consistent with compensatory evolutionary adaptation of the oxygen-transport axis. Results and prior research also show that deletion and disruption of hemoprotein-axis genes of noodlefishes and of the erythrocyte-null Antarctic icefishes were caused by different evolutionary processes but likely shared a common molecular mechanism based on insertional mutagenesis. We propose that historical contingency and neoteny were the evolutionary agents that led to the unusual oxygen-transport phenotypes of Asian noodlefishes.

S8-P15-Inversion Nation: How Chromosomal Rearrangements Shape Local Adaptation in the Flat Periwinkle

Le Qin

Choo Simon F.S. Li Marine Laboratory, School of Life Sciences, Chinese University of Hong Kong

Chromosomal inversions are often linked to local adaptation in various organisms. In Sweden, two ecotypes of the flat periwinkle Littorina fabalis have been observed: a small ecotype on sheltered parts of the shore and a large one on moderately exposed parts. These ecotypes differ by 12 putative inversion polymorphisms, which show strong linkage disequilibrium and a notable deficiency of heterozygotes. The ecotypes appear to have diverged in allopatry before coming into secondary contact. As these small and large ecotypes have been described previously from sites in the UK, France and Swedish, we hypothesised that these inversions are widespread across L. fabalis populations in northern Europe and Iberia and contribute to adaptation to similar environmental pressures. To test this, we used haplotagging—an efficient linked-read method for large populations—along with Illumina whole-genome sequencing. Our results suggest that the inversions are widespread across the range of L. fabalis except in Iberia. As a next step, we investigate the association of inversion polymorphisms with the environment and phenotype of L. fabalis populations, and whether the inversions contribute as a barrier to gene flow. By comparing inversion patterns across populations, we aim to understand the role of these inversions in local adaptation and speciation throughout their species range.

S8-P16-Evolutionary genome dynamics in asexually reproducing mite species

Mohammed Errbii

University of Cologne

Sexual reproduction is often viewed as essential for species evolution and long-term persistence. Yet, the existence of ancient asexual organisms, such as certain oribatid mites, challenges this view. How these organisms cope with challenges associated with asexuality remains unclear.

Here, we generated high-quality genome assemblies for eleven oribatid mites, spanning four independent transitions to asexuality and sexual relatives, to examine genome dynamics associated with asexuality.

Our analyses confirm that asexual mites do not exhibit an increased mutation load, indicating that selection remains as effective as in sexual species. Moreover, the genomes of asexual species are not enriched in rearrangements and contain fewer transposable elements compared to their sexual counterparts, possibly due to their large effective population sizes (Ne) and/or the upregulation of DNA repair mechanisms.

Using PacBio HiFi population re-sequencing, we explore whether Ne differs between sexual and asexual species and analyze segregation dynamics of short- and large-scale genomic rearrangements at the population level. By combining comparative and population genomics, this project will provide insights into how ancient asexuals escape the predicted evolutionary dead-end and how the release from the constrains of sex might offer evolutionary advantages.

S8-P17-Whole-genome duplication and its role in anhydrobiosis in fungal-feeding nematode Aphelenchus avenae

Wen-Sui Lo

Northwest A&F University

One of the challenges in evolutionary biology is understanding how genome architecture changes drive adaptation to novel environments. Aphelenchus avenae, a fungal-feeding nematode with robust drought tolerance, provides a unique model for studying these processes. Here, we present a telomere-to-telomere genome assembly of A. avenae, revealing an autopolyploid whole-genome duplication (WGD) event expanded its genome to 16 chromosomes. Conserved nematode genes in A. avenae largely cluster on 10 chromosomes and have remained under purifying selection after the WGD. In contrast, putative species-specific genes, characterized by frequent fragment gains or losses and high non-synonymous substitution rates, are located in separate chromosomal regions. This distinction between conserved and species-specific genes has not been observed in other sequenced nematodes. To explore how genes contribute to stress tolerance, we profiled gene expressions across 13 time points as the nematodes gradually entered anhydrobiosis. Conserved genes, such as lea genes, and trehalose biosynthesis genes, which are important for drought resilience, may provide individual benefits through dosage effects. In contrast, species-specific genes, some of which may have undergone subfunctionalization or emerged as novel genes, exhibit expression shifts during anhydrobiosis. Future functional studies will explore how genes with distinct evolutionary fates, shaped by whole-genome duplication and varying selective pressures, contribute to A. avenae's drought tolerance. Our work highlights the evolutionary significance of WGD and newly evolved genes in promoting adaptation to extreme environments.

S8-P18-Parallel sympatric divergence driven by disruptive selection acting on the magic reproductive timing trait in a wild fish

Hao Yang

Institute of Oceanology, Chinese Academy of Sciences

Sympatric speciation has been of key interest among biologists investigating how selection drive speciation without geographic isolation. However, the evolutionary trajectories and genetic architectures that underlie sympatric ecological divergence remain poorly understood. For the annual fish, Neosalanx brevirostris, from lakes in the Yangtze River basin, two sympatric ecotypes exhibit differences in reproductive timing, with one breeding in spring and the other in autumn. Reproductive timing is considered a 'magic trait' in sympatric speciation, as it is both ecologically relevant and facilitates assortative mating. Using population genomic approaches, we investigated the genetic mechanisms underlying rapid reproductive timing divergence by comparing two breeding ecotype pairs from two lakes and their ancestral anadromous population. Population structure results provided strong evidence that sympatric divergence has evolved independently and repeatedly in both lakes. Furthermore, we identified a set of genome-wide candidate adaptive loci, which were present as standing genetic variation in ancestral population with high frequencies, and quantitative results supported disruptive selection driving the reproductive timing divergence. These adaptive loci were involved in various biological functions, including photosensory, circadian entrainment, temperature sensing and hormone signaling, which closely aligned with the polygenic architecture of adaptive divergence between the two breeding ecotypes involving multiple complex physiological traits. These results provide valuable insights into the genetic architecture underlying the sympatric formation of ecological species.

S8-P19-Recurrent allopolyploidy outruns rediploidization in Potamogeton aquatic plants

Nikita Tikhomirov

Max Planck Institute for Plant Breeding Research

The extent and rate of genomic changes following polyploidy are extremely variable. To achieve a more general view on the interplay between polyploid genome evolution and various life history aspects, we study the genomes of Potamogeton, the biggest aquatic plant genus on Earth with diverse and broadly distributed diploid, tetraploid and octoploid lineages. Our dataset includes 15 long-read genome assemblies at all three ploidy levels as well as 433 resequenced samples, together representing all of the major species groups in the genus.

The analysis of gene tree topologies proved both tetra- and octoploids to have hybrid (allopolyploid) origin. Genus-wide phylogenetic analysis placed putative ancestors of the tetraploid lineage as sister to the two highly divergent present-day lineages. For the tetraploid lineage, we have separated the subgenomes using specific sequence motifs and dated the polyploidization event at 2–5 million years ago using LTR transposons insertion times, with parental genomes having diverged 6–23 million years ago according to the fossil-calibrated phylogeny.

Despite the age, both tetraploid and octoploid Potamogeton genomes harbor many duplicated BUSCO genes, indicative of the slowly ongoing rediploidization process. We have detected few major genomic rearrangements and no substantial downsizing in the polyploid genomes, which agrees with frequent interspecific hybridization in the genus and genomic traces of preferential outcrossing in the tetraploids.

With this valuable genomic resource, we are currently investigating the lineage-specific rediploidization patterns and the role of interspecific gene flow in the establishment of Potamogeton polyploids.

S8-P20-Chromosomal inversions and the genetic architecture of parallel adaptation

Maddie James

The University of Queensland

When populations diverge in the face of gene flow, there is a tug-of-war between divergent natural selection and the homogenising effects of genetic recombination. This antagonism can be resolved when locally adaptive loci reside within regions of reduced recombination such as chromosomal inversions, which facilitate population divergence. How chromosomal inversions contribute to parallel evolution, especially whether different structural variants can achieve similar adaptive outcomes, remains an area of intense exploration. In this talk I will show how inversions have played a pivotal rule during parallel adaptive trait evolution in contrasting ecotypes of an Australian wildflower, Senecio lautus. I will first overview how we are using a Multiparent Advanced Generation Inter-Cross (MAGIC) population to break down genetic linkage and trait correlations, enabling us to directly uncover the genetic architecture of parallel adaptive traits in the system. I will show that inversions play an increasingly important role in adaptive divergence as the rates of gene flow between populations intensifies. In addition, I will reveal how inversions do not always arise in the same genomic locations across replicate populations, yet capture sets of functionally equivalent genes. Our research furthers our understanding of the deterministic and stochastic nature of natural selection, and highlights how genome architecture shapes evolutionary trajectories during the early stages of ecotypic divergence.

S8-P21-Evolutionary innovations for marine life: Genomic insights from intertidal spiders

Fan Li

Peking University

While the majority of spiders display remarkable adaptability across terrestrial environments, the obligate intertidal species in the genus Desis have evolved unique adaptations for coastal life. Despite their intriguing adaptations, the molecular mechanisms underlying their success in intertidal zones remain largely unexplored. Here, we employed an integrated multi-omics approach to explore the genomic landscape of Desis jiaxiangi, a prominent representative of intertidal spiders from China. We constructed a high-quality, chromosome-scale genome assembly for this species and uncovered a range of molecular adaptations supporting its specialized intertidal lifestyle. The findings shed light on the detailed molecular mechanisms enabling intertidal spiders to thrive in their unique habitats. Furthermore, this research highlights the potential of multi-omics approaches to advance spider biology, provides deeper insights into the evolutionary and ecological significance of these remarkable arachnids, and bridges important gaps between marine and terrestrial evolutionary processes.

S8-P22-Unequally Abundant Chromosomes and Unusual Collections of Transferred Sequences Characterize Mitochondrial Genomes of Gastrodia (Orchidaceae), One of the Largest Mycoheterotrophic Plant Genera

Hanchen Wang

Institute of Botany, the Chinese Academy of Sciences

The mystery of genomic alternations in heterotrophic plants is among the most intriguing in evolutionary biology. Compared to plastid genomes (plastomes) with parallel size reduction and gene loss, mitochondrial genome (mitogenome) variation in heterotrophic plants remains underexplored in many aspects. To further unravel evolutionary outcomes of heterotrophy, we present a comparative mitogenomic study with 13 de novo assemblies of Gastrodia (Orchidaceae), one of the largest fully mycoheterotrophic plant genera, and its relatives. Analyzed Gastrodia mitogenomes range from 0.56 to 2.1 Mb, each consisting of numerous, unequally abundant chromosomes. Size variation appears to have evolved through chromosome rearrangements followed by stochastic loss of "empty" chromosomes. The discovery of a hyper-abundant (~15 times intragenomic average) chromosome in two assemblies represents the hitherto most extreme copy number variation in any mitogenomes, with similar architectures discovered in two metazoan lineages. Asymmetric evolutionary consequences of heterotrophy are evident in transferred sequence components: despite drastically reduced intracellular plastome transfers convergent across heterotrophic plants, their rarity of horizontally acquired sequences sharply contrast parasitic plants, where massive transfers from their hosts prevail. Rates of sequence evolution are markedly elevated but not explained by copy number variation, extending prior findings of accelerated molecular evolution from parasitic to heterotrophic plants. Putative evolutionary scenarios for these convergence and divergence fit well with the common (e.g., plastome contraction) and specific (e.g., host identity) aspects of the two heterotrophic types. These idiosyncratic mycoheterotrophs expand known architectural variability of plant mitogenomes and provide mechanistic insights into their content and size variation.

S8-P23-Mutation rate and selection in the unusual duplicated mitogenomes of Australasian stingless bees (Apidae: Meliponini)

Genevieve Law

The University of Sydney

The overall architecture of the mitochondrial genome (mitogenome) is highly conserved, particularly in animals. However, there are exceptions: the mitogenomes of the Indo-Malay/Australasian (IM/AA) clade of stingless bees (Apidae: Meliponini) were recently discovered to exhibit long inverted repeats. Depending on the species, these mitogenomes contain two identical copies of either all, or some, mitochondrial genes. Here, we investigated the ratio of mitochondrial to nuclear mutation rate and selection pressures experienced by the amphimeric mitogenomes found in the IM/AA clade, in comparison to related stingless bees, in order to investigate links between mutation rate and the amphimeric mitogenome structure. This analysis was conducted in silico using assembled mitochondrial and nuclear genes from species taken from across the three main clades of Meliponini, with the use of several newly sequenced mitogenomes. Given the potential for mitogenome structure and mutation rate to impact co-evolved nuclear genes, we examined the possibility of mito-nuclear co-evolution in the stingless bees and the role it may have played in divergence within the IM/AA clade. Through this comparative examination of stingless bee genomes (both mitochondrial and nuclear), we contribute to an increased understanding of mito-nuclear co-evolution and speciation of the Australasian stingless bees and beyond.

S8-P24-The effect of chromosomal inversions on gene expression in Mimulus guttatus

Paris Veltsos

Evology, Evolution and Genetics Research Group, Biology Department, Vrije Universiteit Brussel, Pleinlaan 2, Brussels, Belgium

We comprehensively evaluate 64 inversions that segregate within a single natural population of the yellow monkeyflower (Mimulus guttatus). Nucleotide diversity patterns confirm low internal variation for the derived orientation (predicted by recent origin), elevated diversity between orientations (predicted by natural selection), and localized fluctuations (predicted by gene flux). Sequence divergence between orientations varies idiosyncratically by position, not following the "suspension bridge pattern" predicted if the breakpoints are the targets of selection. Genetic variation in gene expression is not inflated close to inversion breakpoints but is clearly partitioned between orientations. Like sequence variation, the pattern of expression variation suggests that the capture of coadapted alleles is more important than the breakpoints for the fitness effects of inversions.

S8-P25-Chromosomal Inversions Drive Genome Evolution and Population Diversification in Common Quails

Santiago Montero-Mendieta

Conservation and Evolutionary Genetics Group, Doñana Biological Station (EBD-CSIC), 41092 Seville, Spain

Cave environments impose extreme selective pressures that drive radical genetic and phenotypic adaptations in cave-dwelling organisms. In the allotetraploid cavefish Sinocyclocheilus microphthalmus (endemic to Guangxi, China), these pressures intersect with the genomic complexity of polyploidy, providing dual subgenomes that may respond asymmetrically to selection. Using whole-genome resequencing of 47 individuals across six isolated cave populations, we dissect how genome architecture, selection, and neutral processes interact to shape adaptation in this extreme environment. Our population structure analyses reveal pronounced genetic differentiation, consistent with allopatric divergence in fragmented caves. We identify seven single outlier clusters (SOCs) linked to local adaptation, with signatures of balancing selection and elevated linkage disequilibrium. Although we observe no substantial genome-wide divergence between subgenomes, asymmetric evolution within SOCs and among homoeologs suggest divergent evolutionary trajectories. Notably, one subgenome often retains ancestral functions (e.g., core cellular processes under purifying selection), while the other accumulates cave-associated innovations, particularly in vision-related genes due to relaxed selection. Altogether, these findings show how allotetraploidy has created a genomic "playground" for extreme-environment adaptation, where cave-driven selection leverages structural variation to balance redundancy with innovation.

S8-P26-Genome architecture and organisation as a source of heritable epigenetic variation

Maximilian Fitz-James

Department of Biology, University of Oxford

As well as specifying differences between cell types within an individual, epigenetic information is also a major source of variation between individuals within a population. While this epigenetic variation is often restricted to a single generation, it can in some cases become heritable by “epimutation”, raising questions about the role of such non-genetic information in adaptation and evolution.

The mechanisms and prevalence of these epimutations are not clear, and likely vary. In several cases across diverse organisms, epimutation is triggered by a transient genetic change: a mutation lasting only one or a few generations with effects that persist for much longer. In particular, copy number changes through deletion, duplication or polyploidy can lead to epigenetic changes that last for generations. This suggests a possible mechanism by which large scale genetic changes such as duplications or transpositions may be a source of heritable epigenetic variation, as well as a source of genetic novelty.

In one such case in Drosophila, transient heterozygosity of a regulatory locus lasting only a single generation leads to epimutation of another distant locus, resulting in heritable differences in histone modifications with effects on gene expression and phenotype. We demonstrate that this transient genetic change leads to an increase in chromatin contacts between the two distant loci, implicating 3D chromatin organisation in epigenetic inheritance. We artificially recapitulate these chromatin contacts using a synthetic biology system that is able to trigger epigenetic inheritance without any genetic change, providing a mechanism for this case of epimutation through trans effects.

S8-P27-Evaluating the Viability of Macrosynteny as a Phylogenetic Trait Using Lophotrochozoa

Pia Merete Eriksen

University of Oslo (UiO)

With each passing year, we increase the amount of chromosome-scale full-genome data available to us, allowing us to stretch the boundaries of modern molecular methods as far as they can go. However, before this wealth of genomic data, we relied exclusively on morphological analyses; comparing conservation of limb structure, developmental patterns, and other phenotypic traits. While strict morphological analyses of extant taxa have now taken a backseat to molecular analyses, new genomic data grants us opportunities to examine genomic morphology, by treating chromosomal macrosynteny and arrangement as phenotypic traits: karyo-morphology.

We explored the viability of macrosynteny as a phylogenetic trait in Lophotrochozoa, a superphylum comprising well-studied phyla such as Mollusca and Platyhelminthes, alongside under-studied phyla such as Gnathostomulida and Gastrotricha. By clade-wise incrementally building on macrosyntenic patterns across the superphylum, we were able to establish the effectiveness of macrosynteny as a phylogenetic trait for studying divergences in deep-time. While previous research has been conducted into the role of macrosynteny as an evolutionary character in groups, such as Ctenophora and Porifera, and, more recently, within Annelida, this research represents the first comprehensive and cohesive study of macrosynteny's role as a trait, and its congruence with conventional molecular hypothesis.

S8-P28-Novel Retrovirus Integrations in Bats - The enERVate Pipeline

Jorge Gonzalez Garcia

Trinity College Dublin

Mammalian genomes contain remnants of ancient viral infections, known as endogenous viral elements (EVEs), which serve as molecular "fossils", providing valuable insights into host-virus evolutionary dynamics over millions of years. Bats represent an exceptional model for the study of EVEs due to their unparalleled tolerance to viral infection, unique immune system adaptations, and their role as reservoirs for a diverse array of viruses, making them ideal for studying host-virus co-evolution.

Endogenous viral integrations provide a historical record of past infections, shedding light on viral lineages that certain bat species have co-evolved with and tolerated. This information can guide future viral surveillance efforts in key bat species and populations. For this purpose, we developed enERVate, a robust computational pipeline designed to identify, filter, and annotate integrations of endogenous retroviruses (ERVs) in host genomes.

In this study, we examined the chromosome-level genomes of 25 bat species, publicly available through the Bat1K and Genomes on a Tree (GoaT) resources. Our analysis uncovered an extensive diversity of endogenous retroviruses, including novel sequences not previously reported in mammals. We identified and classified over 10000 ERV fragments from 10 distinct retrovirus families, revealing that many of these sequences form phylogenetically distinct clades compared to extant retroviruses. These findings highlight a long and complex history of interaction between bats and retroviruses, showcasing unique patterns of viral integration.

Overall, our work provides critical insights into the evolution of bat retroviruses, advances the understanding of host-virus interactions, and underscores the unique genomic and ecological roles of bats in shaping viral diversity.

S8-P29-Study of Evolutionary Genomics of Swifts and Swiftlets (Apodiformes: Apodidae): A Comparative Genomic approach

Heiman Ho

Sun Yat-sen University

The role of genetics in species evolution and its impact on biodiversity is a central topic in evolutionary biology. Advances in sequencing technology now provide opportunities to generate high-quality reference genomes for a variety of species. However, in-depth investigations into how species adapt to diverse habitats through genetic mechanisms remain limited. Among animal taxa, avian species exhibit remarkable lifestyle diversity (aquatic, terrestrial, aerial), driven by specialized adaptations in body structure. Swifts (Apodiformes: Apodidae) are small, highly aerial birds, while swiftlets, which share similar morphology, typically inhabit caves and are renowned for producing nests using saliva. The genetic and evolutionary mechanisms underlying this phenomenon remain unclear. To explore this, we collected three species—Apus nipalensis, Cypsiurus balasiensis, and Aerodramus fuciphagus—and performed de novo sequencing using the PacBio HiFi platform. This yielded a total of 240.7 Gb of data, complemented by re-sequencing and Hi-C sequencing. The resulting genomes were assembled to a chromosomal level and annotated. To investigate their environmental adaptations, we incorporated reference genomes of Apus apus and Aerodramus brevirostris for comparative analyses. Our findings revealed two chromosomal fission events in C. balasiensis and one chromosomal fusion event along with two inversions in A. apus. Through these comparisons, we aim to identify signatures of adaptation related to flight performance, vision, hearing, and immune systems, shedding light on the genetic basis of their ecological and behavioural specializations.

S8-P30-A telomere-to-telomere genome assembly of cotton provides insights into centromere evolution and short-season adaptation

Guanjing Hu

ICR/AGIS, CAAS

Cotton (Gossypium hirsutum L.) is a key allopolyploid crop with global economic importance. Here we present a telomere-to-telomere assembly of the elite variety Zhongmian 113. Leveraging technologies including PacBio HiFi, ONT ultralong-read sequencing, and Hi-C, our assembly surpasses previous genomes in contiguity and completeness, resolving 26 centromeric and 52 telomeric regions, 5S rDNA clusters and nucleolar organizer regions. A phylogenetically recent centromere repositioning on chromosome D08 was discovered specific to G. hirsutum, involving deactivation of an ancestral centromere and the formation of a unique, satellite repeat-based centromere. Genomic analyses evaluated favorable allele aggregation for key agronomic traits and uncovered an early-maturing haplotype derived from an 11 Mb pericentric inversion that evolved early during G. hirsutum domestication. Our study sheds light on the genomic origins of short-season adaptation, potentially involving introgression of an inversion from primitively domesticated forms, followed by subsequent haplotype differentiation in modern breeding programs.

S8-P31-Detecting evolutionary constraint in Arabidopsis thaliana genes using a panel of diverse Brassicaceae genomes

Raúl Wijfjes

Faculty of Biology, Ludwig-Maximilians-Universität München, 82152 Planegg-Martinsried, Germany

The Brassicaceae family is a central pillar of evolutionary studies in plants, containing plant species showing a range of environmental adaptations and the model plant Arabidopsis thaliana. High-quality genome assemblies are available for many Brassicaceae, allowing us to measure evolutionary constraint, defined here as the probability of small mutations and larger structural variation being retained at a given genomic position between different species of the family relative to neutral expectations. Such information enables us to identify genomic positions likely involved in molecular function or trait evolution, that would not have been identified using intraspecific measures of constraint only. Despite its importance, our understanding of the degree of evolutionary constraint in the A. thaliana genome is limited, as previous efforts to measure it used a small number of genomes of mostly closely related Brassicaceae. To address this, we present a panel of Brassicaceae genomes tailored to detect evolutionary constraint with high power and resolution. To maximize the phylogenetic diversity represented in the panel, we collected tens of publicly available chromosome-level genome assemblies of a variety of Brassicaceae and complemented these with newly generated ones to cover more clades. We annotated protein-coding genes in each assembly and used conserved gene order to infer orthology groups, leveraging the chromosome-level contiguity of the assemblies. To demonstrate the utility of the panel, we show how it can be used to measure evolutionary constraint, and consequently identify regions of evolutionary significance, in A. thaliana genes.

S8-P32-Genetic variation in a polygenic trait under stabilizing selection and population structure

Juan Li

University of Vienna

We aim to understand polygenic variation across a subdivided population. Using the infinite-island model, we assume that the population consists of many demes that are symmetrically connected by migration. Within each deme, individuals exhibit a polygenic trait subject to stabilizing selection towards the same optimum. We investigate how genetic variation within demes and across the whole population depends on the mutation rate, effect sizes, the number of loci, and migration rate. Through diffusion approximation and simulations, we seek to identify the conditions under which a polygenic trait relies on either the same or different genetic variants across different demes.

S8-P33-The Red Queen's hypothesis of transcriptional start sites: hyper-mutational bias leads to transcription start site hopping in human protein coding genes

Yi Qiu

University of Toronto

It is generally assumed that major features in our genome, such as the transcriptional start site (TSS) of protein coding genes (PCG), are solely shaped by natural selection. However, some evidence suggests non-adaptive evolutionary forces can be important in shaping local genomic architecture. To study mutational bias around the TSS of human PCGs, we extracted single nucleotide polymorphism (SNPs) occurring in frequency of <1% in human populations from the gnomAD database as a proxy for baseline mutation rates. We aligned all human PCGs centred around the TSS and calculated mutation rates with single-nucleotide resolution. Surprisingly, we observe hyper-mutation of most mutation types at the TSS, despite TSSs preferentially having the pyrimidine/purine motif at -1/+1 positions. This signal was also present in RNAPII-transcribed long non-coding RNAs but absent in genes transcriptionally-silent in the germline, suggesting RNAPII-transcription-associated mutational bias. Since nucleotide composition at the TSS is not at equilibrium based on mutational bias alone, we hypothesized that it is either being maintained by selection or undergoing "local shifting". To understand the evolutionary forces at play, we performed site frequency analyses of rare SNPs at the TSS and various locations as controls for negative selection (start codon and synonymous codon) and neutral substitution (intergenic, non-synonymous codons and introns). We do not observe signs of negative selection at the TSS, suggesting these sites are evolving neutrally. Since TSS motifs are mutating away and not being maintained by purging selection, we propose the Red Queen's model for the turnover or "local shifting" of TSSs.

S8-P34-Impacts of Genome Architecture (Chromosomal Fusions) on Selection Response during Laboratory Experimental Evolution

Carol Eunmi Lee

University of Wisconsin-Madison

Understanding the predictability of evolutionary pathways and genetic mechanisms underlying parallel evolution can help uncover mechanisms of adaptive evolution. Recently, populations of the copepod Eurytemora affinis species complex have invaded many freshwater habitats throughout the world. During these invasions, distinct clades (sibling species) with different chromosome number display remarkable parallel evolution, with the same loci (and SNPs) repeatedly under selection during salinity transitions. In laboratory evolution experiments, we imposed salinity decline for 10-20 generations on selection lines from clades with 15 chromosomes (E. affinis proper, ancestral state) and 4 chromosomes (E. carolleeae, fused state). In both clades, ion transport-related genes were the predominant functional category under selection. However, selection responses differed radically between selection lines from the two clades, with strong signals of positive epistasis among beneficial alleles in the 15-chromosome clade but selection on standing variation in the 4-chromosome clade. In the 15-chromosome genome, beneficial alleles are concentrated near the telomeres, where recombination and positive epistasis could bring them together within genomes. Consistent with these results, theoretical simulations showed that under strong positive epistasis, higher chromosome number can increase parallelism between replicate selection lines. In contrast, in the 4-chromosome genome, beneficial alleles are already linked near fusion sites (especially centromeres) where recombination is low, such that selection acts on balanced polymorphisms of large linkage groups. Our results illustrate that genome architecture has profound impacts on how selection acts on populations.

(Funding: NSF IOS-2412790, DEB-2055356; Refs: Du et al. 2024, GPB; Stern&Lee, 2020, NatureEcol.Evol.; Stern et al. 2022, NatureComm.)

Symposium 11: Phylogenomics in the presence of gene flow: opportunities and challenges

S11-P01-Integrating Bayesian and Heuristic Tests of Introgression to Infer Phylogenetic Networks

Yang Yang

Beijing Forestry University

Interspecific hybridization is a common and significant phenomenon in plants, playing a fundamental role in evolution and speciation. Heuristic methods, such as the D-statistic (a.k.a. ABBA-BABA test), are frequently employed to detect gene flow between species. However, these methods have inherent limitations, including their inability to infer the directionality of gene flow and to distinguish ghost introgression from introgression between sampled non-sister species. In contrast, full-likelihood approaches, Bayesian Phylogenetics and Phylogeography (BPP), make comprehensive use of the information embedded in multilocus sequence data by leveraging both gene-tree topologies and branch-length information. Despite their advantages, these methods are computationally intensive, which restricts their broader application. This study focuses on a stepwise framework integrating heuristic methods with the full-likelihood approach (BPP) to infer phylogenetic networks in plants. By combining the strengths of both methodologies, this framework seeks to address the limitations of individual methods and provides a more robust understanding of reticulate evolutionary histories among plant species.

S11-P02-Surprising effects of differential loss in genome evolution

Nico Bremer

Heinrich Heine University Düsseldorf

Gene loss is an important process in genome evolution, though its power is often underestimated. If a gene is present at the root of a phylogenetic tree and can be lost in one lineage across the tree, it can potentially be lost in all, leading to gene extinction. Just before gene extinction, there will be one lineage that still retains the gene, generating a 'last-one-out' distribution. Such an isolated gene presence will emulate the result of recent lateral gene acquisition, even though its distribution was generated by loss. How probable is it to observe 'last-one-out' distributions in real data? Here we mathematically derive this probability and find that it is surprisingly high, depending upon the tree and the gene loss rate. Examples from real data show that loss can readily account for observed frequencies of last-one-out gene distributions that might otherwise be attributed to lateral gene transfer.

S11-P03-The Last Universal Common Ancestor was not a hyperthermophile

Antoine Corne

University of Bristol

Reverse gyrase, resulting from a gene fusion event of a SF2-like helicase and a type IA topoisomerase domain, is the only hyperthermophilic-specific enzyme known. As a result, it has been widely studied to understand the ecology and physiology of early life. Previous studies have been based on phylogenetic analyses of whole-genes, leading to conflicting views on how reverse gyrase emerged and evolved and whether it was present in the Last Universal Common Ancestor (LUCA). We investigated the evolution of reverse gyrase performing domain-level phylogenetic analyses instead. Our results reveal that the multidomain reverse gyrase gene-level phylogenies are misled due to the distinct evolutionary histories of its constitutive domains. Our results suggest that the fusion event that gave rise to the multidomain reverse gyrase gene occurred deep in the bacterial lineage, perhaps in the bacterial stem-lineage, after horizontal transfer of the helicase domain from Archaea, supporting a bacterial origin of hyperthermophily. The resulting multidomain reverse gyrase was then transferred back to Archaea. Consequently, the reverse gyrase originated after LUCA and, as such, there is no longer evidence for a hyperthermophilic LUCA.

S11-P04-Inference of Gene Flow between Species under Misspecified Models

Jun Huang

School of Biomedical Engineering, Capital Medical University, Beijing 100069, P.R. China

Interspecific gene flow plays an important role in adaptive evolution and species diversification, and it is a hot spot and a difficult point in the research of population genetics. Despite recent advances in genomics and statistical methods, it remains difficult to infer gene flow. Furthermore, one may have to estimate introgression rates and times under misspecified models due to the presence of sampling errors, the complex evolutionary history, and the limitations of analytical methods. Therefore, we use mathematical analysis and computer simulation to examine estimation bias when the model of gene flow is misspecified. Such misspecifications include misassigning introgression to incorrect lineages, assuming introgression in the wrong direction, and neglecting recombination events and ghost species. When gene flow occurs continuously through time but in the analysis is assumed to occur at a fixed time point, common evolutionary parameters such as species divergence times are surprisingly well estimated. When introgression is assumed to occur in the wrong direction, the time of introgression tends to be correctly estimated. Additionally, we found that recombination, at rates comparable to estimates from humans, has little impact on the estimation of parameters. Overall, our analyses suggest that the simple introgression model is useful for extracting information concerning gene flow and divergence, even when the model may be misspecified. However, for reliable inference of gene flow, it is crucial to include multiple samples per species, in particular, from hybridizing species.

S11-P05-Untangling the phylogenetic discordance in Pancrustacea

Ka Yan Ma

The Chinese University of Hong Kong

Pancrustacea, encompassing Crustacea and Hexapoda, holds significant ecological and economic importance, but its evolutionary history remains unclear due to unresolved deep-level phylogeny, particularly on the recalcitrant relationships within Allotriocarida. To resolve pancrustacean phylogeny and analyze the potential reasons for recurrent topological discordance, this study analyzed genomic data from 106 species representing all major extant pancrustacean lineages, focusing on class-level phylogeny. Results showed Xenocarida and Multicrustacea as artifacts of long-branch attraction (LBA), and supported Allotriocarida and Copepoda as sister groups. Low phylogenetic support at deep nodes of Allotriocarida + Copepoda revealed incomplete lineage sorting (ILS), contributing to topological incongruence. Additionally, 42 ancient introgression (AIN) events were identified among major clades. AIN, ILS, and gene tree estimation errors accounted for over 30% of gene tree variation. Furthermore, putative whole genome duplications (WGDs) were identified in ten crustacean species across three clades, impacting ortholog inference and potentially causing false groupings. This study identified the four factors contributing to Pancrustacea's evolutionary relationship inconsistencies, offering insights into the biological processes and errors affecting species tree accuracy. Future studies should consider these sources of incongruence for more robust and precise phylogenetic relationships.

S11-P06-Modeling the mosaic structure of bacterial genomes to infer their evolutionary history

Peter Arndt

Max Planck Institute for Molecular Genetics

The reconstruction of the chronology and phylogeny of bacterial evolution is a challenging endeavor, primarily due to the paucity of fossil records available for analysis. The analysis of bacterial genomes remains challenging because of large sequence divergence, the plasticity of bacterial genomes due to frequent gene loss, horizontal gene transfer, and differences in selective pressure from one locus to another. Therefore, taking advantage of the rich and rapidly accumulating genomic data requires accurate modeling of genome evolution. An important technical consideration is that loci with high effective mutation rates may diverge beyond the detection limit of the alignment algorithms used, biasing the genome-wide divergence estimates toward smaller divergences. In this article, we propose a novel method to gain insight into bacterial evolution based on statistical properties of genome comparisons. We find that the length distribution of sequence matches is shaped by the effective mutation rates of different loci, by the horizontal transfers, and by the aligner sensitivity. Based on these inputs, we build a model and show that it accounts for the empirically observed distributions, taking the Enterobacteriaceae family as an example. Our method allows to distinguish segments of vertical and horizontal origins and to estimate the time divergence and exchange rate between any pair of taxa from genome-wide alignments. Based on the estimated time divergences, we construct a time-calibrated phylogenetic tree to demonstrate the accuracy of the method.

S11-P07-Hybridization pattern between the brown bear and the polar bear reveals multi-layered genetic structure of the brown bear.

Yu Endo

National Institute of Genetics

Hybridization between species is observed across many taxa, but the dynamics across geographical space and timing remain poorly understood due to sampling limitations. The brown bear (Ursus arctos), widely distributed in the Northern Hemisphere, has hybridized multiple times with the polar bear (Ursus maritimus). Previous studies have shown that modern individuals from island populations retain more alleles from this admixture compared to those from continental populations, suggesting differential hybridization patterns. However, direct evidence detailing the changes in hybridization proportions over time has been lacking, despite the availability of ancient brown bear genomic data.

To elucidate how recent geographical hybridization patterns were shaped by the population demography of brown bears, we calculated $f4$ statistics using whole-genome sequencing data from 49 modern and 5 ancient brown bears. Our analysis revealed a higher proportion of polar bear alleles in modern brown bears from previously unreported island populations and in ancient brown bears compared to modern continental individuals. Furthermore, we found that modern continental brown bears were significantly influenced by post-glacial expansion from the Eurasian continent, unlike their island counterparts. These results suggest that the genetic structure of modern brown bears is multi-layered, reflecting multiple demographic events over time.

S11-P08-A 'Spectacular' Saga: Reconstructing the Evolutionary History of the Indian Spectacled Cobra (Naja naja)

Paulomi Dam Kanunjna

Indian Institute of Science, Bengaluru

Snakes are some of the most feared and revered animals, with innumerable references in folklore and pop culture alike. Cobras (Naja spp.) are some of the most iconic snakes with great medical significance. India is an abode to the spectacled cobra (Naja naja), one of the 'big four' medically relevant snakes with a country-wide distribution. These snakes exhibit remarkable geographic variability in morphology and venom composition. However, the evolutionary history and gene flow across geographically disparate populations have remained elusive. This study aims to unravel the phylogenetic history and population genetics of N. naja across its distribution in the Indian subcontinent by employing phylogenetic and phylogenomic frameworks. We aim to test the monophyly of the species, to evaluate the genetic diversity, population substructure, and gene flow, and to construct time-calibrated phylogenies to understand the geological events dictating their diversification and distribution patterns. The findings suggest that despite a wide distribution across different habitats and ecological conditions, N. naja is monophyletic, with island populations (from Sri Lanka) being different. Our results will further provide insights into local adaptation and historical demographic changes. The multi-faceted approach, leveraging recent advancements in omics technologies, presents an interesting avenue to decipher the evolutionary trajectory of N. naja.

S11-P09-Fine-Resolution Asymmetric Migration Estimation

Hao Shen

University of Chicago

The inference of gene flow has been a central task in spatial population genetics since its inception. Early research focused on the theoretical properties of migration networks with specific topologies, typically restricted to small networks. Advances in large-scale gene flow inference has been facilitated by using approximations to pairwise coalescent times using commute times in random walks (which can be computed using “resistance distances”; for example, as in the EEMS and feems software packages). These approaches enable rapid inference across networks with hundreds of demes. However, the commute time approximation breaks down in the presence of strongly asymmetric migration and in such cases solving the pairwise coalescence time exactly is preferable (Lundgren and Ralph 2019). Due to its high computational complexity O(d^6), standard solutions for exact pairwise coalescent times are limited to small scales in practice (e.g. <20 demes).To overcome these challenges, we introduce FRAME (Fine-Resolution Asymmetric Migration Estimation). FRAME employs techniques from computational linear algebra and solves the structured coalescent equation with computational complexity of O(d^4). By integrating this approach with penalized-likelihood methods in FEEMS, FRAME enables the inference of fine-resolution asymmetric migration patterns with significantly improved computational efficiency. The efficacy of our FRAME is validated through both equilibrium and non equilibrium simulations, and its application to several empirical datasets (North American gray wolves, white-crowned manakins, poplar trees, and ancient humans) reveals significant features of asymmetric gene flow.

Symposium 13: Synthetic approaches to evolutionary theory: A dialog between the past and the future

S13-P01-In situ directed evolution on native mammalian genome

Zeyu Chen

Gene Regulation Observatory, Broad Institute of MIT and Harvard

Directed evolution has been applied to enhance molecular function in context-specific scenarios and create new biomolecules.Yet, this method has not been widely used on the native genomic DNA to understand the base-level function of the mammalian genome. The major technical challenges remain lacking of base-level long-range mutagenesis tools on the native chromatin. Here we present a programmable platform for long-range, locus-specific hypermutation called helicase-assisted continuous editing (HACE). HACE leverages CRISPR-Cas9 to target a processive helicase-deaminase fusion that incurs mutations across large (>1000–base pair) genomic intervals. We applied HACE to identify mutations in mitogen-activated protein kinase kinase 1 (MEK1) that confer kinase inhibitor resistance, to dissect the impact of individual variants in splicing factor 3B subunit 1 (SF3B1)–dependent missplicing, and to evaluate noncoding variants in a stimulation-dependent immune enhancer of CD69. HACE provides a powerful tool to process in situ directed evolution on native mammalian genome, for investigating coding and noncoding variants, uncovering combinatorial sequence-to-function relationships, and evolving new biological functions.

S13-P02-Identification of Cancer Mini-Drivers by Deciphering Selective Landscape in the Cancer Genome

Xunuo Zhu

Zhejiang University

Traditional selective pressure analysis methods have treated all sites within a gene equally, overlooking the variation in selection pressures across different sites. Also, these methods predominantly focused on positively selected driver genes, categorizing the remainder as passengers. Such a dichotomous, gene-level model oversimplifies the complexity of cancer evolution. In this study, we introduced CN/CS-calculator, a site-specific molecular evolutionary analysis method that can capture selective pressures acting across different gene sites. By deciphering the interplay between a gene's selection pattern and its function in oncogenesis, this method uncovers a unique class of mini-driver genes, which exhibit weak positive selection, with certain critical sites providing context-dependent promoter effects on the fitness of cancer subclones while others are constrained by evolutionary conservation. This method emphasizes the importance of site-specific analysis in uncovering how subtle evolutionary forces shape cancer biology. The refined understanding offers new insights into the mechanisms of cancer heterogeneity and molecular evolution, with potential implications for advancing therapeutic strategies and prognostic assessments.

S13-P03-A comprehensive double-mutant scan of a steroid receptor DNA-binding domain

Benjamin Epley

Department of Ecology and Evolution, University of Chicago

Understanding epistasis, the non-additivity in the effects of substitutions, is central to the study of molecular evolution because it entrenches established changes, shifts the effects of mutations over evolutionary timescales, and shapes the network of accessible sequence states. Our study quantifies all main effects and pairwise genetic interactions from a complete library of all possible single and double amino acid mutations of two extant steroid receptor DNA-binding domains with distinct functions and their reconstructed ancestor. We find that significant pairwise epistatic interactions are uncommon and tend to be small in magnitude relative to main effects. Less than half of all pairwise interactions significantly change over each historical interval, suggesting few significant higher-order interactions. Our dataset, for the first time, allows us to determine all historical pairwise effects of substitutions along the evolution of a protein and, more broadly, uncover the role of epistasis over evolutionary history, clues to the prevalence of higher-order epistatic interactions, and the genetic architecture of specific pairwise epistatic interactions.

S13-P04-Exploration of the dynamic evolutionary rates of early animals

Zhiqing Guo

University of Bristol

The timing and rate of animal diversification remain subjects of contentious debate, with perspectives polarized between an early and late Neoproterozoic origin of animals. Rates of morphological evolution in the past can be inferred from phenotypic differences among living organisms-just as molecular evolutionary rates in ancestral lineages can be inferred from genetic divergences. To explore the implications of evolutionary innovation, we employed a Bayesian molecular clock model to estimate molecular evolutionary rate dynamics across three different timescales, using a single amino acid alignment comprising 30,031 positions from 203 nuclear coding genes. These results were compared to morphological rates estimated with a Bayesian morphological clock model based on a dataset of 2,232 phenotypic characters encompassing all aspects of animal anatomy. Our findings reveal that molecular evolutionary rates consistently increase as the Cambrian period approaches, peaking at its onset. However, morphological rates exhibit a pattern of acceleration prior to the Cambrian period. Moreover, molecular rates across different timescales show a similar trend, with a gradual increase as they approach the Cambrian period. These results indicate that molecular and morphological clocks capture fundamentally different dimensions of evolutionary history, with morphological innovations decoupled from genomic evolutionary rate shifts. This comparative framework underscores the complex interplay between genomic and phenotypic evolution during the early diversification of animals.

S13-P05-Engineering a Toxin-Antidote Dominant Sperm Gene Drive for Population Control in Flies

Bingke Jiao

IGDB, Chinese Academy of Sciences

Synthetic gene drives, which mimic natural selfish genetic elements exhibiting super-Mendelian inheritance, offer a promising solution for modifying or suppressing disease vectors such as mosquitoes and flies. Most available gene drives in insects rely on homing mechanism, which converts heterozygotes to homozygotes. However, this approach is highly dependent on the efficiency of the homology-directed repairing (HDR) pathway, which varies among species and cell types. An alternative gene drive design is to selectively eliminate sperms not carrying the gene drive, thereby ensuring only desired sperms survive or fertilize. Since this design bypasses the allele conversion step required in homing drives, it overcomes the reliance on the HDR pathway. In our study, we use a toxin-antidote model to realize the selective elimination of sperms. Biefly, we utilize CRISPR-Cas as a "toxin" to disrupt male sterile genes in the genome, while a recoded, CRISPR-resistant male sterile gene, expressed haploidly, serves as an "antidote" to rescue the otherwise non-viable sperm. This design, termed as Toxin-Antidote Dominant Sperm(TADS), has been realized in plants by two recent studies, but has not yet been implemented in any animals ,likely due to the short haploid stage in animal sperm and the presence of cytoplasmic bridges between spermatids that allow the sharing of the antidote. To address these challenges in animals, we propose modifying the promoters or untranslated regions(UTRs) of male sterile genes to adjust the timing of antidote expression and prevent the sharing of antidote RNA or protein through cytoplasmic bridges between sperms.

Symposium 14: The roles of hybridization and introgression in evolution

S14-P01-Haplotype-level assembly for a hybrid species Saccharomyces bayanus reveal substantial introgression from multiple species

Junhao Chen

Saint Louis University

Interspecies hybridization is prevalent in various eukaryotic lineages and plays important roles in phenotypic diversification, adaptation, and speciation. To better understand the changes that occurred in the different subgenomes of a hybrid species and how they facilitate adaptation, we have completed chromosome-level de novo assemblies of all chromosomes for a recently formed hybrid yeast, Saccharomyces bayanus strain CBS380, using Oxford Nanopore Technologies' MinION long-read sequencing. We characterize the S. bayanus genome and compare it with its parent species, Saccharomyces uvarum and Saccharomyces eubayanus, and other S. bayanus genomes to better understand genome evolution after a relatively recent hybridization event. We observe multiple recombination events between the subgenomes in each chromosome, followed by loss of heterozygosity (LOH) in nine chromosome pairs. In addition to maintaining nearly all gene content and synteny from its parental genomes, S. bayanus has acquired many genes from other yeast species, primarily through the introgression of Saccharomyces cerevisiae, such as those involved in the maltose metabolism. Finally, the patterns of recombination and LOH suggest an allotetraploid origin of S. bayanus. The gene acquisition and rapid LOH in the hybrid genome probably facilitated its adaptation to maltose brewing environments and mitigated the maladaptive effect of hybridization. This paper describes the first in-depth study using long-read sequencing technology of an S. bayanus hybrid genome, which may serve as an excellent reference for future studies of this important yeast and other yeast strains.

S14-P02-Patterns of interspecific introgression in ornamental and feralized Xiphophorus strains from southern China

Jialin Geng

school of ecology,SYSU

Adaptive introgression can propel the rapid evolution of species. Ornamental swordtails are the result of century-old hybridization among multiple ancestries, which is a precious system for understanding interspecific gene flow. We combine HMM-based local ancestry inference and ABC simulations to infer the hybridization parameters and identified candidate regions for adaptive introgression. Candidate regions involved in high-frequency introgressions are enriched for functions such as immunity, reproduction, survival, and life history traits. Moreover, the study provides some initial evidence for deleterious mutations and mito-nuclear interactions in shaping the genomic landscape of these hybrid strains, offering insights into their evolutionary dynamics in captivity.

S14-P03-Hybridization and incomplete lineage sorting in the evolution of Apiales

Jun Wen

Institute of Botany, Jiangsu Province and Chinese Academy of Sciences (Nanjing Botanical Garden Mem. Sun Yat-Sen)

Apiales, a key group of flowering plants extremely rich in species and taxonomically complex, contains some 5489 species. The order has both herbaceous and woody groups. There are seven families in the order, the two largest of which are Apiaceae and Araliaceae. These two families are of great interest because of their abundance of edible (carrot, parsley, and so on) and medicinal (Angelica, ginseng family, and so on) species. The diversity of species and morphological traits makes these two families favored. The speciation and evolution history of these groups are the focus of current research. However, the long-standing controversy in the two families concerning relationships among genera and species has prevented the comprehensive studies of their evolution. In recent years, we have carried out phylogenetic studies on the important groups of Apiaceae and the genus Hydrocotyle (a transitional group between Apiaceae and Araliaceae) by transcriptome (for extracting nuclear genes) and chloroplast genome data. The results showed that the incongruence between chloroplast-based and nuclear-based phylogenetic trees was common in the phylogenetic studies of these groups. Combined with the morphological studies and the divergence time estimation, we speculated that there should be widespread hybridization and introgression among Apiales species. The estimation of species trees based on the coalescent model suggested possible rapid evolutionary divergence and incomplete lineage sorting. Based on existing studies, we speculate that the diversity of species and morphological traits in Apiales is largely due to extensive hybridization and incomplete lineage sorting.

S14-P04-Multiple bovine species introgression contributes to the complex history and adaptation of East Asian cattle

Ningbo Chen

Northwest A&F University

East Asia is rich in domestic cattle resources and other species of the Bos genus, including yak, gaur, and extinct aurochs. Hybridization between cattle and other species occurs frequently. All these events contributed to the complex genetic background and adaptations of East Asian cattle. Here, we used modern and ancient genomes of these bovine species to reveal the origin of East Asian cattle and the historical and geographical distributions of domestic cattle, extinct aurochs, yak, and gaur. Furthermore, we identified evidence of pervasive introgression from East Asian aurochs, gaur, banteng and yak into ancient and present-day East Asian cattle from southern China and the Tibetan Plateau, suggesting their potential role in facilitating the environmental adaptation of domestic cattle. For Tibetan cattle, we provide evidence of hybridization of cattle, yak, and aurochs on the Tibetan Plateau from ~3750 years ago and found that recently selected genes and introgressed yak alleles jointly facilitated cattle rapid adaptation to high-altitude environments. We also found that the range of the present-day tropical gaur extended as far north as the margins of the northeastern Tibetan Plateau during the late Neolithic from ~29°N to ~34°N. Our results indicate that the rapid and successful adaptation of East Asian indicine cattle to hot, humid environments was promoted by localized introgression from banteng and/or gaur. Our findings provide new insights into the evolutionary history of cattle and the importance of introgression in the adaptation of cattle to new environmental challenges in East Asia.

S14-P05-Negative selection in females against heterozygosity on X chromosome drives the evolution of reproductive barrier during incipient speciation

Qinglong Li

Yunnan University

The evolution and impact of X chromosome in speciation is a cornerstone of evolutionary biology, with Haldane's rule widely regarded as a hallmark of the evolution of reproductive barrier. Using wild boars and Eurasian domestic pig as researchful materials, we reveal that the X chromosome promotes reproductive barrier formation in females via a unique mechanism that diverges from Haldane's rule during early speciation. Phylogenomic analyses of wild suids and gene flow assessments in wild boar populations pinpointed a 50-Mb X-linked region centered on the Xist locus, characterized by accelerated evolution and restricted gene flow. This faster-X region exhibits its most pronounced evolutionary acceleration between closely related wild boars and unexpected power in dividing wild boars into populations with geographical isolations, suggesting its role as a speciation island. The faster-X region is a recombination coldspot with enrichment of structure variations, which undergoes strong negative selection that significantly reduces female heterozygotes in both wild boars and F2 population in a Eurasian pig cross, indicating conserved evolutionary dynamics spanning from within-population processes to inter-population divergence. The X:autosomal interactions play a minimal role in the observed female heterozygote deficiency in F2 population. These findings provide critical insights into X chromosome evolution, underscoring its pivotal role in driving reproductive barriers through female-specific mechanisms.

S14-P06-Unlocking Asexual Survival in Hybrid Clariid Catfish (Bighead catfish (Clarias macrocephalus) and North African catfish (Clarias gariepinus)) through Understanding Meiotic Deviations and Endoreplication

Kornsorn Srikulnath

Kasetsart University

Reproductive isolation is essential for maintaining species integrity by preventing the production of offspring between different species. Yet, interspecific hybridization occurs frequently, leading to important evolutionary outcomes, such as the formation of new species and genetic exchanges. Although hybrid sterility often results from genomic incompatibilities, some hybrids are found to reproduce asexually, bypassing these challenges. This study investigates how asexual reproduction is achieved in clariid catfish hybrids, specifically focusing on hybrids of Asian and African clariid catfish (bighead catfish (Clarias macrocephalus) and North African catfish (Clarias gariepinus)). Male hybrids are rendered sterile due to widespread chromosome asynapsis. By contrast, female hybrids demonstrate potential fertility and produce viable backcross offspring. Meiosis, here, was analyzed in hybrid female catfish using immunostaining to detect chromosome pairing and FISH to identify parental chromosomes. Analysis shows that over 95% of oocytes in female hybrids exhibit chromosome asynapsis, yet they successfully develop into mature eggs. Additionally, some oocytes undergo premeiotic endoreplication, resulting in clonal reproduction. These findings illustrate how hybridization disrupts traditional gametogenesis, leading to new reproductive pathways. It is suggested that intrinsic mechanisms from sexually reproducing parental species facilitate this clonal reproduction among hybrids. This research enhances our understanding of the similarities between sexual and asexual reproduction and indicates that certain clonal lineages may emerge from the reproductive strategies of their parent species. Overall, the insights gained from this study have significant implications for evolutionary biology and hold promise for advancing practices in aquaculture.

S14-P07-Widespread remote introgression events in grass genomes

Yujie Huang

Zhejiang university

Genetic transfers are pervasive in the reticulated evolution of both prokaryotes and eukaryotes, which usually includes genomic introgression between phylogenetically close species within a genus, as well as the well-recognized horizontal gene transfer that takes place across kingdoms. However, remote introgression (RI), namely the gene transfer between divergent lineages above the genus level, remains largely unexplored so far. Here, we develop an integrated method named RIFinder, by combining phylogeny and comparative genomics approaches, to identify RI events. By analyzing 122 grass (Poaceae) genomes, we detected 622 RI events occurring between subfamilies, functionally enriched in stress resistance. A significant preference was observed among the subfamilies, with the Pooideae subfamily having the highest number of receiver genes while bamboo subfamily (Bambusoideae) had the lowest. Specifically, we observed a 30-Kbp genomic segment in the genome of the desert grass species Cleistogenes songorica (Chloridoideae) that was transferred from the Triticeae tribe (Pooideae), with genes associated with drought tolerance. Additionally, our results reveal that RI events have led to the dispersed distribution of gramine biosynthetic gene clusters across grass genomes. In short, our study provides an efficient tool for detecting potential RI events, demonstrates the potential significance of RI in the environmental adaption of grass species, and deepens our understanding of the diversity and complexity of reticulate evolution in plant species.

S14-P08-An autosomal genome informs about the evolutionary history of the Columbian mammoths in Mexico

Rigoberto Padilla-Bustos

UNAM

Paleogenomic studies suggest that Mammuthus columbi, the only mainland mammoth species endemic to the Americas, originated from an ancient hybridization between woolly mammoths (Mammuthus primigenius) and steppe mammoths (Mammuthus trogontherii).

While M. columbi's habitat ranged from North to Central America, available genetic data are limited to boreal and temperate regions of Canada and the USA. In fact, the only M. columbi nuclear genome available to date derives from an individual from Wyoming (USA). The discovery of Columbian mammoth remains in the Basin of Mexico provides a unique opportunity to further understand its evolutionary trajectory in the continent by adding information from tropical latitudes.

We screened 72 libraries and generated three low-coverage nuclear genomes. We jointly analyzed these data with other Columbian, woolly and steppe mammoths genomes from America and Eurasia. We perform f-statistics analyses to investigate the admixture dynamics of M. columbi to the other mammoth lineages. Additionally, we carried out a hidden Markov model approximation to identify genomic regions of putative M. trogontherii-like origin, and generate phylogenetic tree restricting to such introgressed regions. We observed that M. columbi samples from the USA and Mexico present different genetic affinities to American M. primigenius. Furthermore, we observe that Mexican samples display a similar admixture proportion of trogontherii-like regions to the Columbian mammoth sample from Wyoming.

Our study revealed a complex evolutionary origin of the Columbian mammoths and underscores the importance of wider spatial sampling to better comprehend M. columbi evolutionary history

S14-P09-Ignoring population structure in evolutionary models can lead to the inference of spurious admixture events

Lounes Chikhi

CNRS Toulouse

Genomic and ancient DNA data have revolutionized our vision of human evolution. They have also revolutionized our vision of admixture and hybridization across taxa. Yet, using genetic data to identify, date and quantify evolutionary events is not straightforward, as inferences may depend on model assumptions. In the last two decades, the idea that Neanderthals and members of the Homo sapiens lineage interbred, or that admixture is common across taxa has gained momentum. However, ancient admixture events are mainly supported by statistical approaches that depend on demographic models minimizing or ignoring intra-continental population structure, despite its indisputable occurrence and the fact that, when ignored, population structure can lead to the inference of spurious demographic events. We simulated genomic data under structured and admixture-free scenarios of human evolution, and found that all the tested admixture approaches identified long Neanderthal fragments in our simulated (admixture-free) genomes and an admixture event that never took place. We also observed that several published admixture models failed to predict important empirical diversity or admixture statistics, and that several scenarios from our structured model better predicted these statistics jointly. Our results suggest that admixture between Neanderthals and H. sapiens needs to be re-evaluated in the light of structured models. We also simulated structured models for chimpanzees and bonobos and found that admixture events identified between bonobos and chimpanzees, might also be false positives. This suggests that hybridization events, which are increasingly documented in many species, may benefit from such a re-evaluation.

S14-P10-DNA Methylation and transcriptomic patterns in hybridizing Neotropical Bromeliads

Clarisse Palma da Silva

Unicamp

Epigenetic variation as a mechanism for reproductive barriers and speciation is emerging, yet little is known about DNA methylation patterns in plants and hybrids. We investigate the association between differential gene expression and DNA methylation in reproductive isolation and assortative mating in Pitcairnia species (Bromeliaceae). We analyzed sympatric and allopatric populations, and hybrids between P. staminea and P. albiflos, using PacBio HiFi methylation patterns with the Jasmine software, pbmm2 and pb-CpG-tools, revealed global CpG methylation rate of 73.38% and 75.13% for P. albiflos and P. staminea respectively. Additionally, enzymatic methyl-seq (EM-seq) with Illumina short reads was analyzed with Bismark software for 59 individuals. No significant differences were found in global methylation levels across CpG, CHG, and CHH contexts. However, MethylKit revealed 5,846 differentially methylated regions (DMRs) in P. albiflos compared to P. staminea (3,872 hyper- and 1,974 hypomethylated). Sixty-one DMRs were exclusive to hybrids. Pure individuals from allopatric populations exhibited fewer DMRs than sympatric populations, suggesting that genomic interactions in hybridizing species may trigger genome shock, activating transposons. Most DMRs were in repetitive elements, comprising ~70% of both genomes. These findings suggest DNA methylation may contribute to reproductive isolation and speciation in these neotropical species. Future research should explore the functional roles of these regions in the evolution of RI and adaptation.

S14-P11-Extreme Levels of Genetic Variant Sharing in Southwestern United States Myotis Bats: Implications for Hybridization and Viral Adaptation

Genavieve Gray-Sandoval

University of Arizona

Bats are a natural reservoir for many viruses, including coronaviruses, filoviruses, and henipaviruses, making them an excellent model for studying virus-driven adaptation. Investigating evolutionary pressures exerted by these viruses offers valuable insights into mechanisms underlying bat resilience and viral tolerance. Recent research highlights hybridization's role in introducing genetic variation, particularly in genes associated with viral immunity. In Myotis bats, recurrent hybridization may facilitate the modulation of immune genes, including viral interacting proteins (VIPs), enhancing viral tolerance. Unlike previous gene flow studies in Myotis bats, which rely only on single genome assemblies, we performed large-scale resequencing and generated high-quality reference genomes. We conducted de novo whole-genome sequencing for nine Myotis species from the southwestern United States, producing near-chromosome level reference genomes used to call variants in ~20 resequenced individuals per species. Our analyses uncovered extreme genetic variant sharing, with distantly related species sharing up to ~50% of higher frequency SNPs, contradicting species phylogeny and highlighting extensive gene flow. Immune and non-immune-related loci displayed distinct patterns of genetic variation, revealing the interplay of selection and hybridization in shaping genomic diversity. Our approach establishes Myotis bats as a powerful model for studying hybridization, adaptation, and immune evolution. Their species richness, low divergence, and ecological diversity offer a robust framework for understanding how natural processes shape genetic variation. By integrating genome-wide analyses with modified adaptation detection tools, our study advances knowledge of hybridization's genomic consequences and the role of viral pressures while providing insights into bats' viral tolerance and role in zoonotic disease dynamics.

S14-P12-Contribution of homoeologous exchange to domestication of polyploid Brassica

Tianpeng Wang

University of California, Irvine

Background: Polyploidy is widely recognized as a significant evolutionary force in the plant kingdom, contributing to the diversification of plants. One of the notable features of allopolyploidy is the occurrence of homoeologous exchange (HE) events between the subgenomes, causing changes in genomic composition, gene expression, and phenotypic variations. However, the role of HE in plant adaptation and domestication remains unclear.

Results: Here we analyzed the whole-genome resequencing data from Brassica napus accessions, representing the different morphotypes and ecotypes, to investigate the role of HE in domestication. Our findings demonstrate frequent occurrence of HEs in Brassica napus, with substantial HE patterns shared across populations, indicating their potential role in promoting crop domestication. HE events are asymmetric, with the A genome more frequently replacing C genome segments. These events show a preference for specific genomic regions and vary among populations. We then identify candidate genes in HE regions specific to certain populations, which likely contribute to flowering-time diversification across diverse morphotypes and ecotypes. In addition, we assemble a new genome of a swede accession, confirming the HEs signals on the genome and their potential involvement in hypocotyl-root tuber development. Finally, by analyzing HE in another allopolyploid species, Brassica juncea, we characterize a potential broader role of HE in allopolyploid crop domestication.

Conclusions: Our results provide novel insights into the domestication of polyploid Brassica species and highlight homoeologous exchanges as a crucial mechanism for generating variations that are selected for crop improvement in polyploid species.

S14-P13-A comparative assessment of methods to estimate gene flow and adaptive introgression using individual-based simulations

Yulin Sun

The University of Queensland

Gene flow, the exchange of genetic material between populations or closely related species, is widespread in nature and plays a crucial role in speciation and adaptation. Building on advances in population genetics and numerous empirical studies on gene flow in parallel evolution, the interplay between gene flow and natural selection in shaping ecotypes continues to be actively explored. Given the prevalence of gene flow and introgression in nature, it is essential to review and evaluate widely used methods for detecting and characterising these processes. Based upon an extensive literature review, I developed a catalogue of ~25 methods for characterising gene flow and introgression, grouping them into five categories based on their underlying principles, including gene tree frequencies, coalescent theory, and phylogenetic network. Simulation, where the ground truth is known, provides valuable means to assess whether these methods accurately characterise gene flow, especially at low levels. However, research gaps persist due to the use of inconsistent simulation approaches in previous method evaluations, and the comparison between only two or three methods at a time. In this study, I used an individual-based forward genetic simulation tool, SLiM, to conduct a comparative assessment of the methods selected from the catalogue to test their performance under different evolutionary scenarios. The results reveal that most methods perform poorly at low levels of gene flow, with their accuracy improving as the gene flow level increases. These findings provide insights into selecting appropriate methods to study how gene flow shapes adaptation and speciation in natural systems.

S14-P14-Selection increases mitonuclear DNA discordance but reconciles incompatibility in African cattle

Xian Shi

Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China.

Mitochondrial function relies on the coordinated interactions between genes in the mitochondrial (mtDNA) and nuclear genomes. Imperfect interactions following mitonuclear incompatibility may lead to reduced fitness. MtDNA introgressions across species and populations are common and well documented. Various strategies may be expected to reconcile mitonuclear incompatibility in hybrids or admixed individuals. African admixed cattle (Bos taurus × B. indicus) show sex-biased admixture, with taurine (B. taurus) mtDNA and a nuclear genome predominantly of humped zebu (B. indicus). The legacy of multiple male-mediated introgressions over time within the continent persists. Here, we leveraged local ancestry inference approaches to discern the ancestry and distribution patterns of nuclear functional genes associated with the mitochondrial oxidative phosphorylation process in the genomes of African admixed cattle. We show that most of the nuclear genes involved in mitonuclear interactions are under selection and of humped zebu ancestry. Variation in mtDNA copy number (mtDNA-CN) may have contributed to the recovery of optimal mitochondrial function following admixture with the regulation of gene expression, alleviating or nullifying mitochondrial dysfunction. Interestingly, some nuclear mitochondrial genes with enrichment in taurine ancestry may have originated from ancient African aurochs (B. primigenius africanus) introgression. They may have contributed to the local adaptation of African cattle to pathogen burdens. Our study provides further support and new evidence showing that the successful settlement of cattle across the continent was a complex mechanism involving adaptive introgression, mtDNA-CN variation, regulation of gene expression, and selection of ancestral mitochondria-related genes.

S14-P15-Correlation between Hybrid Ancestry and Ecological Niches in Invasive Tilapias from the Dongjiang River Basin in South China

Jianyu Zhang

SUN YAT-SEN UNIVERSITY

Hybridization can be an innovative process, allowing species to occupy new ecological niches. Tilapia is a commercially important group of African cichlids, which frequently hybridize in nature. Multiple species and their hybrids in the genera Oreochromis, Sarotherodon and Coptodon have established populations in diverse habitats in southern China and constituted a widespread biological invasion. We collected tilapia and environmental data from ~60 small river habitats in the lower reaches of the Dongjiang River Basin in Guangdong Province. Over one thousand tilapia were subjected to genome sequencing and phenotyping. We conduct hybrid ancestry inference, genome-wide association, and morphological analysis to investigate the correlations between ancestry, phenotypes, and habitat types. We expect our study to shed light on the role of hybridization in the mechanism of invasive species.

S14-P16-A refined landscape of introgression in great apes

Martin Kuhlwilm

University of Vienna

Great apes experienced a history of admixture, which in some respects resembles what we know about modern humans and their interactions with ancient hominins such as Neanderthals or Denisovans. Chimpanzees and bonobos most likely experienced bi-directional gene flow, and western gorillas from the species of eastern gorillas. Furthermore, bonobos and eastern gorillas received additional gene flow from extinct lineages. However, few genomes are available for these populations, and methods to detect introgressed fragments are often tailored to modern human genomes, which limits our understanding of these events. Here, we compile and curate a comprehensive dataset of high-coverage whole genomes from 77 chimpanzees, 14 bonobos and 77 gorillas published previously across multiple studies, thus leveraging the available knowledge on genomic variation of these species, from wild and captive individuals. We devised coherent pipelines to infer introgressed fragments in these non-human species, including a novel implementation of machine learning algorithms for detecting introgression without data for the source population. Using this state-of-the-art methodology, we refine estimates of introgression landscapes for the admixture events these species experienced. A cross-comparison between these species and archaic introgression in modern humans reveals common pathways under introgression, e.g. in immunity-related genes. Moreover, deeper investigation of signatures for adaptive introgression shows little evidence for beneficial effects of bonobo haplotypes in chimpanzees. In conclusion, we present an updated comprehensive picture of introgression in our closest living relatives.

S14-P17-The case of SLC30A9: ABC-based testing of a widespread Denisovan adaptive introgressed variant

Jorge García Calleja

IBE (CSIC-UPF)

Strong positive selection has been found around the human SLC30A9 gene, especially in East Asia. Selection likely acted in a highly differentiated non-synonymous variant, rs1047626, leading to a methionine-to-valine substitution in the encoded zinc transporter. Our previous work revealed that the valine variant is found in the genome of Denisova while all high-coverage Neanderthal genomes carry the methionine variant. Moreover, we highlighted the striking similarity between the putatively selected SLC30A9 haplotype across several current human populations and the Denisovan individual. Since up to 7% of the Yoruba chromosomes also carry this haplotype, we applied the Approximate Bayesian Computation (ABC) statistical framework to investigate the most plausible evolutionary origin of this variant under different scenarios and parameter estimation of both the selection coefficient in different populations and the timing of the mutation. Using forward simulations we modelled three potential scenarios: Denisovan origin, Neanderthal origin, or ancestral human origin; accounting in each of them for Back-to-Africa migrations. In order to describe the genetic variability present in the simulated datasets and the observed data, we used a variety of summary statistics that account for linkage disequilibrium patterns, haplotype and nucleotide diversity, and admixture.

Our results show a higher posterior probability for a Denisovan origin. After an introgression event in East Asia, the substitution likely spread later to European populations and was finally introduced into African populations. We highlight the power of ABC using forward simulations to interrogate candidate variants under adaptive introgression with complex demographic histories.

S14-P18-Evidence for reinforcement between humans and Neanderthals

Manuel Ferrando-Bernal

University of California San Francisco

The sequencing of the first Neanderthal genome over a decade ago revealed genetic variation in non-Africans inherited from gene flow from Neanderthals. This Neanderthal ancestry is unevenly distributed along human genomes, with large regions lacking Neanderthal variants, especially on the X chromosome and in genes expressed in the testicles. Given the divergence of the two species over half a million years ago, these patterns suggest potential for hybrid incompatibility when humans and Neanderthals came into secondary contact. In such situations,theory predicts that selection may act to increase reproductive isolation by reinforcing traits that had started to diverge between the species to aid mate choice—a process known as reinforcement speciation.

To test for evidence of reinforcement between humans and Neanderthals, we analyze high- and low-coverage genomic variation data from both sympatric and non-sympatric populations. We applied a genomic statistical method that has been recently published that compares changes in allele frequency among sympatric and non-sympatric populations. We find many loci with patterns of divergence consistent with reinforcement. Analyzing the functions of these regions, we discovered enrichment for changes in genes related to the face and body shape, hair distribution, and reproduction, consistent with the action of reinforcement selection on traits that would influence mate choice and fitness. These loci also include variants related to hernia and scoliosis that could explain prevalence differences among individuals with African and non-African ancestry.

To our knowledge, this is the first genomic study to identify a potential reinforcement speciation process in the genus Homo.

S14-P19-The genomic landscape of archaic introgression: insights from the largest nationwide genomic cohort of admixed individuals from Mexico

Valeria Añorve-Garibay

Center for Computational Molecular Biology (CCMB), Brown University

Analyses of genomic data have shown that out-of-Africa modern humans interbred with Neanderthals and Denisovans about 50,000 years ago. The detection of archaic introgressed segments has been carried worldwide mainly leveraging samples from the 1,000 Genomes Project populations, which include individuals from the Americas such as Mexicans in Los Angeles (MXL). However, this cohort may not fully capture the vast genetic diversity present in Mexican individuals, limiting our ability to explore archaic ancestry variation and to accurately assess its contribution to complex traits. To address this, we integrated data from the Mexican Biobank Project (MXB) and the Mexico City Prospective Study (MCPS) to leverage genome-wide genotype data from over 100,000 individuals residing in Mexico. This resulted in the largest cohort of individuals from the Americas ever studied, providing an unprecedented resource for exploring archaic ancestry in this region. Using a reference-free method to infer archaic introgression, we identified variants putatively from Neanderthal or Denisovan origin. Our analyses reveal that individuals from southern Mexico have on average more archaic alleles compared to those in the north. Additionally, we observed that MXB individuals carry more Denisovan alleles than MXL individuals, correlating with MXB having a high proportion of Indigenous American genetic ancestries. Notably, we observed a 1.58-fold increase in archaic ancestry coverage in MXB compared to MXL, further highlighting the importance of large-scale, diverse cohorts in uncovering archaic variation within Mexico. These findings provide a valuable resource for future research on the evolutionary and functional impact of archaic introgression in underrepresented individuals.

S14-P20-Hybrid zone mediating clinal adaptation of yellowcress herbs to high-altitudes

Tingshen Han

Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences

Hybridization and introgression are pivotal mechanisms driving adaptive evolution, particularly in extreme environments. This study explores the ecological and genomic underpinnings of high-altitude adaptation in Rorippa elata, an allotetraploid plant endemic to the Hengduan Mountains. Through integrative approaches in population genomics, quantitative genetics, and ecological genetics, we elucidate how hybridization shapes adaptation in this species. Allopolyploidization in R. elata facilitated genomic innovations, including the selective sweep of a kinetochore protein gene critical for stable mitosis and meiosis. Population genomic analyses revealed significant intraspecific divergence along latitudinal gradients, forming a cline shaped by intraspecific hybridization, geographic heterogeneity, and divergent selection. Notably, genes involved in translesion DNA repair were positively selected across the cline and during northward colonization, likely as a response to increased UV radiation. To further investigate the evolutionary dynamics of hybrid clines, we developed a multiple advanced generation inter-cross (MAGIC) population, mimicking natural hybrid evolution. These hybrids, harboring clinal variation, were transplanted into field common gardens, where life-history fitness traits were collected over multiple years. Genome-wide association studies identified key fitness genes associated with adaptive traits such as cold-induced translation, heat stress tolerance, and flowering time regulation. This research highlights the interconnected roles of hybridization and clinal processes in adaptive evolution, providing novel insights into how hybridization enables plants to thrive in challenging environments. Our findings offer a roadmap for leveraging hybrid vigor to address the challenges of a rapidly changing world, emphasizing the potential of hybridization as a driver of evolutionary innovation and resilience.

S14-P21-Hybridization and Introgression Patterns in Two Closely Related Wader Species

Xiaotong Niu

School of ecology, Sun Yat-sen University

Hybridization occurs when closely related populations or species come into secondary contact. On one hand, the fitness of hybrid offspring may be reduced, thereby strengthening reproductive isolation between species; on the other hand, hybridization promotes genetic introgression between species, enhancing the genetic diversity and adaptability of populations and species. Therefore, the exploration of hybridization contributes to a better understanding of the mechanisms underlying species divergence and the formation of reproductive isolation. This study focuses on two early-diverged, closely related Charadriiformes bird species, the Kentish Plover (Charadrius alexandrines) and the White-faced Plover (C. dealbatus), and integrates population genetics analysis to investigate the dynamics of hybrid zones and patterns of introgression. We collected phenotypic data and whole-genome resequencing data from 189 individuals from 11 sampling sites of the Kentish Plover, White-faced Plover, and their contact zone along the Chinese coastal areas. Our results reveal that the hybridization and introgression between the Kentish Plover and the White-faced Plover have experienced two different patterns of hybridization history. Gene flow has been detected in both modern secondary contact zone and ancient Bohai Sea-Taiwan isolation period. Individuals hybridizing in sympatric breeding area contain more genetic component of the White-faced Plover, and there is a segregation of feather color traits. This study focuses on the hybrid zone of closely related species in secondary contact from the perspectives of population genomics and phenotype, providing new evidence for affecting species divergence under the background of ongoing hybridization and gene flow, and the generation of species divergence.

S14-P22-A Refined Analysis of Neanderthal-Introgressed Sequences in Modern Humans with a Complete Reference Genome

Shen-Ao Liang

Fudan University

Background

Leveraging long-read sequencing technologies, the first complete human reference genome, T2T-CHM13, corrects assembly errors in previous references and resolves the remaining 8% of the genome. While studies on archaic admixture in modern humans have so far relied on the GRCh37 reference due to the availability of archaic genome data, the impact of T2T-CHM13 in this field remains unexplored.

Results

We remap the sequencing reads of the high-quality Altai Neanderthal and Denisovan genomes onto GRCh38 and T2T-CHM13. Compared to GRCh37, we find that T2T-CHM13 significantly improves read mapping quality in archaic samples. We then apply IBDmix to identify Neanderthal introgressed sequences in 2,504 individuals from 26 geographically diverse populations using different reference genomes. We observe that commonly used pre-phasing filtering strategies in public datasets substantially influence archaic ancestry determination, underscoring the need for careful filter selection. Our analysis identifies approximately 51Mb of Neanderthal sequences unique to T2T-CHM13, predominantly in genomic regions where GRCh38 and T2T-CHM13 assemblies diverge. Additionally, we uncover novel instances of population-specific archaic introgression in diverse populations, spanning genes involved in metabolism, olfaction, and ion-channel function. Finally, to facilitate the exploration of archaic alleles and adaptive signals in human genomics and evolutionary research, we integrate these introgressed sequences and adaptive signals across all reference genomes into a visualization database, ASH (www.arcseqhub.com).

Conclusions

Our study enhances the detection of archaic variations in modern humans, highlights the importance of utilizing the T2T-CHM13 reference, and provides novel insights into the functional consequences of archaic hominin admixture.

S14-P23-Electric organ discharge ontogeny and evolution in an African weakly electric fish: an insight from artificial hybridization

Feng Cheng

University of Potsdam

The African weakly electric fish Campylomormyrus possess specific electric organ to produce electric organ discharge (EOD) for communication. The EOD varies in duration that not only among different species, but also during their ontogeny. In order to identify the behind genetic mechnism that might affect the EOD divergence among different species and ontogeny, we used juvenile and adult of three different purebred species and two artifical bred intragenus hybrids for bulk RNA sequencing. Those purebred species show either short, medium or long EOD duration, and exhibit EOD duration enlogation during development. By comparing the gene expression between different EOD duration species/ontogeny, we identified several important candidate genes that might affect EOD duration divergence. Those candidates also exhibited imbalanced allele specific expression in hybrids. We used Crisp Cas9 to knocked out several candidates.

S14-P24-The Landscape of Adaptive Introgression Networks in High-Altitude Adaptation from 1,039 Tibetan Genomes

Shen-Ao Liang

State Key Laboratory of Genetic Engineering, Center for Evolutionary Biology, School of Life Science, Fudan University, Shanghai 200438, China

Archaic introgression has equipped modern humans with genetic adaptations crucial for survival in diverse environments, leading to substantial phenotypic modifications. A prominent example is the EPAS1 haplotype, derived from Denisovan admixture, which assumes a pivotal role in the regulation of hypoxia-related traits and is a fundamental determinant in facilitating high-altitude adaptation in Tibetan populations. However, beyond EPAS1, the broader role of adaptive introgression in shaping the intricate networks of functional interactions that underlie high-altitude phenotypes remains notably under-explored. In this study, we analyzed genomic data from 1,039 Tibetan individuals using an upgraded version of the reference-free software IBDmix, optimized for detecting introgression signals under natural selection. Our results demonstrated that Tibetans harbor substantially elevated levels of archaic ancestry compared to Han Chinese populations. In addition to the well-characterized high-frequency EPAS1 haplotype, we also identified several other adaptive haplotypes unique to Tibetans, exhibiting variable frequency patterns. Pathway enrichment analysis of these adaptive signals highlights a strong association with metabolic pathways linked to hypoxia and high-altitude adaptation. Importantly, we identified novel, complex regulatory networks influenced by adaptive introgression that are integral to high-altitude resilience. These findings highlight the multifaceted impact of archaic introgression on high-altitude adaptation in Tibetans, offering new insights into the diverse evolutionary trajectories of complex traits and advancing our understanding of the genetic mechanisms underlying physiological resilience to extreme environments.

S14-P25-Influence of Pliocene and Pleistocene climates on hybridization patterns between two closely related oak species in China

Yao Li

Nanjing Forestry University

Contemporary genetic admixture patterns reflect imprints of both ancient and recent gene flow, offering insights into hybridization history in response to palaeoclimate change. Here, we examine the relationships between present admixture patterns and past climatic niche suitability of two East Asian oaks (Quercus acutissima and Q. chenii) to test the hypothesis that mid-Pliocene warm climates promoted while Pleistocene cool climates limited hybridization among local closely related taxa. We analyse genetic variation at seven nuclear microsatellites (1111 individuals) and three chloroplast sequences (576 individuals) to determine present admixture patterns and ancient hybridization history. We apply an information-theoretic model selection approach to explore the associations of genetic admixture with past climatic niche suitability. We found that over 70% of the hybrids determined by nuclear markers and over 90% of the individuals with locally shared chloroplast haplotypes are concentrated within a mid-Pliocene contact zone between ~30°N and 35°N. Climatic niche suitabilities for Q. chenii during the mid-Pliocene Warm Period (~3.264–3.025 mya) and during the Last Glacial Maximum (~0.022 mya) best explain the admixture patterns across all Q. acutissima populations and across those within the ancient contact zone, respectively. Our results highlight that palaeoclimate change shapes present admixture patterns by influencing the extent of historical range overlap. Specifically, the mid-Pliocene warm climate promoted ancient contact, allowing widespread hybridization throughout central China. In contrast, the Pleistocene cool climate caused the local extinction of Q. chenii, reducing the probability of interspecific gene flow in most areas except those sites having a high level of ecological stability.

S14-P26-Neanderthal and Denisovan nuclear-embedded mitochondrial DNA sequences introgressed into modern human genomes

Qiong Zhu

Fudan University

Nuclear-embedded mitochondrial DNA sequences (NUMTs) are molecular fossils of ancestral mitochondrial DNA, providing valuable insights into human evolutionary processes. Despite the significance of NUMTs, their distribution in archaic human genomes and their evolutionary trajectory in modern populations have been underexplored. Here, we identify 483 distinct NUMTs in 2,519 high-coverage modern human genomes across geographically diverse populations. Additionally, we discover 10 different NUMTs from three Neanderthal genomes and one Denisovan genome, with four (40%) of these NUMTs persisting at high frequencies in modern human populations. Phylogenetic analysis reveals that NUMT insertions span a range of evolutionary origins, from pre-dating the divergence of Neanderthals and Denisovans to more recent events within modern humans. Notably, two ancestral NUMTs diverged prior to the Neanderthal-modern human split and were introgressed into contemporary populations. Specifically, two Denisovan-derived NUMTs, dating approximately 198,000 and 427,000 years ago, were later introgressed into Oceanian populations. We find that individuals carrying NUMTs overlapped with introgressed segments by a nonrandom manner, highlighting the role of archaic introgression in the persistence of NUMTs. Some NUMTs are positively selected and associated with genes involved in olfactory function and collagen processing. Furthermore, certain NUMTs may influence 3D genome architecture, notably affecting the immune-associated DOCK8 locus. We show a dynamic evolutionary landscape of human NUMTs and highlight how archaic introgression of NUMTs as previously overlooked genetic legacy contributes to both functional and phenotypic effects in modern humans.

Symposium 15: Transposable elements in genome evolution and biotechnology development

S15-P01-Enhanced DNA methylation associated with amplification of transposable elements during foxtail millet domestication

Wenlei Guo

Institute of Zoology, Chinese Academy of Sciences

Foxtail millet, one of the most ancient crops, has undergone significant changes in its agronomically important traits during domestication. To gain deeper insights into the genetic and epigenetic mechanisms underlying the domestication syndrome, we conducted a comprehensive analysis of DNA methylation using whole-genome bisulfite sequencing (WGBS) on 60 foxtail millet accessions, including 20 wild, 20 landrace, and 20 cultivar varieties. In addition, we employed RNA-seq and small RNA-seq to explore the impact of epigenetic variations on the population. Our population-level comparison revealed a notable increase in methylation levels during domestication, which is distinct from the balanced methylation variations observed in other crops. By integrating structural variation (SV) data from the pan-genome, we discovered that approximately one-third of the differentially methylated regions (DMRs) are attributed to deletions or insertions of transposable elements (TEs). Notably, two amplification events of the DNA/PIF-Harbinger transposon were characterized and found to be over-represented in CHH DMRs, predominantly located in gene-rich regions. 21-/22-nt sRNAs were enriched in the younger burst event and trigger the non-canonical RNA-directed DNA methylation (RdDM) pathway. Interestingly, the TE bursting events, which were more pronounced in the cultivated population, exhibited strong associations with the genes that contribute to the emergence of domestication and agronomically important traits, such as seed shattering and plant architecture. Collectively, our findings offer a comprehensive perspective on the intricate interplay between genetic and epigenetic variations. They enhance our understanding of how transposable element bursts contribute to crop domestication, providing valuable insights for future crop improvement strategies.

S15-P02-Termini-seq: a generalized telomere-targeted sequencing method for high-resolution characterization of telomere diversity and dynamics

Congying Chen

State Key Laboratory of Oncology in South China, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, China

The termini of eukaryotic linear chromosomes possess conserved structures known as telomeres, which safeguard chromosome integrity via end-capping. A comprehensive characterization of the sequence variation, epigenetic modification, and cellular dynamics of telomeres can offer valuable insights for important biological questions such as genome instability, cellular senescence, as well as organismal aging and cancer, yet this is highly challenging due to the repetitive nature and low genomic abundance of telomeres. Here we present a generalized telomere-targeted sequencing method named Termini-seq that does not rely on the prior knowledge of telomeric sequence composition. With this novel method, we were able to obtain a complete picture of telomere length, sequence variation, and DNA methylation profile in a single sequencing run with chromosome-specific resolution. We validated the performance of Termini-seq with various eukaryotic models and telomere characteristics, which revealed consistently high accuracy and efficiency. Finally, we applied Termini-seq to study the dynamic changes of telomeres during organismal aging, DNA replication machinery perturbation, and cancer drug resistance, which revealed mechanistic and evolutionary insights into how alterations in telomeres impact cellular and organismal fitness.

S15-P03-A Comprehensive Study of Dynamics and Impacts of Structural Variations and Transposon Insertions in Populus Species

Shuyu Liu

Lanzhou University

Structural variants (SVs) and transposable elements (TEs) represent a major portion of genetic mutations segregating across eukaryotic genomes, with profound ecological and evolutionary implications. However, the genetic architecture, functional impacts and phenological consequences of these naturally occurring variations are not yet fully understood in natural populations. To address this gap, we constructed a nucleotide resolution catalog of genetic variants, including single-nucleotide polymorphisms (SNPs), SVs, and TE insertions, and presented a rationale for comprehensively exploring their evolutionary genomics using whole-genome resequencing data from multiple widespread Populus species. Our findings revealed that SNPs, SVs, and TE insertions have, to varying extents, shaped the genetic diversity of Populus species through the combined forces of selection and demographic events. Notably, rare TE insertions exhibited a disproportionately large impact on gene expression compared to rare SVs and SNPs. The cis-regulatory roles of SVs and TE insertions facilitated the adaptation and functional divergence of genes retained from whole-genome duplication events. Moreover, SVs and TE insertions may serve as candidate variants for complex phenology influenced by numerous small-effect variants. By including SVs and TE insertions, two novel genomic components, to extend the understanding of Populus population dynamics. Our research highlights the overlooked genome-wide variants produced by SVs and TE insertions, offering a comprehensive view of different types of genetic variations in genome evolution and architectural complexity, along with resources supporting reliable SVs and TE insertions in non-model species.

S15-P04-Widespread of horizontal gene transfer in eukaryotes

Kun Li

Shanghai Jiao Tong University

Horizontal gene transfer (HGT) is the transfer of genetic material between distantly related organisms. While most genes in prokaryotes can be horizontally transferred, HGT events in eukaryotes are considered as rare, particularly in mammals. Here we reported the identification of HGT regions (HGTs) in 10 eukaryotes, including human, mouse, cow, lizard, frog, zebrafish, fruit fly, nematode, Arabidopsis and yeast. By comparing their genomes with thousands of eukaryotes, bacterial and virus genomes, we found between 10 and 243 non-redundant HGTs per eukaryote species. PCR and third-generation sequencing validation of HGTs excluded the probability of contamination. Genes impacted by HGTs are enriched in transmembrane transport. Some HGTs have transformed their host genomes with various numbers of copies and have impacted hundreds, even thousands of genes. Our findings reveal that HGT is widespread in eukaryotic genomes, and HGT is a ubiquitous driver of genome evolution for eukaryotes.

S15-P05-The diversity and evolution of transposons in termite genomes

Cong Liu

Okinawa Institute of Science and Technology

Transposons are mobile elements that can replicate and move across the genome. They are a major component of insect genomes and have played a significant role in the evolution of insects. Here, we investigated the evolution of transposons in the genomes of 45 taxonomically diverse termite species and two cockroaches, Blatta orientalis and Cryptocercus meridianus. Our genomes contained between 5,925 and 18,684 transposon families, accounting for 24.90 to 54.74% of their genome size. Long interspersed nuclear elements (LINEs) and DNA transposons were the most abundant and diverse transposon classes across species. The composition in transposon families was correlated with termite phylogeny. Furthermore, we reconstructed the history of transposon expansion for each species and found evidence of bursts of transposition that took place in parallel in various termite lineages. Overall, our results shed light on the dynamic of transposon activity at million-year-old timescales.

S15-P06-Comparative analysis of the evolution of transposons in two medaka species

Xiaochi Zhou

Sun Yat-Sen University

Hainan medaka (Oryzias curvinotus) and pectoral-fin spot medaka (Oryzias pectoralis) are closely related species to the widely studied Japanese medaka (Oryzias latipes). While their habitats partially overlap, distinct ecological preferences separate them: O. pectoralis is restricted to freshwater environments, whereas O. curvinotus inhabits coastal brackish waters. These species offer valuable opportunities to investigate how effective population size influences genome structure, particularly in relation to transposon dynamics. In this study, we present the first de novo genome assembly of O. pectoralis, achieved through the integration of Illumina paired-end sequencing, Oxford Nanopore sequencing, and Hi-C data. The assembly comprises 49 scaffolds (N50 = 31 Mb) and has a total genome size of 740 Mb. Subsequent whole-genome resequencing of both species suggests that O. pectoralis likely experienced a reduction in effective population size in the recent past. Furthermore, transposon expansion was identified as a significant contributor to the genomic enlargement observed in the Oryzias genus, as evidenced by annotations of available genomes from other species within the genus. Ongoing analyses aim to determine whether the fixation of recently active transposons is influenced by historical population bottlenecks. We expect this study can add new insights to the study of genome structure and transposon dynamics.

S15-P07-The evolution and mechanisms of genome gigantism in Caelifera insects

Xuanzeng Liu

Shaanxi Normal University

The evolution and mechanisms of genome gigantism in Caelifera (Insecta: Orthoptera) were studied from genomic and transcriptomic perspectives. The genome size of the suborder Caelifera ranges from 0.97Gb to 21.48 Gb, within which the Acridoidea have undergone substantial expansion, while those of the Tetrigoidea and Tridactyloidea have undergone slight contraction. The results indicate that the radiation evolution of Acridoidea was accompanied by genome gigantism, with frequent chromosomal rearrangements observed during the same period, leading to an increase in chromosome numbers (n=6 to n=12). Furthermore, ongoing co-evolution between transposable elements (TEs) and host silencing mechanisms has been observed. Contrary to the commonly accepted arms race relationship, the number of C2H2-ZNF domains (KRAB-ZFPs host silencing factors) decreased rapidly in gigantic genomes. Concurrently, a decelerated rate of TE loss was observed in gigantic genomes. At the transcriptomic level, the proportion of TE silencing is lower in gigantic genomes than in species with smaller genome size. The study revealed that the significantly lower expression of piRNA methyltransferase (Hen1 methyltransferase) may underpin low-level piRNA silencing. Furthermore, the findings of the study indicate that PIWI-clade evolution is subject to stronger purifying selection in contracted genomes compared to expanded genomes. In the genomic gigantism observed in Caelifera, the presence of exceptionally long introns (>200 kb) in PIWI-clade genes, resulting from numerous TE insertions, may serve as a mechanism for TE anti-silencing. Overall, low-level host silencing may contribute to genomic gigantism in Caelifera, offering a valuable perspective for exploring the C-value paradox.

S15-P08-Caste-dependent genomic stability and protection against transposable elements in the superorganism

Esther van den Bos

University of Münster, Germany

In multicellular organisms, the germline is the only tissue that contributes to the next generation. Consequently, it is essential to safeguard germ cells from mutagenic environmental hazards in order to maintain their genetic integrity. In superorganismal social insect colonies, the queen serves as the primary reproductive unit, analogous to the germline of multicellular organisms, and should thus be much more strongly protected against mutagenic influences than workers, that represent the soma of the insect colony.

Transposable elements (TEs) are a significant source of novel mutations in eukaryotes and their activity can be induced by external and physiological stressors. The germline is particularly protected against the activity of TEs in multicellular organisms, raising the question whether the same applies to the queens of superorganismal insects. To explore caste-specific differences in genomic robustness under the influence of stress, we pharmacologically suppressed Hsp90, a central player of the cellular canalization machinery. We used ONT long read sequencing of pools of full-sib queens and workers to assess rates of novel TE insertions in germline and somatic individuals. Together, the results of this study shed light on the convergent evolution of germline genomic integrity of superorganisms and multicellular organisms.

S15-P09-From Red to Green: The Evolutionary History of Porphyridium's Phycobilisomes

Hwan Su Yoon

Sungkyunkwan University, Suwon, Korea

Phycobilisomes are complex light-harvesting antenna proteins anchored to the thylakoid membrane of plastids, playing a crucial role in light absorption for photosynthesis in cyanobacteria, glaucophytes, and rhodophytes. Three species of unicellular red algal genus Porphyridium (P. purpureum, P. sordidum and P. aerugineum) show different colors; red, purple, green, respectively, and they are inhabiting mainly in marine water (P. purpureum), brackish-freshwater (P. sordidum), and freshwater habitats (P. aerugineum). To better understand phycobilisome evolution and chromatic adaptation, we sequenced chromosome-level of genomes of these three Porphyridium species. Genome analysis showed that P. purpureum contains 28 phycobilisome genes, including 20 linker proteins, while P. sordidum and P. aerugineum (loss of cpeA) retain only 8 and 6 linker proteins, respectively. Loss of linker proteins were directly correlated with TE insertions, therefore, these genomes feature directly support different color of three species that were adapted in different habitats in which different wavelengths of light are available. Our findings offer novel insights into the diverse evolutionary pathways of phycobilisomes and the intricate mechanisms of chromatic adaptation within the Porphyridium.

S15-P10-Unprecedented independent domestications of horizontally transferred DNA transposons in mammals

Vladimir Kapitonov

Marine Biological Laboratory

The eukaryotic genomes are populated by transposable elements (TEs). While the majority of TEs follows the parasitic mode of their coevolution with the host genome, some of them can be converted into important functional elements. Here I describe evolution of the HSMAR1 transposon (Mariner superfamily) and show that ~47 million years ago, almost simultaneously, HSMAR1 was transferred horizontally 11 times into genomes of ancient mammals belonging to 11 diverse groups: Cingulata, Hipposideridae, Rhinolophidae, Megadermatidae, Pecora (cow, bison, sheep, deer and giraffe), Muroidea (mouse, rat, hamster, gerbil, blind mole-rat), Tarsiiformes, Simiiformes (humans, apes, New and Old World monkeys), Lemuriformes (gray mouse lemur, blue-eyed black lemur, greater bamboo lemur, sifaka), Otolemur (small-eared galago), and Scandentia (northern tree shrew). I identified five genes derived from five independent events of domestication of HSMAR1 copies that happened ~47 MYA. These genes encode proteins derived exclusively from the HSMAR1 transposase. The reported domestication events happened in Lemurifromes, Hipposideros armiger (great roundleaf bat), Rhinolophus ferrumequinum (greater horseshoe bat; two domestications), and Pecora. Most surprisingly, the first three domestication events (Lemuriformes, roundleaf and horseshoe bats) occurred independently at the same orthologous locus in the corresponding genomes. Moreover, the remaining two domestication events occurred at the same position of the second orthologous locus, which bears no similarity to the first one, and is located on a different chromosome. Possible explanations of these highly unusual and unexpected events are presented.

S15-P11-Extrachromosomal circular DNA long-read sequencing and genome assembly of a S. lycopersicum breeding line revealed active LTR retrotransposons derived from S. peruvianum L. introgressions.

Pavel Merkulov

MIPT

The evolution of plant genomes has been significantly influenced by transposable elements (TEs). Even among closely related species, differences in the transposition activity and landscapes of transposons can be quite significant. One method that is frequently used in tomato breeding is interspecific hybridization, which produces unique combinations of TEs from different species. In modern cultivars, the consequences of this mechanism for TE transposition activity have not been investigated. Using extrachromosomal circular DNA (eccDNA) nanopore seq uencing, we found two families of highly active Ty1/Copia LTR retrotransposons in tomato Solanum lycopersicum), called Salsa and Ketchup. These families' members generate thousands of eccDNAs under controlled conditions and epigenetic stress. The majority of the eccDNA produced by Ketchup and Salsa showed little similarity to the genomic sequence of S. lycopersicum, according to eccDNA sequence analysis. Whole-genome nanopore sequencing and de novo genome assembly were used to determine the origin of these TEs. We discovered that interspecific introgression from S. peruvianum caused these TEs to arise in a tomato breeding line. All of our research indicates that interspecific introgressions can bring in novel genetic variations and active transposable elements from other species, which can enhance genetic and phenotypic diversity.

S15-P12-NICOTIANA BENTHAMIANA “VIRUS FACTORY” OBTAINMENT VIA TSWV-BASED VIGE

Alexander Polkhovskiy

Moscow Institute of Physics and Technology

Biological mutagenesis (BM) involves the use of advanced mutagens such as transposable elements to expand the genetic variability of modern crop plants. Viral induced gene editing (VIGE) is emerging as another direction of BM. Recently, tomato spotted wilt virus (TSWV) was used as a biological container to drive CRISPR-Cas9 system into plant cells and cause targeted mutations without actual transgenesis. Viral systemic infection in plant somatic tissues and further in vitro culturing allows to regenerate plants from this tissue and obtain ready-to-use virus-free plant breeding material. In some cases, agrobacteria-free infection of various plant species (including agrobacteria-resistant ones) is needed. For this a special “virus factory” plant with compromised anti-viral immunity is needed.

In this work we adapted and applied TSWV-driven CRISPR-Cas9 mutagenesis of N. benthamiana phytoene desaturase (PDS), Dicer-like enzyme 2 (DCL2) and Dicer-like enzyme 4 (DCL4) genes. Mutagenesis in latter two genes is needed for obtaining of “virus factory” plants. We analyzed dependencies between virus symptoms progression and mutation rate, factors affecting virus infection and mutagenesis efficiency rate (heat-shock, cold stress, NbRDR6 silencing). We applied in vitro culturing of systemically infected plant leaves and performed regeneration of mutated plants. Obtained plants were screened for mutagenesis rates and their seeds were used for analysis of progeny mutation splitting and viral infection inheritance.

S15-P13-Factors driving differences in the distribution of transposable elements in natural populations of diverse species

Anna Haider

Vetmed Uni Vienna

Transposable elements are important drivers of genome architecture. These autonomous genetic elements may replicate and potentially disrupt vital genes. Studying the distribution of TEs in different organisms such as Drosophila melanogaster and Arabidopsis thaliana we noted differences in the distribution of TEs among natural populations. While in D.melanogaster all investigated individuals had a few insertions of each TE family, we found a patchy distribution of TE families in A.thaliana, where several TE families are entirely absent in some accessions. We explore different factors that could drive these differences such as the populations structure or sexual reproduction. We explore the effect of different factors with simulations of TE invasions in spatially distributed populations. Finally, we test whether the factors we identified can account for the distribution of TEs in diverse other organisms.

S15-P14-Transposable Elements (TEs) Control Genes as Targets of Positive Selection During Recent Evolution of Diverse Human Populations

Roudin Sarama

Queen Mary University London

Transposable elements (TEs) are mobile genetic sequences that can insert themselves into new locations within the host genome, influencing gene expression, chromatin structure, and regulatory networks. TEs have played a crucial role in shaping the human genome, contributing to evolutionary innovation and genetic diversity. However, because TEs can proliferate within the genome in a largely uncontrolled manner, they often act as selfish genetic elements, leading to mutagenic effects that can compromise host-genome integrity and cellular function. As a result, cellular organisms have evolved regulatory mechanisms to suppress TE activity and restrict their spread. Recent studies suggest that genes regulated by TEs may be subject to positive selection as a signature of fast evolution in response to challenges from TE. In this study, we analyse a previously curated list of 59 autosomal coding genes involved in TE control processes, across 54 diverse human populations. By integrating advanced summary statistical methods, we identify signals of long-term positive selection acting on these loci. Additionally, we compare our empirical results with neutral expectations generated through msprime-based simulations. Several candidate genes exhibit extreme values of Tajima's D and Fay and Wu's H in multiple human populations, consistent with long-term positive selection. These findings provide new insights into how TE control mechanisms have shaped human genetic diversity and underscore the evolutionary importance of host-TE interactions.

S15-P15-Novel Endogenous Betaretrovirus and how it aids species identification in Slow Lorises (Nycticebus)

Ravinder Kanda

Oxford Brookes University

Endogenous retroviruses (ERVs) are a type of transposable elements which arise when exogenous infectious retroviruses integrate into the host genome, specifically within germline cells. While much of ERV research has focused on humans and great apes, studying ERVs in other primates can offer valuable new insights. In this study, we screened the genomes of two endangered slow loris species, Nycticebus bengalensis and N. coucang, using a previously identified long terminal repeat (LTR). This led to the discovery of a novel ERV, which appeared at varying levels of completeness and included numerous solo LTRs. Phylogenetic analysis revealed that this ERV is a betaretrovirus, most closely related to HERV-K. Interestingly, despite its close relation to HERV-K, this ERV is not found in humans but is present in Asian loris species and absent in other primates. The striking similarity of viral insertions between the two loris species suggests that the current N. coucang genome may either be a hybrid of the two loris species or represent a N. bengalensis from a different population than the reference genome. This study emphasizes the importance of exploring ERVs in more distantly related primates, offering insights into both species' genomes and evolutionary history, as well as viral evolution.

S15-P16-Identification of Novel Transposable Elements in Anuran Genomes: Challenging a Mammal-Centric Perspective

Ravinder Kanda

Oxford Brookes University

The class of transposable elements (TEs) known as endogenous retroviruses (ERVs) have had a pivotal role in vertebrate genome evolution, providing coding and regulatory sequences which can be co-opted by the host genome. Currently there is a mammal-centric view of ERVs, and in vertebrate groups such as Anura (frog and toad species), very few ERVs have been described. Within Anura, the role of TEs in genome evolution is of interest due to Anura having comparable genome sizes to mammals and their habitation of multiple diverse niches. Here, through in silico screening of 46 anuran genomes, we have identified a total of 57 novel ERV families with intact viral genes, represented by multiple loci, in 29 genomes. The structures of these ERVs differ to what is typically described in mammalian genomes, with viral genes differentially classified across retroviral genera. Furthermore, the intactness of the viral genes and estimated ages of the ERV families suggest that many of the novel ERVs identified were recently active and newly inserted in the genomes, with several ERVs shared across multiple species. These findings have provided the foundation for the investigation of ERVs in species genomes beyond mammals, to understand the presence and role of ERVs across the vertebrate tree of life. Additionally, this study supports the systematic screening of further vertebrate species to obtain an accurate representation of TEs, and in particular ERVs, which have shaped genome evolution.

S15-P17-Association Between Locus-Specific HERV Expression and Gene Dysregulation in Hepatocellular Carcinoma

Hui Wen

George Washington University

Background: Human endogenous retroviruses (HERVs) harbor accessory proteins that influence cellular processes and have been linked to various diseases, including cancer. This study investigates locus-specific HERV expression and its association with gene dysregulation in hepatocellular carcinoma (HCC).

Methods: We analyzed 424 HCC samples from The Cancer Genome Atlas (TCGA), which comprised 371 tumor tissues and 50 matched normal tissues. We employed Telescope to identify and quantify HERV expression across the total RNA sequencing data.

Results: The majority of differentially expressed HERVs exhibited reduced expression in tumor tissue (166 downregulated vs. 50 upregulated), suggesting a potential functional role of HERV expression patterns in shaping the pathophysiological landscape of HCC. Specifically, the suppression of HERV-H family members, which are known to regulate cellular differentiation, may contribute to tumor dedifferentiation, increased plasticity, and enhanced metastatic potential. This loss of differentiation control and increased adaptability may play a critical role in driving the progression of liver cancer.

Discussion: Our study highlights the critical role of HERV expression in HCC, highlighting the differential regulation of specific HERV families in tumor tissue. The widespread downregulation of HERV loci suggests a potential loss of regulatory functions that may influence tumor biology, while upregulated HERVs could reflect broader changes in the tumor microenvironment. These expression patterns point to a complex relationship between HERV activity and key cancer-associated processes. Further investigation into these mechanisms may provide new insights into HCC pathogenesis and offer potential therapeutic targets within HERV-associated pathways.

Symposium 16: Thematic symposium: evo-devo

S16-P01-The human-specific mutations of MCPH1 induce sensory-experience-dependent brain neoteny and extended lifespan in mouse

Xin Luo

Kunming Institute of Zoology, CAS

Neoteny is a developmental specialization that contribute to human intelligence and lifespan, although the underlying genetic mechanism remains largely unknown. Here, we demonstrate that the mice carrying two human-specific mutations of the MCPH1 gene show neotenic development in the brain, and prolonged aging with an extended lifespan. Specifically, the MCPH1-humanized mice exhibit cortical folding and a human-like delay of dendritic spine maturation in the sensory cortex. The human version of MCPH1 can regulate activity-dependent gene expression via the DNA double-strand break (DSB) repair pathway to decrease spine density during the juvenile stage, and to improve synaptic plasticity and cognitive abilities through glia-neuron interaction at adulthood. Remarkably, the MCPH1-humanized mice display a healthier aging by keeping better motor and cognitive abilities. These findings reveal the pleiotropic evolutionary outcomes of the human MCPH1, mirroring the characteristics of human neoteny.

S16-P02-Single-cell multi-omics analyses of chondrocytes reveal evolutionary changes of lower limb development and their potential roles in human bipedalism

Bing Su

Kunming Institute of Zoology, Chinese Academy of Sciences

Bipedal locomotion is a hallmark of human evolution. Systemic structural changes in lower limb long bones represent a key phenotype of human adaptation to upright walking, yet its molecular and genetic mechanisms remain elusive. With the use of single-cell multi-omics approaches, we investigated regulatory networks underlying endochondral ossification, a critical process in long bone development. We generated multi-omics data from the developing lower limbs of rhesus macaque, including single-cell RNA-seq (scRNA-seq), bulk ATAC-seq and bulk HiChIP. By integrating these data with the human and mouse datasets, we delineated the conserved and species-specific regulatory features of chondrocytes. In particular, to investigate genetic mechanisms underlying huma bipedalism, we analyzed the human-specific sequence changes, including structural variants (SVs), single-nucleotide variants (SNVs), and human accelerated regions (HARs), and we evaluated their impacts on chondrocyte regulatory networks.

Our results elucidate developmental, epigenetic, and evolutionary mechanisms shaping bipedal locomotion, advancing our understanding of chondrocyte-specific regulatory dynamics in humans.

S16-P03-Investigation on the Evolution and Function of the Barbel Traits in Some Goby Fishes

Yi Shao

Sun Yat-sen University School of Ecology

The study of phenotypic divergence among species is a crucial aspect of evolutionary ecology. The presence of barbels across various fish species suggests that this morphological feature has undergone multiple instances of independent evolution. A notable feature within Gobioidei is the presence of barbels, a trait shared by all members of certain genera (e.g., Gobiopsis) or exhibited by specific species within a genera (e.g., Tridentiger). Therefore, investigating this trait may contribute significantly to understanding the adaptive evolution of gobies. In this study, a comprehensive analysis of five goby species was conducted using genomic, transcriptomic, population genetic approaches and spatial transcriptomics. In genomics, de novo assembly and annotation were performed for five relevant species, followed by a comparative genomic analysis to identify genes involved in the loss of conserved elements between species with and without barbels. We also sought potential convergent genes related to the barbels trait and estimated the divergence times of the species. Analyses of positive and relaxed selection were conducted. In transcriptomics and population genetics, differential expression analysis (DEG) was carried out for three species within the Tridentiger genus. The historical effective population size was assessed, and both recent and ancient genes involved in population sweep were identified. Fst and Dxy values were calculated across the genome, along with recombination rates on chromosomes. Finally, spatial transcriptomics was utilized to conduct a preliminary exploration of the expression levels of genes identified in the aforementioned analyses, examining their spatial distribution.

S16-P04-Identification and Analysis of Thyroid Hormone Synthesis-Related Factors and Their Signaling Pathways in Regulating Ascidian Metamorphosis

Jin Zhang

Ocean University of China

Thyroid hormones (THs) are important for regulating animal metamorphosis, but the molecular mechanism of TH synthesis has not been experimentally investigated in invertebrates, while it is well understood in vertebrates. This is primarily because the Thyroglobulin (TG) protein has not been identified in invertebrates. TG is a structural protein that provides tyrosine for TH synthesis and is unique to vertebrates. Previous studies have attempted to identify TG-like proteins in tunicates, but have not been successful in determining the TG-like sequences or verifying their functions. Here, we identified the the TG-like protein in invertebrate Styela clava and found a new follicle-like structure has function of TH synthesis in S. clava larva. By bioinformation analysis and biochemical experiments, one protein was identified as the potential sc-TG-like protein, named sc-TG-like284-31. Based on the expression pattern of sc-TG-like284-31 protein, we found a special follicle-like structure at front of S. clava larval trunk, with function of TH synthesis. The results of cross-species comparison were shown homology between follicle-like structure in S. clava and papillae structure in Ciona robusta. Further, the in vivo knocking down and in vitro iodination experiments was indicated that the sc-TG-like284-31 protein was regulated S. clava metamorphosis through directly participated in S. clava TH synthesis, like vertebrate TG protein. This work offers new evidence for molecular mechanism of TH synthesis in invertebrates, contributing for understanding the evolution of TG protein.

S16-P05-Evolutionary and Developmental Basis of Ovipositor Specialization in Drosophila suzukii

Xiaodong Xu

Westlake University

The ovipositor, a pivotal reproductive structure in insects, exhibits extensive morphological diversification driven by adaptations to specific ecological niches. Drosophila suzukii, an invasive species native to Southeast Asia, possesses a highly specialized ovipositor characterized by its enlarged structure and increased density of bristles. This adaptation allows the species to deposit eggs in fresh fruits, contrasting with other Drosophila species that primarily target decaying fruits. This unique morphological trait has driven the global proliferation of D. suzukii and its substantial impact on agriculture. From an evolutionary developmental (evo-devo) perspective, this study explores the genetic mechanisms underlying the morphological specialization of the ovipositor in D. suzukii and its closely related species. By integrating transcriptomics (RNA-seq), single-cell RNA-seq, and functional validation approaches, we identified key patterning genes and the Notch/Wnt signaling pathway as central regulators of ovipositor development. These genes coordinate cell size expansion and bristle formation, driving the morphological divergence of the ovipositor. Comparative analyses of D. suzukii and D. melanogaster reveal distinct gene expression patterns and regulatory networks that underpin the evolutionary divergence of this trait. Our findings provide novel insights into the genetic basis of trait divergence and ecological adaptation, shedding light on the evolutionary mechanisms that enable D. suzukii to exploit new ecological niches. By elucidating the gene regulatory networks involved in ovipositor development, this study advances our understanding of the evolutionary processes shaping insect morphology and ecological specialization.

S16-P06-A shared supergene but distinct dynamics regulate mimicry polymorphisms in closely related Papilio species.

Sofia Sheikh

University of Chicago

Despite the fact that sexes share the vast majority of their genomes, males and females can develop not only distinct morphologies, but also multiple alternate phenotypes suited to particular life histories and selective pressures. Sex-limited polymorphism – the ability of one sex to develop multiple discrete morphs – thus represents a striking and widespread case of adaptive phenotypic variation. Such polymorphisms are often controlled by supergenes, which are sets of tightly linked genetic elements or loci that together function to instruct the development of alternate phenotypes. However, the evolutionary origins of supergenes and the mechanisms through which supergene alleles shape distinct developmental programs remain poorly understood. The doublesex (dsx) supergene controls mimicry polymorphism in several species of Papilio butterflies: divergent dsx alleles switch between discrete mimetic and non-mimetic female wing patterns. We leverage this system to first characterize the evolutionary history of dsx alleles using comparative genomic approaches and find evidence for ancestral co-option of dsx in sex-limited polymorphic wing patterning. Additionally, we identify potential functional elements that differentially regulate alternate dsx alleles across multiple species. To further characterize how the dsx supergene functions, we compare antibody staining and RNA-seq data between two Papilio species that diverged 15 million years ago, showing that mimetic dsx alleles have acquired species-specific patterns of spatiotemporal expression and have unique effects on the transcriptomes of developing wings. Collectively, our findings suggest that, despite the shared use of dsx among closely related species, the mimicry switch operates through distinct underlying mechanisms.

Symposium 16: Thematic symposium: evo-devo

S18-P01-Single-nucleus multi-omics analyses reveal cellular and molecular innovations in the anterior cingulate cortex during primate evolution

Jiamiao Yuan

Kunming institute of zoology

The anterior cingulate cortex (ACC) of the human brain involves higher-level cognitive functions such as emotion and self-awareness. We generated the profiles of human and macaque ACC gene expression and chromatin accessibility at single-nucleus resolution. We characterized the conserved patterns of gene expression, chromatin accessibility, and transcription factor binding in different cell types. Combining the published mouse data, we discovered the molecular identities and cell lineage origin of the primate von Economo neurons (VENs). Our in vitro and in vivo experiments identified a group of primate-shared and human-specific VEN marker genes, such as PCSK6, ADAMTSL3 and CDHR3, potentially contributing to VEN morphogenesis. We demonstrated that the human-specific sequence changes account for the 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.

S18-P02-Comparative single-cell multiome identifies evolutionary changes in neural progenitor cells during primate brain development

Kaimin Chen

Kunming Institute of Zoology, Chinese Academy of Sciences

Understanding the cellular and genetic mechanisms driving human-specific features of cortical development remains a challenge. We generated a cell-type resolved atlas of transcriptome and chromatin accessibility in the developing macaque and mouse prefrontal cortex (PFC). Comparing with published human data, our findings demonstrate that although the cortex cellular composition is overall conserved across species, progenitor cells show significant evolutionary divergence in cellular properties. Specifically, human neural progenitors exhibit extensive transcriptional rewiring in growth factor and extracellular matrix (ECM) pathways. Expression of the human-specific progenitor marker ITGA2 in the fetal mouse cortex increases the progenitor proliferation and the proportion of upper-layer neurons. These transcriptional divergences are primarily driven by altered activity in the distal regulatory elements. The chromatin regions with human-gained accessibility are enriched with human-specific sequence changes and polymorphisms linked to intelligence and neuropsychiatric disorders. Our results identify evolutionary changes in neural progenitors and putative gene regulatory mechanisms shaping primate brain evolution.

S18-P03-Evolution of haematopoiesis: regulation of gene expression in vertebrate blood cells

Carolyn de Graaf

WEHI

The major hematopoietic lineages have existed since early in vertebrate evolution, covering the key functions of oxygen transport, clotting and defense against infection. Some of the specifics of how these functions are performed are unique to mammals, such as using platelets formed from polyploid megakaryocytes for clotting rather than thrombocytes.

Using single cell transcriptomes, we created an atlas of gene expression of haematopoietic cells in mammals, birds and fish, and annotated cells to their closest mammalian homologue by lineage. We used both publicly available data and generated our own atlas of chicken haematopoiesis as a representative of bird haematopoiesis. This has allowed us to look at how conserved haematopoietic gene expression is in different vertebrates in relation to human blood cells and find species specific haemopoietic and immune cell types. We have identified genes which are conserved as signatures of blood and immune lineages across all vertebrates and genes which are only associated with mammalian haematopoiesis.

We also investigated the regulatory elements associated with gene expression during haematopoiesis using ATAC-seq data. We took the mammalian specific and vertebrate conserved gene lists and looked at their regulation via ATAC-seq and found that the more broadly conserved genes have more lineage specific open chromatin regions associated with them than genes with mammalian specific gene expression. This suggests that redundancy in regulation helps maintain lineage specific expression in divergent species. Specific transcription factor motifs have conserved enrichments at lineage specific genes, despite there being a high turnover of enhancer sites over time.

S18-P04-Error-tolerant phylogenetic tree construction enables accurate lineage tracing with single-cell data

Qing Yang

Westlake University

Somatic mosaicism is fundamental to understanding human development, aging, and disease progression. However, accurate lineage tracing using single-cell sequencing data is severely hampered by high levels of noise and error in somatic mutation detection. These errors, which can constitute at least 20-40% of identified mutations, significantly distort phylogenetic tree structures and hinder reliable lineage reconstruction.

To address this challenge, we developed PhyloSOLID, an error-tolerant algorithm designed to construct robust phylogenetic trees from noisy single-cell sequencing data. Our algorithm employs an innovative two-step approach combining a statistical penalty-scoring model and a "pivot tree" model to systematically identify and exclude artifact mutations while preserving true biological signal. Testing against both simulated data and real single-cell datasets demonstrates PhyloSOLID's consistently superior accuracy in phylogenetic reconstruction compared to existing methods.

PhyloSOLID seamlessly integrates with various single-cell technologies, enabling comprehensive lineage analysis from genomic, transcriptomic to spatial profiling data. The algorithm incorporates an interactive visualization framework adaptable to various analytical scenarios, from tumor evolution to developmental lineage tracing. This flexible framework allows users to easily inspect tree quality, interactively remove likely false mutations, and interpret complex lineage structures with confidence.

By providing robust lineage reconstruction capabilities, PhyloSOLID addresses a critical need in somatic mosaicism research. Its superior accuracy and comprehensive toolkit make it particularly valuable for studying clonal dynamics in aging tissues and tracking the evolution of somatic mutations in disease progression, offering new insights into how somatic mosaicism shapes human development.

S18-P05-Single-worm genomics: A chromosome-scale genome assembly for Diplolaimelloides wo'ombe from the Great Salt Lake, UT

Julie Jung

University of Utah

Recently, we discovered a novel species of nematode living in the saltiest habitat where nematodes have ever been discovered, in the microbialites of the Great Salt Lake, Utah, USA. Herein we present our novel species, Diplolaimelloides wo'ombe, and describe these worms both morphologically and molecularly. In addition, we present a chromosome scale draft genome, assembled using a combination of Nextera short read sequencing, PacBio long reads, and Oxford Nanopore long reads. In order to assemble this genome, we had to get past a historical problem in nematology, which is that these are really tiny worms. D. wo'ombe tend to be only about 2 ng of DNA, so we borrowed a technique from the protist community (Multiple Displacement Amplification) that allows single-cell genomics using ultra low inputs of DNA. Lastly, we describe culturing conditions that enable the survival and reproduction of these novel nematodes.

S18-P06-Postmortem Cellular Resilience: Transcriptomic and Epigenetic Insights into Somatic Cell Adaptation After Organismal Death

Zihan Liu

GuangZhou Medical University

Currently, research on the regulatory mechanisms by which cells adapt to organismal death is highly limited, particularly regarding how surviving cells transition from a multicellular regulatory mode to an autonomous survival state following the collapse of the body's systems. We systematically investigates changes in gene transcription patterns and chromatin accessibility by comparing human lung tissue samples obtained immediately after clinical surgical resection with those at different post-mortem interval (PMI), aiming to elucidate the dynamic regulatory processes of cells after death. Through comprehensive mRNA transcriptomic and epigenomic analyses of tissues at various PMI, we observed that cells undergo a series of complex gene expression and chromatin remodeling events postmortem. During the early PMI, cells adapt to adverse conditions such as hypoxia and low temperatures by activating stress response and immune-related genes. During the late PMI, significant changes in the expression of development and proliferation related genes suggest that cells may enter a distinct state of repair or regeneration. Furthermore, we identified a set of biomarkers associated with functions such as DNA damage repair, whose expression levels exhibited a strong correlation with PMI, offering potential molecular feature for assessing the biological state of postmortem tissues. These findings reveal that somatic cells exhibit a resilient pattern which adopts single-cell survival strategies after being separated from an organism for a period of time. This model highlights the dynamic adaptability of post-mortem somatic cells to harsh environments, providing new insights into the post-mortem evolution of somatic cells and the potential applications of regenerative medicine.

S18-P07-Characterization of molecular evolution of primate spermatogenesis is improved by an optimized single-nucleus RNA-seq protocol

Mengjun Wu

Aarhus University

Despite relatively short divergence time, due to strong selective pressure, primate testes have undergone rapid evolution on both morphological and molecular levels. This makes spermatogenesis a particularly interesting context to study molecular evolution as well as evolutionary processes such as primate speciation, in which the unique regulatory dynamics of sex chromosomes during spermatogenesis provide a battlefield for sex chromosome competition, thus driving their rapid divergence.

While the development of single-nucleus RNA-seq (snRNA-seq) allows us to study primate spermatogenesis at single cell resolution, obtaining high-quality snRNA-seq data from primate testis tissues is difficult, due to high levels of debris and ambient RNA. In addition, published snRNA-seq data of primate testes have relatively low gene sensitivity due to limitations of early 10X chemistry. To improve data quality and resolution, we established a new snRNA-seq preparation protocol using the latest 10X GEM-X chemistry. The optimized protocol significantly reduced the ambient background signal and is able to generate high quality data from archival frozen tissues. Compared to previous data, the newly generated data have much higher sensitivity and are thus more accurate in capturing gene expression trajectories during spermatogenesis. Moreover, the new data allowed us to clearly distinguish X and Y spermatids and infer their ratios across several primate species, thus providing an unprecedented view of primate sex chromosome evolution.

Symposium 20: Thematic symposium: evolution of immunity

S20-P01-Deciphering the Evolution of Immune Cells at the Single-Cell Transcriptome Level

Xiang Liu

Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences

During the evolution of invertebrates, immune cells have reached an increasingly complex level, while invertebrates, they exhibit a high degree of diversity. Due to the limitation of conventional approaches in identifying a broader spectrum of immune cells and molecules from various species, the details of immune cell evolution are still unclear. Here, we constructed a sophisticated immune cell atlas across multiple invertebrates to vertebrates and comprehensively analyzed the origin and evolutionary process of immune cells involved in innate and adaptive immunity. We revealed conserved and divergent features during the evolution of immune cells and molecules, particularly highlighting the contribution of phagocytosis and coagulation function cells to the immune defense in almost all species including invertebrates and vertebrates. Additionally, we identified distinctive lymphocyte-like subtypes of VLR-based AIS, most of which showed similar transcriptome to the TCR/BCR/MHC-based lymphocyte lineages. However, AIS undergoes reorganizations during the evolution from jawless to jawed vertebrates. We found that these changes are not only reflected in anticipatory receptors but also involve the lymphocyte functional subtypes. A novel unconventional T-like cell subtype was enriched in jawless vertebrates that carry the common antigen response of T cells, but their activation is associated with the MPL-TPO signaling. Although the lymphocytes of TCR/BCR/MHC-based AIS are highly conserved, they show limited cellular differentiation in the oldest jawed vertebrates, cartilaginous fish. Our evidence suggests that CD8+ T cells evolve earlier than CD4+ T cells, even though cartilaginous fish possess genes for differentiation into CD4+ T cell subsets.

S20-P02-When Cells Rebel: a comparative genomics investigation into marsupial cancer susceptibility

Cleopatra Petrohilos

University of Sydney

Cancer is ubiquitous in multicellular life, yet susceptibility varies significantly between species. Previous studies have shown a genetic basis for cancer resistance in many species, but few studies have investigated the inverse: why some species are particularly susceptible to cancer. The Dasyuridae are a family of carnivorous marsupials that are frequently reported as having high rates of cancer prevalence. We hypothesised that this high susceptibility also has a genetic basis. To investigate this, we created the first reference genome for the kowari (Dasyuroides byrnei), a dasyurid species with one of the highest rates of reported cancer prevalence among mammals. We used this to perform a comparative genomics analysis alongside nine previously assembled reference genomes: four dasyurid species and five non-dasyurid marsupial species. Genomes were annotated using FGENESH++, then orthologs of genes from the human Cancer Gene Census identified. Genes were assigned to orthogroups for input to CAFE (Computational Analysis of gene Family Evolution) analysis to identify gene families that had undergone significantly rapid expansions or contractions in each lineage. In the dasyurids, we identified large expansions in Ras genes, a family of oncogenes. Interestingly, a similar expansion of Ras genes was also identified in the bandicoot and bilby. These genes were primarily expressed in tissues such as testes, ovaries and yolk sac, so we hypothesise they serve a reproductive role. Future work is required to identify the potential roles of oncogene expansions in cancer susceptibility in these marsupial species.

S20-P03-The Diversity and Evolution of MHC Class I and II in Raptors: Insights from the Accipitridae Family

Jumaporn Sonongbua

Genomics and Bioresource Research Unit, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

The major histocompatibility complex (MHC) is a highly variable gene family encoding antigen-presenting molecules for intracellular (MHC class I) and extracellular (MHC class II) pathogen recognition in the adaptive immune system. In birds, MHC evolves through duplications, gene conversion, and concerted evolution, particularly in MHC class II, while MHC class I evolution remains unclear. However, the MHC diversity across avian lineages, especially in non-model species, is poorly understood. The Accipitridae family exhibits significant ecological and evolutionary diversity, with variation in life history, migration, and diet influencing pathogen exposure. Studying MHC diversity in this group is essential to the understanding of adaptive immune responses. In this study, we genotyped MHC-I exon 3 and MHC-IIB exon 2, both encoding the peptide-binding domains, across 27 Accipitridae species and analyzed polymorphism and selection patterns among subfamilies. Our results showed comparable allele counts and nucleotide diversity between the two MHC classes, with a positive correlation of MHC-I exon 3 and MHC-IIB exon 2 in nucleotide diversity in Buteoninae and Aegypiinae. Notably, Aegypiinae (scavenging raptors) exhibited higher MHC class I allele diversity than other subfamilies. In addition, distinct evolutionary patterns were observed between the two classes: MHC-I exon 3 evolution was shaped by pervasive negative selection, while MHC-IIB exon 2 was influenced by both pervasive and episodic positive selection. These findings highlight the distinct selection pressures driving MHC evolution in Accipitridae, providing a broader perspective on its role in avian adaptation and diversification.

S20-P04-The Underlying Evolutionary Dynamics of the Human Immune Response.

Rutvi Rajpara

Institute of Genomics, University of Tartu, Estonia

Modern humans have been exposed to a wide variety of pathogens throughout their existence. Being able to adapt to them was a matter of life and death. Genetic variants that helped to fight against pathogens remained in the human gene pool. Today, we still carry some of these genetic modifications in our DNA that impact human health. However, linking specific pathogen outbreaks to their impact on shaping modern human genomes is challenging. This research addresses that challenge by investigating how past episodes of natural selection and admixture with archaic humans shaped the underlying genetics of the modern human immune response.

To achieve this, we use genomic and immune-related datasets to identify loci associated with pathogen-specific immune responses and investigate their evolutionary history. For example, by reconstructing allele frequency changes, examining signatures of selection, we aim to understand how pathogens shaped these genetic adaptations. Furthermore, we investigate past admixture events and the contribution of archaic introgressed alleles from Neandertals and Denisovans to immune response loci to determine their role in pathogen-specific adaptations. The integration of genomic data, GWAS, eQTL data, and computational tools helps identify associations between evolutionary processes and immune response phenotypes.

The findings of this research will yield novel insights into the relationship between pathogens and disease-associated variants in present-day individuals. They will highlight how historical evolutionary events shape human health and disease susceptibility.

S20-P05-The evolutionary divergence of primate immunoglobulin genes

Katalin Voss

Cornell University

Immunoglobulin (IG) genes are an important part of the adaptive immune system. They consist of variable, diversity, and joining gene segments that encode the light and heavy chains of antibodies essential for pathogen recognition and neutralization. IG genes exhibit remarkable diversity both within and between species but despite their critical role in immune defense, the evolutionary history and forces shaping IG diversity remain poorly understood. Understanding these forces can provide valuable context for how the adaptive immune system has adapted to pathogen challenges over time. A major obstacle in the study of IG has been that until very recently it has been very hard to assemble the IG locus, owing to its structural complexity and repeat-rich nature. In this study we focus on primate IG genes aided by the recent availability of highly accurate genomic data, including five primate telomere-to-telomere genomes and two additional high-quality assemblies from the Vertebrate Genomes Project. These species serve as an optimal model system for investigating the evolutionary processes driving intra- and interspecific divergence in IG genes. We assessed gene diversity by examining the number of genes per locus and species, as well as the genetic distances both within and between species. We further quantified the degree of incomplete lineage sorting and genes and amino acid residues that are under selection. Finally, we integrated expression data to test whether macroevolutionary divergence patterns were related to the gene expression levels within species. This work provides valuable insights into the evolutionary dynamics of the adaptive immune system.

S20-P06-Evolution, Regulation, and Function of Type I Interferon Responses in Yinterochiroptera Bats

Xiaomeng Li

Zhejiang University

Bats (Chiroptera) are known to host multiple pathogens, with Yinterochiroptera and particularly Rhinolophid bats serving as the primary reservoir for SARS-related coronaviruses. Previous studies suggest a distinct type I interferon pattern in bats including constitutive IFN expression. Using high-quality genomes from 8 Yinterochiroptera bats and 7 Yangochiroptera bats, we characterized the type I IFN locus without assembly gaps, providing the first complete type I IFN atlas for these species. Phylogenetic annotation revealed the contraction of IFNAs and expansion of IFNWs, as seen with some other bat species, along with the identification of IFNDs and IFNXs. Viral infection assays highlight the regulation of endogenous IFNs in bat and human cell lines. Notably, all Rhinolophus IFNWs showed weaker induction of downstream ISG pathways and antiviral responses compared to human IFNW and IFNA, while Rhinolophus IFNAs were largely inactive. Using AlphaFold modelling and molecular docking, we evaluate the binding affinities of bat/ human IFNs with IFNAR1/2. Rhinolophus bats exhibit a dampend type I IFN loop compared to humans.The attenuated immune response may enable them to coexist with a broad range of viruses, suggesting a controlled immune response that allows prolonged viral persistence without significant immune activation.

S20-P07-Adaptive evolution and chromatin remodeling during primate immune evolution

Tauras VILGALYS

University of Chicago

Humans and nonhuman primates differ in their susceptibility to many inflammatory and infectious diseases. For instance, humans are extremely susceptible to lipopolysaccharide (LPS, an outer membrane component of gram-negative bacteria), going into septic shock at orders of magnitude lower exposure than other primates. These differences likely result, in part, from genetically controlled divergence in the innate immune response. However, the underlying evolutionary and gene regulatory mechanisms remain poorly understood.

Here, we investigate the immune response to LPS in four primate species (humans, baboons, rhesus macaques, and ring-tailed lemurs). To measure the immune response, we generated single-cell multiome data (gene expression + chromatin accessibility) on peripheral blood mononuclear cells before and after stimulation with LPS. Human cells, particularly monocytes, mounted the strongest immune response to LPS with both more differentially expressed genes and larger magnitude effects. Genes that were upregulated in humans, but not other primates, were enriched for human-specific open chromatin regions before and after immune stimulation. Our data show that the same transcription factors orchestrate the LPS-response across primates, but which genes are targeted by these key immune regulators has changed during primate evolution. Among these changes, we observe signatures of adaptive evolution at transcription factor binding sites near human-specific upregulated genes, particularly for members of the REL and AP1 transcription factor families. Thus, adaptive evolution of transcription factor binding sites contributes to divergence in the primate immune response via chromatin remodeling. We are expanding this work to understand the response to viral stimuli.

S20-P08-Adaptive evolution of inflammasome-related genes in amniotic vertebrates: Divergence and plesiomorphy explain the patterns of surface charge variation in mammalian NLRC3

Tao Li

Department of Zoology, Faculty of Science, Charles University

Inflammasomes are cytoplasmic multiprotein complexes comprising sensor proteins, inflammatory caspases, and in some (but not all) cases adapter proteins. Although they form key part of the innate immune system that activates expression of inflammatory cytokines, such as interleukin (IL)-1B and IL-18, in response to infectious microbes and host-derived danger signals, little is known about their molecular evolution in vertebrates. Here, we analyzed adaptive sequence variation of 6 genes involved in inflammasome formation from RIG-I-like receptors and NOD-like receptors families across 105 species, covering all major clades of mammals, birds and reptiles. In total, we detected 86 positively selected sites based on a consensus of three dN/dS-based detection methods (MEME, FUBAR, and PAML). Aiming to describe the patterns of the molecular adaptation in the inflammasome-related genes , we further modelled out their 3D protein structures and explored the mismatch between phylogeny and the phenotype clustering of surface charge distribution. A convergence-like pattern was observed between mammals and rodents in the NLRC3 gene, prompting further analysis with a broader range of species. We found that this convergence-like pattern in NLRC3 is given by a plesiomorphic state of the molecular phenotype shared between lemurs and rodents. Adopting predictive bioinformatics, our results provide insights into evolution of interspecific differences and similarities in inflammasome structures that contribute to immune response optimization throughout evolution.

S20-P09-Diversifying selection on Complement Factor H in a wild rodent

Rocco Notarnicola

Adam Mickiewicz University

Complement Factor H (CFH) is a crucial regulator of the innate immune response in vertebrates, protecting self-cells by inhibiting the constitutively expressed alternative complement pathway. Bacterial pathogens can hijack this immune inhibitor to evade clearance, opening the stage for host-pathogen coevolution. However, how CFH evolves in response to selective pressure from parasites remains unclear. Here, we explore it by characterising polymorphism in CFH domains 19-20, known to interact with bacterial pathogens, and test for signals of selection in wild populations of bank voles (Clethrionomys glareolus). First, we searched for signatures of positive selection by calculating rates of synonymous to non-synonymous substitutions. Second, we analysed population differentiation in SNP frequencies at CFH compared to a background of >60,000 random genomic markers generated via Restriction site-Associated DNA sequencing (RAD-seq). We found six CFH variants among 13 bank vole populations. Signatures of positive selection were found in codons involved in interactions with pathogens, particularly within the domain 20. CFH was more differentiated among populations than genomic background, implying divergent selection acting on it, possibly because of adaptation to locally evolving pathogens. Other genomic regions differentiated between populations included genes involved in, among others, neurogenesis, oogenesis, and the alternative pathway for complement activation, which is regulated by CFH. This suggests that additional genes may be under diversifying selection, including immune genes involved in complement activity. Our findings underscore the involvement of key innate immunity genes, traditionally considered evolutionarily conserved, in evolutionary arms races with pathogens, leading to their diversification.

S20-P10-Elucidating PD-1 and SHP-2 Interactions: Structural, Evolutionary, and Computational Perspectives

Hugo Samano Sanchez

Zhejiang University-University of Edinburgh Institute

Cell signaling relies on precisely balanced activation and repression events to ensure robust yet transient responses, thus preventing accidental irreversible cellular commitments. One such example is T cell activation, which is positively modulated by antigen presentation and co-stimulation, and negatively regulated by inhibitory receptors such as PD-1.

PD-1 is a cell receptor with a disordered intracellular region with two conserved tyrosine-containing motifs: ITIM and ITSM. When phosphorylated by Lyn or other Src family kinases, these motifs recruit and form a stable complex with SHP-2, a protein phosphatase involved in cell proliferation and anti-apoptotic signaling and frequently found dysregulated in cancer. SHP-2 itself is a modular protein comprising two SH2 domains and one phosphatase domain. Recent structural biology studies have shown the critical contributions of specific amino acids within and between the ITIM and ITSM motifs.

Furthermore, PD-1, encoded by the Pdcd1 gene, is conserved across mammals, retaining the ITIM and ITSM motifs. Notably, some distant species possess Pdcd1 homologs with unique mutation patterns that reshape the interaction surfaces. Through computational analyses, we investigated these alternative motifs, revealing potential roles in divergent protein–protein interactions and offering deeper insights into the evolutionary plasticity of motif-domain interactions.

S20-P11-Quantifying differential host adaptation during the viral life cycle

Mary Reed-Weston

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. 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. Additionally, though previous research has made it clear that there is significant adaptation in host genomes in response to viral infection, little is known about what host proteins related to steps of viral infection are under the strongest selection. In order to address this gap, we have characterized adaptation by looking at genes encoding viral-interacting proteins (VIPs) at different stages of viral infection (e.g. entry into the cell, viral genome replication, intracellular transport). Using an ABC extension of the McDonald-Kreitman test, ABC-MK, we identified viral entry, viral release from the cell, and intracellular viral transport as hotspots of adaptation during the viral life cycle, using human data as a proof-of-concept. We intend to repeat this analysis in 9 species of Myotis bat to compare patterns of adaptation between humans and bats. Given the large population sizes and higher genetic variation in bats, we expect to be able to detect more subtle patterns of adaptation across the viral life cycle and see how these patterns differ from humans. Investigating these mechanisms will allow for increased understanding of bat resistance to and coevolution with viruses, which is invaluable in the prevention of spillover to humans.

Symposium 28: Human history and past social organisation in the light of palaeogenomes: new methods, new findings session B

S28-P01-From fragments to foodways: Challenges in palaeodietary reconstruction using ancient metagenomics

Jonas Niemann

Globe Institute, University of Copenhagen

Shotgun DNA sequencing of archaeological samples, such as dental calculus, paleofeces, or ancient mastics, is a powerful tool for reconstructing ancient microbial communities and studying the evolution of the human microbiome. In some cases, this method can also provide valuable insights into the diets and subsistence strategies of past populations. However, detecting DNA from dietary sources in ancient shotgun metagenomic datasets presents significant challenges, primarily due to the low abundance of ancient eukaryotic DNA in these samples and the potential for false positive assignments. In this study, we re-analyzed five previously published ancient metagenomic datasets using various bioinformatic workflows commonly employed for identifying eukaryotic taxa in archaeological samples. Our findings emphasize the risks associated with false positive results in ancient metagenomic analyses and highlight the critical importance of rigorous validation to ensure reliable interpretations.

S28-P02-Extended enamel database for hominin palaeoproteomics allows for more accurate identification of hominin groups

Viridiana Villa Islas

University of Copenhagen

In contrast to ancient DNA, proteins survive for longer periods and in wider regions on Earth. Ancient proteins have been recovered even from skeletal remains dating to more than 2 million years from tropical regions, especially from dental enamel, allowing us to explore the distant past of human evolution with molecular evidence. The analysis of such data requires the usage of an appropriate protein sequence database since it assists in accurately reconstructing an ancient protein sequence through de novo sequencing and determines the sequence information identified and retained during phylogenetic analysis. Palaeoproteomic studies of dental enamel have mainly focused on analysing a set of core proteins (n≈13). However, additional proteins have been reported in protein extracts from dental enamel, and these potentially contain single amino acid variants (SAAVs) that are informative for the taxonomic assignment of unidentified hominin teeth. To explore this potential, we characterise protein sequence variation among 86 proteins reported in the dental enamel literature for global, present-day human populations (1KGP and HGDP) as well three Neanderthals and one Denisovan. We catalogued the number of proteins containing unique SAAVs for Denisovans, Neanderthals, and modern humans. The addition of present-day human variation to the database led to the recognition of the specific archaic SAAVs and the several SAAVs shared between modern humans and archaic hominins - these would otherwise not be identified by relying on the human protein reference alone. A deeper analysis and enrichment of the present database will assist in improving palaeoproteomic analysis of hominin skeletal material.

S28-P03-Paleogenomics of Human Remains from Nenggiri Valley in Malaysia

Ahmad Faiz Shazwan Bin Abdul Malik

School of Humanities, Archaeology and Cultural Resources Program, Kanazawa University

The historical, religious, cultural, and linguistic diversity of Southeast Asia is shaped by multiple human migration waves in the past. Various theories, supported by archaeological, linguistic, anthropological, and genetic evidence, have been proposed to explain these migrations. The Malay Peninsula played a crucial role in prehistoric human movements. However, ancient DNA studies remain limited despite the discovery of over 50 burials in the Nenggiri Valley, linked to pre-Neolithic (Hoabinhian) and Neolithic farming populations. The main challenge in genetic analysis is the poor preservation of collagen and endogenous DNA. The site's location, situated between Negrito (traditionally hunter-gatherers) and Senoi (traditionally farmers) territories, provides key evidence of prehistoric human admixture.

We have successfully obtained new genomic sequences from Neolithic human remains at the Gua Cha site. Principal Component Analysis (PCA) and outgroup-f3 statistics indicated high genetic affinity between these individuals, previously sequenced ancient populations, and modern Indigenous groups in Malaysia. Admixture tests were carried out to identify potential migration groups contributing to the cultural development of these sites and examine gene flow among ancient and modern populations including the Indigenous populations.

This research enhances our understanding of human migration and admixture in the region. The findings provide new genetic insights into ancient population dynamics and their connections to modern Southeast Asian populations.

S28-P04-1000 years of solitude: Tracing local lineages and ancestry in a prominent Albanian BA-IA tumulus

Xiaowen Jia

Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology

Situated at the interface of the Aegean and the Adriatic in southeastern Albania, the prominent tumulus of Kamenice functioned as an inhumation burial site from 1700 to 500 BCE. In this work, we generated genome-wide SNP data for 230 individuals buried in Kamenice over its complete historical span, alongside 19 Iron Age individuals from North Macedonia and 2 Late Bronze Age individuals from southwestern Bulgaria. This comprehensive dataset together with fine-scale genetic analysis using IBD segment networks provide pioneering insights into, on the one hand, genetic continuities and changes of Late Bronze Age and Iron Age Albania; on the other hand, the function and meaning of tumuli in prehistoric Europe.

Contrasting societal transformations identified in the wider region, here we document a genetically stable and even genealogically continuous population at Kamenice over a timespan of more than 1000 years. The local population maintained a distinct homogenous genetic profile while participating in a regional kinship network extending over 300 kilometers, as demonstrated through identified biological relatedness up to the 10th degree. Moreover, for the first time in a time transect of a single continuously used archaeological site, we observe a partial turnover of male lineages. Occurring in a generally patrilineal system, this shift around 750 BCE also co-occurs with a substantial change in the structure and material symbols of the tumulus.

S28-P05-Moving Beyond Migration: Using Archaeogenetics to Study Diverse Modes of Mobility in the Ancient Mediterranean

Hannah Moots

Centre for Palaeogenetics Stockholm

Migration has often been invoked as the primary mechanism of genetic change in archaeogenetic studies. While this term has frequently been used as an umbrella term for any type of human movement, mobility occurs and has occurred in many diverse patterns. This project examines how archaeogenetic studies can explore various modes of mobility, beyond migration, especially through interdisciplinary collaborations and the integration of historical, archaeological and bioanthropological sources. We identify challenges in the study of mobility patterns - such as equifinality that arises from working with small, geographically heterogeneous datasets and the spectre of culture-historical archaeology's fixation on migration - and suggest possible solutions, such as developing a shared terminology across disciplines. To illustrate the theoretical points raised above, we reanalyze archaeogenetic data from port cities of the Mediterranean in the context of complementary information from allied disciplines. Port cities are an excellent arena to shed light on the diverse patterns of mobility, as they are important sites of interaction between the maritime world and the hinterlands beyond, between local and diasporic populations. The primary case study will be insights gained from archaeogenetic and mobility isotope data for 20 individuals from Isola Sacra, the necropolis for Rome's primary trading port. The complementary nature of these two analyses, combined with insights from archaeological and textual sources, provides a complex picture of mobility in the Roman world which could not have been fully understood from either source alone.

Session 3

Symposium 12: Population genetics through time session A

S12-P01-Population genomics reveals genetic load and selection in a bioluminescent Flashlight fish

Jiawen Li

Sun Yat-sen University

The evolutionary process between the Flashlight fish Anomalops katoptron and its endosymbiotic bacterium Candidatus Photodesmus katoptron has not been fully studied. To address this, we have employed 12 samples of flashlight fish to obtain next-generation sequencing results of the flashlight fish's population genomic DNA. Based on this data, we conduct population genetic analysis to explore the population size, population history, genomic mutations, recombination events, and genes under recent positive selection in A. katoptron. Furthermore, by integrating key genes involved in the symbiosis between A. katoptron and bacteria, we analyze the selective pressures they experienced and their evolutionary processes to clarify their roles and adaptive changes during evolution, aiming to more fully reveal the co-evolutionary process between A. katoptron and bacteria.

S12-P02-Higher African ancestry proportion is associated with an attenuated increase of serum phosphorylated tau181 in a longitudinal study of aging African Americans

Lu Wang

University of Georgia

African Americans have a higher risk of Alzheimer's disease (AD) and related dementia (ADRD) than non-Hispanic Whites. This study examines three blood biomarkers for ADRD – phosphorylated tau181 (p-Tau181), glial fibrillary acidic protein, and neurofilament light – in the context of genetic ancestry for an African American cohort.

METHODS

The Family and Community Health Study (FACHS), started in 1996, recruited and followed 5th graders and their caregivers from several hundred African American families. Using the 1000 Genomes Project as a reference dataset, we estimated genetic ancestry proportions using PCA and ADMIXTURE analysis. We tested the associations of African ancestry proportion with cross-sectional biomarker levels (N = 573) and their longitudinal changes over 11 years (N = 225).

RESULTS

African ancestry proportion was not associated with cross-sectional biomarker levels but was associated with an attenuated increase in p-Tau181 over time. Genetic factors associated with AD and biomarkers in European cohorts showed no associations with biomarkers in our cohort.

DISCUSSION

Our results suggest the longitudinal p-Tau181 changes differ across ancestry. More research is needed to characterize the longitudinal dynamics of ADRD biomarkers across ancestry and the driving biological or sociocultural factors.

S12-P03-Structured population of Eurasian woolly mammoths driven by isolation-by-distance and periodic gene flow

Jiaming Hu

State Key Laboratory of Geomicrobiology and Environmental Changes, China University of Geosciences, Wuhan, China

The woolly mammoth (Mammuthus primigenius) presents an ideal case for studying biogeographic patterns across the Northern Hemisphere. In this study, we examined the population dynamics of Eurasian woolly mammoths by generating 12 new whole-genome sequences and integrating them with existing datasets. Our analysis revealed three distinct population clusters (China, Europe and Siberia) across Eurasian that diverged simultaneously during the Middle Pleistocene and were primarily shaped by isolation-by-distance. After their initial split, recurrent gene flow from Siberia into Europe and China added complexity to within-population dynamics and increased the genetic diversity of the recipient populations. Additionally, stronger gene flow events, mostly during the Last Glacial Maximum (LGM), possibly facilitated mitochondrial introgression, contributing to the prolonged presence of mitochondrial Clade II in China and a shift from Clade III to Clade I in Europe. These findings underscore the role of intraspecific gene flow in shaping the evolutionary history of woolly mammoths and present testable hypotheses regarding the impacts of gene flow accompanying population dispersals under changing climates.

S12-P04-Maternal genomic legacy indicates multiple ancestral founding lineages linked to ancient farmers and pastoralists in the formation of Altaic people

Guanglin He

CQMU

Advances in ancient human DNA research have transformed our understanding of human evolutionary history, the role of admixture and introgression in adaptation, and the genetic underpinnings of human diseases and traits at spatiotemporal resolutions. However, the evolutionary dynamics of ancient East Asian populations remain fragmented and require systematically evaluated via large-scale ancient genomes. We imputed 800 Chinese ancient genomes via HuaXi haplotype reference panel and discussed the population interactions and admixture histories of ancient East Asian populations spanning 160,000 years to the historical period. Anatomically modern East Asian ancestries had limited influence on present-day East Asians. Post-glacial population transitions and Holocene expansions played pivotal roles in shaping modern East Asian populations. Agricultural innovations originating in China drove the expansion of the founder population of millet and rice farmers, facilitating the spread of proto-Sino-Tibetan and broader language families across northern and southern China. The genetic links between southern Chinese farmers, ancient Southeast Asians, and Lapita populations in Oceania support a co-dispersal model of farming, language, and people from South China. Genomic analyses of Tibetan Plateau populations revealed that early Neolithic Zongri hunter-gatherers were closely related to ancient northern East Asians. Subsequent Tibetan Plateau groups displayed pronounced population stratifications and gene flow with Tibetan-Yi Corridor populations. Additionally, migrations and admixture events linked to shifting subsistence strategies in North China enriched the genomic diversity of ancient northern East Asians and contributed to phenotypic diversity. Adaptive signatures derived from spatiotemporally diverse ancient genomes elucidate the evolutionary underpinnings of high-altitude adaptation, pigmentation, and morphological traits.

S12-P05-MtDNA genomes reveal relaxed purifying selection in human population expansions after the Last Glacial Maximum

Hong-Xiang Zheng

Fudan University

Modern humans have experienced explosive population growth in the past thousand years. We hypothesized that recent human population have been accompanied by environments with relaxation of selective constraints, possibly due to the more abundant food supply after the Last Glacial Maximum.

The ratio of nonsynonymous to synonymous mutations (N/S ratio) is a useful and common statistic for measuring selective constraints. Though phylogenetic analysis over 35,000 global mtDNA genomes — including those of East Asians, Africans, Europeans, and indigenous Americans — we observed a higher N/S ratio in the expanded lineages, indicating the weaker effect of purifying selection on the expanded populations.

While larger effective population size and longer coalescence time in the expanded lineages might theoretically strengthen purifying selection, our findings suggest that the observed higher N/S ratio in these lineages was due to relaxed selective constraints. Further computer simulation showed that a higher N/S ratio was unlikely when all individuals were under the same selective constraints, consistent with our observation.

Further analysis of ~1,000 Han Chinese individuals from the Human Phenome Atlas revealed that the recently accumulated deleterious mutations were significantly associated with lower oxygen intake and reduced resting metabolic rate, indicating the crucial role of mtDNA variants in mitochondrial oxidative phosphorylation. Thus, we argue that the expanded populations subjected to weaker selective constraints than the non-expanded populations did. The findings of this study may have profound implications for understanding the role of relaxed selective constraints that shaped human phenotypes and diseases in extant human populations.

S12-P06-Genetic origins of the Kiritimati population from central-eastern Micronesia

Kai Tätte

Human Evolution, Department of Organismal Biology, Uppsala University, SE-752 36 Uppsala, Sweden

The migration of Austronesian-speaking populations through Oceania has intrigued researchers for decades. The Kiribati islands, situated along the boundaries of Micronesia and Polynesia, provide a crucial link in this migration. We analyzed the genome-wide data of the Kiritimati population of Kiribati to uncover their genetic origins and connections with other Oceanian groups. Our study reveals that the Kiritimati population primarily exhibits Remote Oceanian-related ancestry associated with ancient Lapita and present-day Polynesian populations. In addition, our identity-by-descent analysis identifies populations from Philippines as their closest relatives in Island Southeast Asia. The genetic links between Kiritimati, ancient Lapita, and modern Polynesians underscore the shared ancestry and continuous gene flow across these regions. This genetic continuity and ongoing links are supported by linguistic and cultural evidence, illustrating a complex history of migration and admixture in Oceania.

S12-P07-From Eurasia to North America: The Key Role of Equus dalianensis in Late Pleistocene Equid Interactions

Shiwen Song

China University of Geosciences, Wuhan

Equids migrated out of the Americas around 2.5 million years ago, rapidly diversifying across the Old World. Although many lineages have gone extinct, their genetic legacy persists in extant species through ancient admixture. Here, we investigate Equus dalianensis, a now-extinct equid species from northeastern China, using 18 mitochondrial genomes (0.6-10× coverage) and two nuclear genomes (0.8× coverage). Our analyses confirm that E. dalianensis represents a distinct species at both mitochondrial and nuclear levels. Gene flow was detected between E. dalianensis and North American equids as well as Eurasian E. lenensis, with the latter's admixture signal diminishing within a few thousand years. Additionally, E. dalianensis exhibited a historically broader range but lower genetic diversity compared to E. lenensis and modern horses (E. caballus and E. przewalskii). Isotopic evidence supports the hypothesis that low genetic diversity and dietary specialization contributed to its extinction. This study advances our understanding of equid evolution, extinct species interactions, and extinction mechanisms, offering valuable insights for conservation biology.

S12-P08-Maximum likelihood-based, base call quality-aware population genetic statistics estimators

Andrey Vaulin

NTU

With finite data, classic asymptotically unbiased estimators for population genetic parameters may yield impossible values, and introduce a median bias driven by sample size, complicating the comparison of populations of different sizes. Futhermore, estimation of allele frequencies and population genetic statistics from genotyping data may cause errors when the data is of bad quality or sequencing data has a low coverage – for example, in metagenomes and paleogenetics. The estimators can be improved by constructing the estimators from raw data with base call qualities.

Here we develop a maximum likelihood-based approach that more accurately estimates the population genetic parameters and addresses the question of comparison (median) bias. Our approach provides three improvements to classic statistical estimates: (i) we directly estimate the population genetics statistics from base call accuracies in terms of phred scores instead of requiring genotype calls, (ii) we show how the pursuit for an unbiased estimator can lead to impossible and/or less accurate values, and provide a corrected estimator for p, and (iii) we provide two new estimators: first one minimizes the median bias for Watterson θ and the second one that reduces error for Tajima's D, compared to standard estimator .

Overall, this work improves standard estimators that are central to the population genetics field.

S12-P09-Ancient genomics reveal Chinese pig domestication history associated with human migration and management

He Yu

State Key Laboratory of Gene Function and Modulation Research, School of Life Sciences, Peking University, Beijing, China

As one of the earliest domesticated animals, the history of pig domestication has remained controversial, especially in East Eurasia. Here, to trace the origin and dispersal of East Eurasian domestic pigs, we generate 21 nuclear genomes and 23 mitogenomes from ancient pig/boar remains spanning from 5,800 BCE to 1,300 CE across China. By combining published Eurasian pig genomes, we demonstrate that a continuous pig population domesticated in northern China played a major role in shaping the genetic diversity and contributing to the unique black coat color of modern Chinese domestic pigs. The expansion of ancient pigs from northern China mirrors the spread of Yellow River millet farmers to the Yangtze River Basin and Southwest China, accompanied by genetic introgression from local wild boars. We also observe a loss of genetic diversity and an increase in inbreeding in northern pigs, potentially linked to captive breeding practices as early as 3,000 BCE. Our findings illuminate the domestication and dispersal history of pigs in China, which is closely associated with the human migration and societal development.

S12-P10-Chromosome-level genome provides insights into parallel latitude adaptation in closely related cyprinid fish Hemiculter leucisculus and Siniichthys bleekeri

Qiaoling Deng

Institute of Hydrobiology, Chinese Academy of Sciences

Parallel evolution is usually expected among closely related taxa exposed to similar selective pressures, and suggested as a mechanism of biodiversity origin. Hemiculter leucisculus and Siniichthys bleekeri are two closely related cyprinids that have similar geographic distribution patterns in China, which might have genetic parallelism. Here, by analyzing individual whole-genome resequencing data of 223 samples with a south-north gradient in China, we investigated whether latitude adaptation occurred in these two species, and whether parallelism existed between them. A high-quality reference genome of Hemiculter leucisculus, based on PacBio HiFi long read sequencing and HiC-directed scaffolding was presented first. Our further analysis revealed parallel lineage differentiation in Hemiculter leucisculus and Siniichthys bleekeri populations between Southern and Northern China, but no differentiation between Eastern and Western regions along the Yangtze River. We then identified 33 loci showing signals of selection that overlap between the two cyprinids in the "north vs south" contrast. These loci were located within, or near 12 genes. Three of these genes—RHO, PRPH, and ATP1A1—are involved in photoreceptor function, a key component of the visual system. In addition to the findings that light levels can influence visual system in birds and primates, with northern individuals bearing larger eyes coping with weaker light, our study indicated that latitudinal light levels might influence visual system in fish as well. Our study provides more examples to parallel adaptation to latitudinal variations with genomic data, and may enhance our knowledge to the mechanism of biodiversity origin.

S12-P11-Bayesian model-averaging of parametric coalescent models for phylodynamic inference

Yuan Xu

University of auckland

Bayesian phylodynamic models are important for reconstructing population histories from genetic data, yet their accuracy depends on suitable demographic models. To handle model uncertainty, we propose a Bayesian Model Averaging (BMA) framework that integrates multiple parametric coalescent models—including constant, exponential, logistic, and Gompertz growth—along with "expansion" variants for nonzero ancestral populations. Implemented with Metropolis-coupled MCMC, this approach allows the sampler to switch among candidate growth functions, capturing demographic histories without a single pre-specified model. Simulation studies confirm that logistic and Gompertz models may require specialized sampling strategies (e.g., adaptive multivariate proposals) for robust mixing. We demonstrate these models' calibration and performance on datasets simulated under diverse substitution models, showing that joint inference of genealogy and population parameters remains well-calibrated when incorporating correlated-move operators and BMA.

Applying this method to real-world datasets, analysis of Egyptian Hepatitis C virus (HCV) sequences indicates that models with an early founder population followed by rapid expansion receive comparable support, with a slight preference for Gompertz-like expansions. In contrast, analysis of a metastatic colorectal cancer (CRC) single-cell dataset suggests that near-exponential growth remains plausible at advanced disease stages, revealing how tumor subclones may retain substantial proliferative capacity. Overall, our unified BMA framework reduces the need for restrictive model selection procedures and yields deeper insights into epidemic spread and tumor evolution. By integrating multiple growth hypotheses within a Bayesian setting, it naturally avoids overfitting and offers a powerful tool for inferring population histories across diverse biological contexts.

S12-P12-Transcriptomic variation in metazoans follows the drift-barrier hypothesis

Kai Mi

Shanghai Jiao Tong University (School of Life Sciences and Biotechnology)

Transcriptional variation is primarily driven by alternative transcription starts (ATS), alternative polyadenylation (APA), and alternative splicing (AS), greatly expanding the complexity and diversity of gene expression. Despite these phenomena are widespread in eukaryotes, whether most transcriptional variations are adaptive remains debatable. To the end, we tested if transcription variation follows the drift-barrier theory in this study. According to the theory, transcription variation in a species, if predominantly nonadaptive, should decrease with the effective population size (Ne) of the species due to stronger natural selection. We calculated the usage of minor ATS, AS, and APA sites to measure the extent of transcription variation for a species, and correlated the usage with the Ne of the species. By exhaustively analyzing 166 transcriptome samples from 100 metazoan species, we indeed found the usage of minor ATS, APA, and AS decreases with increasing Ne. These findings indicate that transcription variation follows the drift-barrier theory and provided novel evidence supporting the nonadaptive hypothesis of transcription variation.

S12-P13-Ancient Human Extinction Risk: Implications from High-Precision Computation and Fossil Evidence

Haipeng Li

Shanghai Institute of Nutrition and Health

Reconstructing ancient population size dynamics is essential for understanding human evolution. Using the fast infinitesimal time coalescent process (FitCoal), we identified a severe population bottleneck approximately 930,000 years ago, which pushed human ancestors to near extinction but left only subtle genomic traces in present-day humans. Detecting this signal requires precise computational methods. Compared with mushi and ten billion simulations performed using msprime, FitCoal provides the fastest and most accurate inference of expected branch lengths across site frequency spectrum categories. Independent fossil and paleoclimate records corroborate the timing and evolutionary impact of this bottleneck, highlighting the need to integrate genetic, paleontological, and environmental evidence. These findings refine our understanding of hominin evolution and the demographic events that shaped modern human ancestry.

S12-P14-Genomic Insights into Population Dynamics and Balancing Selection in Pioneer Mangrove Species

Qipian Chen

Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences

Population history and dynamics play a crucial role in shaping the genetic diversity patterns of species, which are key to their survival. Mangroves with their special geographic distribution and high susceptibility to climate and environmental changes provide a valuable model for studying the impact of historical events on variation. However, previous studies have mostly been based on limited genetic markers, which have constrained the full assessment of the impacts. Here we utilized a high-quality genome of the pioneer mangrove species Sonneratia alba and resequencing data from 14 populations across the IWP region to investigate the population dynamics shaping the current species patterns. We found that the strong population structure in S. alba was shaped by historical geographic isolation during the Pleistocene glacial periods, particularly due to the merging of Sunda and Sahul shelves. Reduced gene flow was observed between oceanic regions. The evidence of significant genetic differentiation is further supported by strong selective sweeps. We observed extremely low genetic diversity in isolated populations, likely resulting from a shared bottleneck during the Last Glacial Maximum. This loss of polymorphism appears to have been partially mitigated by long-term balancing selection at multiple genomic regions and heterozygous gene loci associated with resistance responses. Despite the highly structured populations, considerable genetic variation in NLR genes, shared across these populations and driven by balancing selection, suggests these variations play a crucial role in maintaining genetic diversity. These findings indicate that, despite historical bottlenecks and strong population structure, adaptive mechanisms have helped maintain essential genetic diversity.

S12-P15-Evolution of alternative polyadenylation after gene duplication

Wenshuai Ma

Shang Hai Jiao Tong university

Gene duplication (GD) is a primary driver of genetic diversity and innovation, while alternative polyadenylation (APA) is a critical mechanism of gene regulation. However, how APA evolves following gene duplication remains unclear. In this study, we investigated the evolutionary trends of APA after gene duplication and identified the key determinants. Our analyses revealed that duplicated genes exhibit greater polyadenylation diversity, increased usage of minor APA sites, and longer 3' untranslated regions (UTRs) compared to single-copy genes. Notably, these differences become more pronounced with the expansion of gene family size. Furthermore, we found that these differences are significantly associated with duplication timing, gene expression levels, duplication mechanisms, and gene functions. These findings offer valuable insights into how genetic and regulatory diversity emerges, adapts, and influences organismal evolution and function.

S12-P16-Sex-biased regulatory changes in the placenta of native highlanders contribute to adaptive fetal development

Yue Tian

Shandong University

Reproduction is an eternal theme of human beings. Hypoxia leads to increased miscarriage rate and fetal growth restriction. As the first population to enter the Qinghai-Tibet Plateau and successfully settle there, the Tibetans have obtained higher reproductive success rate and better reproductive phenotype after long-term natural selection, and the genetic and molecular regulatory mechanisms behind it remain to be elucidated. Therefore, in order to explore the genetic mechanism of the Tibetan population to maintain high fertility in the plateau hypoxic environment, the author team conducted a systematic reproductive phenotype and whole genome association analysis on more than two thousand pairs of mothers and newborns. It was found that compared with the Han people who gave birth on the same plateau, the Tibetans had better reproductive performance. And these results are produced by polygenic adaptation. Through the experiment of EPAS1 gene heterozygous knockout mice exposed to hypoxia, it was found that EPAS1 heterozygous knockout mice had a higher birth rate and had a protective effect on embryonic development. To further explore the molecular mechanism of plateau hypoxia affecting reproduction, through the comparative analysis of histology and transcriptome of placental tissues of plateau people, the authors found that the Tibetan had better histological and molecular characteristics, among which the plateau adaptation candidate genes may have a regulatory role, which still needs to be deeply explored. The above results reveal the phenotypic and genetic characteristics of reproductive adaptation of population living in high altitude for generations.

S12-P17-A modern human-specific regulatory variant of ACSF3 influence human height and basal metabolic rate via macronutrients metabolism

Shaohua Fan

Fudan University

Anatomically modern humans (AMHs) exhibit a 15-50% higher basal metabolic rate (BMR) and significantly increased height compared to non-human apes. This study investigates the genetic basis underlying these traits in AMHs. Our analyses reveal a strong genetic correlation between height and BMR. In addition, we identified an AMH-specific mutation, rs34590044-A, that is associated with the increased height and BMR in AMH. rs34590044-A upregulates the expression of ACSF3 via increasing its enhancer activity, leading to increased mitochondrial function, body length, and BMR in mice fed essential amino acids, specifically threonine-enriched diets, which are characteristic of meat-based diets. These results suggest that the emergence of rs34590044-A may have facilitated the adaptation to a meat-enriched diet in AMHs, with increased height and BMR emerging as byproducts of this dietary shift.

S12-P18-The evolutionary genomic study of Korean Kaolinonychus harvestmen (Arachnida, Opiliones, Paranonychidae)

Dongyoung Kim

Seoul National University

The Korean endemic harvestmen in the monotypic genus Kaolinonychus, consisting of two known subspecies, are presumed to be troglophile due to their frequent presence in the complex system of Karst caves in Korea but also found in surface environments outside caves. This makes them an intriguing system for studying incipient cave adaptation as well as biogeography of low-mobility small invertebrates. Producing a draft genome assembly and whole-genome resequencing data from 187 individuals from 51 sites in South Korea, we confirmed that the two previously described morphological subspecies (K. coreanus coreanus and K. coreanus longipes) are genetically distinct. Surprisingly, we also identified a previously unknown genetic lineage in the Mountain Gariwangsan region, which is more divergent than the two known subspecies. Our analyses reveal that K. c. longipes was in fact formed by an admixture between K. c. coreanus and the Gariwangsan lineage. Furthermore, we detected genetic contributions from the Gariwangsan in several Kaolinonychus lineages with varying degree, likely resulting from multiple gene flows between divergent lineages. Lastly, we compared individuals collected from cave and surface environments at both genomic and transcriptomic levels. Highly differentiated and differentially expressed genes include those associated with lipid metabolism and fatty acid synthesis, hinting at potential metabolic adaptations to nutrient-scarce cave environments. This study provides novel insights into the demographic history of low-dispersal invertebrates and the genomic basis of incipient cave adaptation.

S12-P19-Indirectly inherited archaic variants played a key role in facilitating adaptation in modern Peruvians

Susanna Gutierrez

Postdoctoral Researcher

Archaic ancestry, such as from Neanderthals and Denisovans, was indirectly introduced to the South American gene pool via shared ancestry and modern contact with other non-African modern humans, despite the absence of direct contact with archaic humans in the Americas. The selection history of archaic variants in the American population is therefore informative of the key timelines of the initial peopling of America and their recent admixture history can provide profound insight into refining a demographic history of South Americans.

We evaluated evidence for adaptive introgression from both Neanderthal and Denisovan in 20 genomic regions in modern Peruvians. We leveraged whole-genome sequences from modern, ancient, and archaic humans to determine which ancestral modern human population introduced the archaic haplotype to Peruvians. We then applied ancestral recombination graph methods to jointly infer the allele frequency trajectory, allele age, selection strength, and the onset time of positive selection. All archaic regions were introduced to the gene pool long before modern contact, which supports their inheritance from ancestral East Asian populations. Notably, positive selection on the archaic variants started mainly around two time points, including eight to ten thousand years ago and three to five thousand years ago. The latter of which is coincident with the settling of Peru, indicating that archaic variants played a crucial role in facilitating the settling of new environments.

This study underscores the ability of adaptive introgression to act on standing archaic variants inherited without direct contact, that subsequently contribute to the basis for local adaptation.

S12-P20-Museum specimens establish baselines for historical antimicrobial resistance in mammalian microbiomes unaffected by anthropogenic contamination

John Richards

University of Edinburgh

With the advent of global production and use of antibiotics in agriculture and medicine in the mid 20th century, anthropogenic contamination of the environment with antibiotics has created new selection pressures, promoting the evolution and spread of antimicrobial resistance (AMR). The threat of AMR is now a cause of public concern, and many governments have enacted controls on antibiotic use. Unfortunately, anthropogenic contamination is so widespread that there exist virtually no unaffected environment that could serve as a point of comparison with contemporary AMR levels. However, fossilized host-associated oral microbiota, in the form of dental calculus, are abundant in museum collections and can provide a "baseline" of environmental AMR prior to anthropogenic influence. We used ancient DNA methods to generate metagenomes from >300 dental calculus samples of 34 mammalian species with varying ecologies, diets, and geographic ranges. We characterized historical repertoires and abundance of antibiotic resistance genes (ARGs), unaffected by human mass antibiotic production, in host species with different lifestyles (marine, terrestrial, and semi-aquatic), distributed from the equator to the poles. The diversity and abundance of ARGs varies strongly among the host species, likely affected by their exposure to diverse environmental microbiomes. This comprehensive dataset will be instrumental for quantifying the impact of human antibiotic use in the last century, informing our understanding of the spread of AMR through the environment, and ultimately for evaluating the efficacy of governmental programs to combat the rising AMR.

S12-P21-Inferring the genetic regulatory landscape of hominin lineage-specific sites and their phenotypic effects

Jiahui He

Fudan University

Advances in high-coverage archaic human genome data enable the exploration of genetic differences between modern humans and archaic hominins, shedding light on their phenotypic effects. However, the regulatory landscape of lineage-specific mutations and their role in shaping human-specific traits during recent human evolution remains underexplored. In this study, we identified 9,910,635 hominin lineage-specific sites (LSS) that are divergent and highly frequent in one group among modern humans, archaic hominins, and chimpanzee. We investigated how LSS influence gene function and subsequently contribute to phenotypic differences among hominin lineages through regulatory mechanisms of gene expression and alternative splicing. Our results reveal that LSS impact regulatory networks associated with brain and neurodevelopment, embryonic development, and sensory perception. By assessing the effects of LSS on transcription factor (TF) binding affinity and TF binding site motif alterations, we highlight the role of regulatory elements in human evolution. Notably, LSS in promoters appear critical for modern human-specific traits, particularly immune responses, including hemopoiesis and viral defense. Additionally, LSS modulate gene expression via alternative splicing, especially in the nervous and olfactory systems. Furthermore, LSS modify the structure and function of RBP networks, influencing embryonic organ development in modern humans and contributing to regulatory distinctions between modern and archaic humans. Our findings provide novel insights into how LSS drive regulatory changes and shape modern human phenotypes in recent human evolution.

S12-P22-Genomic testing of hypotheses on the origins of the Grecìa Salentina population, a Greek-speaking enclave in Southern Italy

Martina Gulì

National Research Council of Italy

The origins of the Grecìa Salentina (GS) population, a Greek-speaking enclave in Southern Italy comprising 55,000 individuals across eleven communes, remain a topic of significant debate, with hypotheses pointing to either the Magna Graecia period (8th–5th century BCE) or the Byzantine era (early Middle Ages).

We performed whole-genome sequencing of 28 present-day consenting GS inhabitants. The results were compared with a subset of the large public dataset of ancient and modern individuals in the Allen Ancient DNA Resource. A total of ~7,000 European and Middle Eastern samples from the Paleolithic to the Middle Age were selected, with a particular focus on Italian, Greek and Anatolian data from the historical period critical to discriminate between alternative hypotheses on the formation of the GS community.

Principal Component Analysis (PCA), maximum likelihood genetic clustering methods such as Admixture, and F-statistics were employed to identify ancestry patterns.

As one of the last opportunities to study this community before its linguistic heritage disappears, this research provides critical insights into the history of human migration and cultural transmission in the Mediterranean. By integrating genetics, linguistics, and anthropology it addresses long-standing questions about the origins and legacy of this unique enclave. Work approved by Comitato etico ASL

Lecce, and Unione dei Comuni della Grecìa Salentina. Supported by PRIN2022 PC2TSX.

S12-P23-Ancient Beluga Genomes: Unraveling Long-Term Evolutionary Dynamics in Arctic Marine Mammals

Wenxi Li

Copenhagen University

The global environment has changed dramatically over the past few decades, especially in the Arctic, and Arctic sea ice is predicted to decline by up to 94% by 2050. By losing their habitats, ice-associated marine mammals are highly vulnerable to changes in sea ice conditions. Historically, Arctic ecosystems have experienced dramatic shifts in sea levels and sea ice cover during the Late Quaternary. Yet, how climate change has affected Arctic marine mammals on a long-term scale is still not well understood. Although next-generation sequencing technologies have significantly advanced ancient DNA research, most work has focused on macrofossils from terrestrial species. As a result, relatively little is known about marine mammals, as they live and die at sea and thus have relatively limited fossil chronologies.

In this ongoing project, we employ whole-genome shotgun sequencing and mitogenome capture on more than 50 ancient beluga whale (Delphinapterus leucas) specimens dating from the late Pleistocene to the early Holocene. By integrating genomic data with isotopic analyses, we aim to elucidate past population dynamics, migrations, and adaptations of beluga whales in response to the climate shifts. These ancient genomic resources enable us to go beyond present-day snapshots, using cutting-edge population genetics to reconstruct evolutionary histories across thousands of years. Our work not only advances our knowledge of how Arctic marine mammals coped with historical climate fluctuations, but it also offers insights into how ongoing global warming and diminishing sea ice could shape future conservation strategies.

S12-P24-The Neandertals from Stajnia Cave, Poland

Pengpeng Chen

Max Planck Institute for Evolutionary Anthropology

To date, genome-wide DNA data have been recovered from a total of 34 Neandertals from 16 archaeological sites, spanning large parts of their temporal and geographical range. These data have yielded a broad overview of Neandertal populations and their diversity, indicating the existence of multiple distinct populations over time and space, with low effective population size and small group size. However, obtaining genomic data of multiple individuals belonging to the same Neandertal group has so far been successful at only one archaeological site, Chagyrskaya Cave, at the easternmost edge of Neandertal distribution.

Located in Częstochowa Upland, Poland, the Palaeolithic site of Stajnia Cave is important both for its archaeological assemblage, as well as the discovery of multiple Neandertal remains. The site features 15 lithostratigraphic layers, with previous excavations recovering hominin remains from different units.

We screened nine Neandertal teeth from Stajnia Cave for ancient DNA preservation and successfully reconstructed eight complete mitochondrial genomes. All of the Stajnia mtDNA sequences cluster closely together, and with the previously published genomes of Stajnia S5000 and Mezmaiskaya 1 Neandertals. Our preliminary results further suggest that these nine Neandertal teeth represent a minimum of six individuals, at least three of which are male, making Stajnia one of the few sites in Central Europe with multiple Neandertal individuals. The ongoing nuclear genome analyses of these individuals aim to further refine their phylogenetic placement, and provide deeper insights into Neandertal genetic diversity at both inter- and intra-group levels.

S12-P25-Investigating the functional relevance of Denisovan introgressed DNA among New Guineans

Danat Yermakovich

Centre de Recherche sur la Biodiversité et l'Environnement, Université de Toulouse, Centre National de la Recherche Scientifique, Institut de Recherche pour le Développement, Toulouse Institut National Polytechnique, Université Toulouse 3 – Paul Sabatier, cedex 9, Toulouse 31062, France

Neandertals and Denisovans inhabited distinct regions of Eurasia for over 200,000 years before the arrival of modern humans in those territories. The following admixture events between these archaic and modern humans resulted in ~ 2% of Neandertal DNA in the genomes of present-day non-Africans, and an additional 2–5% and ~0,2% of Denisovan DNA in the genomes of Oceanians and Asians, respectively. The functional analysis of introgressed Neandertal DNA in present-day humans has identified significant links to multiple traits related to immunology, behavior, skin and hair. In turn, introgressed Denisovan DNA has shown links to metabolism, high-altitude adaptation, and immunity. However, there is still a lack of comprehensive understanding of the phenotypic impact of Denisovan DNA on their carriers due to the shortage of large genotype-phenotype association studies for Oceanian populations. As an alternative to direct measurements, recent advances in machine learning approaches applied in genomics can help to identify genetic variation with a potential phenotypic role.

In this project, we apply a new prediction-based annotation approach (Benegas et al, 2025) to explore Denisovan DNA for its predicted functional activity and distinctions compared to Neandertal and non-archaic DNA. We will annotate genetic variation in more than 300 whole genomes of present-day New Guineans, living in environmentally distinct regions of the island. In combination with other genomic features such as frequency patterns and proximity to functional units, we believe this project will enable us to discover the impact of introgressed Denisovan DNA on phenotypic variation in present-day people.

S12-P26-Genome-wide patterns of runs of homozygosity in Near Oceanic populations

Audrey Tjahjadi

Yale University

Near Oceania is a key region for understanding human dispersals, the range of human genetic variation, and the adaptive potential of archaic hominin introgression. Still, this region of the world remains underrepresented in genomic studies, leaving unanswered questions about their population history. Analyzing runs of homozygosity (ROH) can shed light on the demographic and evolutionary processes shaping genome-wide genetic diversity in human populations. We generated 177 new Near Oceanic genomes from the Bismarck Archipelago and analyzed them alongside 1,044 individuals from global populations to examine population-specific patterns of ROH and perform a comparative, quantitative analysis of ROH in Near Oceania. Our results show substantial amounts of extended ROH segments and heterogeneity in the geographic distribution of ROH, corroborating previous findings of historical bottlenecks and population structure in these island populations. By characterizing the distribution of ROH segments along the genome, we also identified regions of high ROH frequency (ROH "hotspots" or "islands") that overlap targets of selection and expression quantitative trait loci (eQTLs) affecting genes involved in adaptive immune response (e.g. the IGK and IGH loci) and metabolism (e.g. MTMR9). Our results contribute to a more detailed understanding of how the complex population history of Near Oceania has affected genome-wide patterns of homozygosity and provides insight into the impact of ROH on human biology and functional genetic variation.

S12-P27-Exploring the genetic basis of dispersal propensity in rapidly colonizing endangered wading bird, the Oriental Storks Ciconia boyciana

Qing Chen

School of Ecology, Sun Yat-sen University, Shenzhen, China

Dispersal behavior plays a crucial role in shaping population demography and genetic structure, significantly influencing ecological and evolutionary process. Despite theories propose that dispersal behavior serves as an adaptive strategy for individuals to respond to competition, inbreeding, and fluctuations in resource availability, the genetic mechanisms governing dispersal propensity during rapid colonization processes remain insufficiently understood. In this study, we collected the 66 samples of the Oriental Stork (Ciconia boyciana) from traditional breeding regions (Sanjiang Plain and Songnen Plain, China) and newly established breeding region (Yellow River Delta, China) and performed whole-genome resequencing to explore the genetic basis and consequences of dispersal behavior, individuals in the Yellow River Delta exhibit both sedentary and migratory behaviors. Our results showed that there was no significant genetic differentiation from different breeding regions, but individuals within the same breeding region exhibited large genetic differences, indicating that dispersal behavior may be influenced by complex genetic factors. Ancestry inferences revealed that groups in the Songnen Plain and the Yellow River Delta likely originated from different ancestral populations, suggesting distinct demographic histories. Although overall genetic homogeneity, the group in the Yellow River Delta showed a significantly increase inbreeding level in recent breeding. For the group in the newly established breeding region, we found that genes related to nervous system development were selected, which may be associated with enhanced learning and cognitive abilities to adjust dispersal and migratory behaviors. These findings provide new insights into the early mechanisms associated with changes in migration behavior.

S12-P28-Ancient and modern genomes of the common vole, Microtus arvalis (Rodentia, Cricetidae), reveal the complex evolutionary history of Western Carpathian populations

Alicja Anna Kaźmierkiewicz

University of Warsaw

Microtus arvalis is a rodent species widespread in Europe present and past. In much of Central Europe extant population belongs to the Eastern mtDNA lineage. Previous studies have identified a Central mtDNA lineage in the same region in the Late Pleistocene and Early Holocene suggesting a population turnover at the Pleistocene/Holocene transition. However, the course of this replacement remained unclear. To further understand the evolutionary history of the common vole we generated 13 Late Pleistocene and early Holocene low coverage genomes from paleontological sides across Western Carpathians, and 39 mid- to high- coverage modern genomes from Central Europe. Principal Component Analysis (PCA) of newly generated and previously published data revealed two main clusters: one comprising modern samples belonging to Eastern mtDNA lineage, and the other containing ancient samples from both Eastern and Central mtDNA, and a few modern Eastern samples. This suggests that the expanding Eastern lineage mixed with the local Central population; however, traces of this admixture are not readily visible in the extant population. These results raise several questions, including the reasons for population expansion into areas occupied by other population of the same species, the potential advantageous adaptations of the Eastern population to the environment, whether the signal of such adaptations can be detected in the genome, and why do modern populations in the Western Carpathians show no trace of the Central population. Overall, preliminary results of our project suggest a more complex evolutionary history of the common vole in the Western Carpathians than previously speculated.

S12-P29-Prehistoric Global Migration of Vanishing Gut Microbes with Humans

Zhiru Liu

Stanford University

The gut microbiome is crucial for health and is greatly affected by lifestyle. Many microbial species common in non-industrialized populations are disappearing or extinct in industrialized populations. Understanding which microbes have been long-term residents of the human gut, and thus may have co-evolved with humans over evolutionary time, has implications for the importance of microbial biodiversity loss for health. However, the genetic complexities of bacterial evolution and the plasticity of the gut microbiome have made it challenging to define these long-term associations. Here, we performed deep metagenomic sequencing of the Tsimane horticulturalists of Bolivia and compared their gut microbiomes to those of the Hadza hunter-gatherers of Tanzania. These two populations, whose ancestors were separated for tens of thousands of years, share 1,244 bacterial species, most of which are absent in industrialized populations. Population genetic analyses in 541 of these shared bacterial species revealed patterns of microbial divergence and gene flow consistent with prehistoric human co-migration, with estimated split times that approximately align with human migration out of Africa and into the Americas. Our findings indicate that a diverse gut microbiome co-migrated with humans around the globe, persisting over millennia. However, many of these species are now vanishing in industrialized populations, and the consequences for human health remain uncertain.

Symposium 27: Population genetics through time session B

S27-P01-The genomic history and adaptation of the chickens

Yuemin Li

Queen Mary University of London

The chicken (Gallus gallus domesticus) is one of the most numerous and wide-spread domestic animal species today. Nonetheless, little is known about the early history of chicken, including the timing and circumstances of its domestication and the spatiotemporal patterns of the dispersal following the domestication. Furthermore, the gene pools of modern domestic chickens have been rapidly and radically shaped by the process of recent breed formation involving strong selection and admixture between breeds, which often distorted the deeper genetic history. To reconstruct chicken population history and investigate the early global dispersal events at high temporal and spatial resolution, here we analysed 115 ancient chicken genomes spanning the last 2,500 years from Eurasia and South America imputed with a comprehensive modern genome panel of around 5,000 domestic chickens and wild junglefowl. These ancient samples represent a vast transect of Eurasia, longitudinally from Uzbekistan to the Atlantic Coast and latitudinally from Northwestern Russia to Southern Iran. The population continuity observed has provided a great opportunity to detect natural selection across the time series. We show that some selective sweeps only occurred very recently in the European chickens. This extensive time series dataset allowed us to characterise the genetic composition and transition in fine details, shedding light on the genomic history of chickens.

S27-P02-From North Asia to South America: Tracing the longest human migration through genomic sequencing

Hie Lim Kim

Nanyang Technological University

During the late Pleistocene, humans expanded across Eurasia and eventually migrated to the Americas. Those who reached the southern tip of South America, completed the longest migration out of Africa. To reveal this history of human migration, we analyzed genome sequencing datasets of 1,537 individuals from 139 ethnic groups including understudied indigenous populations in North Asia and the Americas. Our analysis demonstrated that the nomadic hunter-gatherer West Siberians, represented by the Kets and Nenets, contributed to the genetic ancestry of most Siberian populations. This is supported by our estimation that West Siberians were the largest population in North Eurasia 10,000-13,900 years ago. We also found that West Beringians, including the Koryaks, Inuit, and Luoravetlans, exhibit genetic adaptation to the Arctic climate. In South America, our analysis showed that the earliest inhabitants of South America split into four ancestral groups, Amazonians, Andeans, Chaco Amerindians, and Patagonians, ~13,900 years ago. This population structure corresponds closely with biogeographic boundaries in South America, suggesting the impact of environmental factors on population history. Their longest migration and spatial isolation due to the vastness of the South American continent led to population decline and reduced genetic diversity. The Patagonian ethnicities, including the Yagan, Kawésqar, and Chaco Amerindians, show the smallest effective population size in our dataset. This, in turn, has resulted in reduced immunogenic diversity, likely increasing their vulnerability to past and current pathogens. These findings highlight how population history and environmental pressures shaped the genetic architecture of human populations across North Asia and South America.

S27-P03-HLA Genetic Diversity and Immune Response in Ancient Mexican Populations: Insights from the Prehispanic and Colonial Periods Walter

Nicolas Ortega

LIIGH, UNAM

The colonial period in New Spain brought profound demographic and genetic shifts due to the blending of Indigenous, European, and African populations. However, epidemics between 1520 and 1580 caused catastrophic population declines, with over 80% of the Indigenous population perishing. Differences in adaptive immunity, particularly in the HLA system, may have contributed to disparities in survival. To explore HLA diversity and its role in immune responses during these epidemics, we analyzed ancient DNA (aDNA) from 59 individuals (54 colonial, 5 pre-Hispanic). HLA regions were enriched using myBaits in-solution capture and sequenced via next-generation sequencing (NGS). HLA genotyping was performed with Optitype and TARGT pipelines. We further applied NetMHCPan and NetMHCIIpan to predict the binding affinity of HLA variants to pathogen-derived peptides, providing insight into immune responses that may have influenced survival during the colonial period. Our results yielded high-quality data with complete coverage of 26 HLA genes and an average sequencing depth of ~25x in high-quality libraries. We identified multiple HLA alleles, including HLA-A02:01, HLA-B35, HLA-C04:01, HLA-DPA101:03, HLA-DPB104:02, HLA-DQA103:01, HLA-DQB103:02, and HLA-DRB108:02. Their frequencies align with those observed in present-day Mexican populations, suggesting genetic continuity from the colonial period to modern times. Additionally, NetMHCPan and NetMHCIIpan analyses identified HLA variants with strong predicted binding to pathogen-derived peptides, highlighting potential immune advantages during historical epidemics. Overall, this study provides insights into genetic factors that may have influenced survival outcomes during one of the most significant demographic collapses in the Americas.

S27-P04-Evolutionary analysis of diet adaptations in anatomically modern humans using ancient genomes

Carla Casanova Suárez

Institute of Evolutionary Biology

Anatomically modern humans rapidly adapted to diverse environments following the Out of Africa event in a very short period of time from an evolutionary point of view. This rapid adaptation was facilitated by an interplay between cultural and biological factors, including significant dietary changes over the past millennia due to cultural revolutions. Such changes are hypothesized to have influenced human genomes, exemplified by the FADS gene family.

Diet-related traits encompass a complex interplay of behavioral, metabolic, and anthropometric phenotypes and are defined by a large number of pleiotropic genomic factors. Previous efforts to infer and detect polygenic selection often faced limited statistical power due to weak signals, confounding demographic histories, or stabilizing selection effects. However, the rise of large ancient DNA (aDNA) datasets provides a genetic snapshot of past populations and enables novel approaches for studying these complex phenotypes.

In our study, we retrieved Genome-Wide Association Studies (GWAS) and available aDNA data from European populations to explore the evolutionary history of diet-related traits. By applying specifically designed machine learning methods, we examined the relationships between genetically correlated traits. Our findings suggest a dynamic interplay of evolutionary forces driven by pleiotropy that have shaped human dietary adaptations over the past 30,000 years.

S27-P05-A novel Bayesian method to perform demographic inference from genomic data

Tommaso Stentella

Max Planck Institute for Molecular Genetics

Segregating sites and their statistical distribution along genomes contain rich information about the degree of divergence of homologous sequences in a population. The Sequentially Markov Coalescent approximation to the Ancestral Recombination Graph and Hidden Markov Models are commonly used to infer Identity by Descent segments jointly with the population history.

Here we introduce a novel methodology which circumvents the inference of recombination events and focuses on the demographic history. We argue that it is in fact sufficient to consider only the distribution of distances between segregating sites. This quantity is directly observable and inherently a marginal distribution over both the ancestral recombination and coalescence events. Using our theoretical framework, we derive an analytical formula for the distribution of distances between segregating sites conditional on an arbitrary panmictic demography which allows us to compute the likelihood function and leverage Bayesian analysis to reason about model complexity, i.e. only consider histories with sufficient evidence from data.

We validate our method with extensive simulations, showing that it is possible to precisely retrieve the true demographic model. Further, we can show that our analysis is well powered to accurately recover information about the recent past as well as the ancient past of a population. The simplicity of our approach makes it possible to analyze demographic histories to unprecedented accuracy for many species, not limited to humans, but also including endangered species for which only one genome is available.

S27-P06-Genomic Insights into Demographic History, Inbreeding, and Mutation Load of Juglans

Jiamei Peng

Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, College of Life Sciences, Beling Normal University

Tertiary relict plants, as unique branches on the Tree of Life, would profoundly impact seed plant lineage diversity patterns if driven to extinction. However, we know little about how demographic histories influence the accumulation and purging of deleterious mutations that determine their long-term survival and extinction risks. We analyzed genomic data from 310 individuals across 20 species of Juglans, a Tertiary relict genus, employing multiple approaches to reconstruct their demographic histories over the past approximately 5 million years and potential recent hybridization events. Further analyses were conducted on genetic load, inbreeding, and genetic diversity. Results reveal that repeated population expansions and contractions in most Juglans species were driven by Quaternary glacial-interglacial cycles combined with species-specific factors such as human domestication and co-evolutionary interactions with pathogens. American species with small isolated populations showed prolonged population decline accompanied by severe inbreeding or elevated genetic load. Notably, the widely distributed Eurasian Juglans regia exhibited significant inbreeding depression, reduced genetic diversity, and pronounced mutation load in stress-resistance-related genes, potentially linked to prolonged domestication. Our study highlights how demographic histories shape deleterious mutation dynamics in Tertiary relict plants, providing crucial insights for genetic management of deleterious variation in small isolated populations.

S27-P07-Investigating the Global Distribution, Evolution, and Selection of ADH1B *2 and *3 Alleles in Modern Human Populations

Aine MacDermott

University of Richmond

The alcohol dehydrogenase enzyme, encoded by ADH1B, plays a crucial role in alcohol metabolism. The derived *2 (rs1229984) and *3 (rs2066702) alleles increase the turnover rate of the alcohol dehydrogenase complex, conferring protective effects against alcohol dependency and influencing the risk of various cancers. Evidence of positive selection has been found in several populations with these alleles, particularly the *2 allele. Here, we aimed to characterize the geographical distribution and evolutionary trajectories of the *2 and *3 alleles worldwide. For both SNPs, we calculated allelic and haplotypic frequencies across modern human populations and generated haplotype networks and phylogenetic trees to visualize the evolutionary relationships between these haplotypes. We also assessed local ancestry around the ADH1B gene to study the impact of recent admixture on haplotypic distributions. Finally, we inferred the presence, timing, and strength of positive selection at the *2 and *3 loci. The systematic analysis of the ADH1B gene in this study refines our understanding of introgression and selection at this gene in worldwide modern human populations.

S27-P08-Exploring how past geological events influenced the genetic structure of an endemic lizard in the Baja California Peninsula.

Diego Cortez

University of Mexico

S27-P09-The Zagros Mountain-Kavir Desert Migration Corridor: How the Iranian Plateau Shaped Eurasian Ancestries Throughout Millennia.

Nada Elsharkawy

New York University Abu Dhabi (NYUAD)

The Iranian Plateau, bisected by the Zagros Mountains and Kavir Desert, has served as a pivotal migration corridor and genetic crossroads since humans began migrating out of Africa. Despite Iran's centrality to human dispersals, its populations remain underrepresented in genomic studies. Thus, we hypothesize that studying Iranian genomes may reveal ancient and recent traces of Eurasian migrations. To investigate this, we analyzed 5 million variants from whole-genome sequences (WGS) of 73 Iranians alongside WGS data from 4,769 individuals worldwide and 600 ancient DNA (aDNA) samples imputed with high accuracy. Our analysis revealed how the Zagros-Kavir corridor structured genetic exchange, partitioning the plateau into zones that drove subpopulation differentiation through distinct genetic admixtures with Europe, the Caucasus, Western Asia, South Asia, and Central Asia. Integration of aDNA demonstrates temporal variations: Neolithic Iranians were highly distinct from neighboring groups 10,000 years ago; by the Chalcolithic, intensified corridor interactions drove admixture, homogenizing regional ancestries around 6,500 years ago. Modern Iranians exhibit pronounced genetic clines mirroring the Zagros-Kavir axis: northwestern groups display ancestry from ancient Caucasus and Anatolian sources (seen in modern Levantine, Caucasian, and European populations), while populations flanking the desert-mountain interface, closer to South and Central Asia, retain stronger genetic continuity with Neolithic Iranians. Interestingly, these patterns align with linguistic distributions. Thus, we reconstruct two crucial spatiotemporal processes: Eurasian influxes into the plateau and Iranian ancestry spreading via northern and eastern migrations, demonstrating how the plateau's dual role as conduit and barrier sculpted Eurasia's genetic landscape, leaving enduring imprints over millennia.

S27-P10-Evolutionary processes affecting genomic and phenotypic variation in natural populations of the green anole

Mariana Lyra

New York University Abu Dhabi

Untangling the factors that drive morphological evolution has long been a central focus in the study of evolutionary biology. However, the interplay between genomic and environmental variation that mediates phenotypic change and adaptation remains poorly understood in most species. In this study, we used natural populations of the green anole (Anolis carolinensis) to explore how demography and selection have shaped the current genetic and phenotypic diversity in this species. For that, we used whole genome sequencing approach on four distinct geographical populations distributed across the current range that were phenotypically characterized. The primary axis of phenotypic variation among these populations falls along a physiological and life-history gradient, likely driven by latitudinal changes in mean temperature. Our results shows that the Florida population is highly divergent from those in Louisiana, South Carolina and Tennessee. Our results also shows that Louisiana and Tennessee populations are closely related. The demographic analysis indicates a rapid population expansion from southern populations to north following a bottleneck, as suggested by previous studies. Genomic scans of selection have identified genes related to metabolism, inflammation, growth, and exploratory behavior to be associated with the environmental cline. This study brings new insights into the process of local adaptation in Anolis carolinensis.

S27-P11-Unraveling 500 Years of Evolutionary History and Admixture Through IBD Genomic Analyses in the Mexican Biobank

Carmina Barberena Jonas

CINVESTAV

Mexico is a country characterized by its rich diversity and a complicated history of human migrations and genetic mixing over various periods. The impact of colonization introduced European and African genetic influence throughout Latin America, resulting in a modern genetic profile that reflects a blend of different ancestries. Coupled with the scarcity of archaeological evidence, this complexity presents challenges in understanding the evolutionary shifts that have occurred over time. In this study, we utilize genomic signals from varying lengths of identity-by-descent (IBD) segments to reconstruct the demographic changes that have influenced the Mexican population, particularly in recent centuries. To achieve a comprehensive analysis at a national level, we examined IBD segments in 6,057 individuals from all 32 Mexican states, leveraging SNP array data from the MX Biobank. Our unique methodology involved analyzing different lengths of IBD in conjunction with local ancestry, revealing insights into the interconnections of time, geography, and ancestry. Using network analyses, we identified distinct clusters of individuals across different regions of Mexico, reflecting the country's population structure. By integrating IBD and local ancestry inference, we leveraged modern genetic data to explore temporal patterns and historical influences on population dynamics. This approach demonstrates the power of clustering methods in uncovering genetic structure and its evolution over time.

S27-P12-High-Resolution Genomic Insights into the Peopling of Hawai’i

Edward Huang

Stanford University

The settlement of the Hawaiian Islands, located ~3000 kilometers away from the nearest continental landmass, represents one of the most impressive navigational feats in human history. Archaeological evidence variably indicates that the Hawaiian Islands were peopled by Polynesian explorers from the Marquesas Islands or the Society Islands, as early as the 3rd century CE or as late as the 13th century CE.

Modern whole-genome sequences from Hawai’i and other Polynesian islands offer many insights into the settlement of the Hawaiian Islands. Although Polynesian whole genomes are underrepresented in existing datasets, Ancestral Recombination Graph (ARG)-based methods such as Relate and Twigstats offer a promising approach to improve the statistical power of genomic analyses of these founder populations. We use these ARG-based population inference methods to uncover genomic insights into the peopling of Hawai’i, including estimates of Hawaiian population size through time. We also leverage these methods to infer the founding population of Hawai’i, as well as its settlement date and ensuing admixture events. As such, we show that the Hawaiian archipelago offers a fitting canvas for ARG-based methods and their ability to infer the population genetics of the past at a high resolution.

S27-P13-Genomic Insights into North Eurasian and South American Indigenous Populations Shaped by Diverse Climatic Conditions

Amit Gourav Ghosh

Nanyang Technological University

Eurasia and America harbour diverse human populations with distinct cultural and lifestyle backgrounds shaped by their diverse environmental conditions. Despite such diversity, these populations remain understudied, and their demographic history and adaptive responses to environments are largely elusive due to underrepresentation in large-scale genome studies and reliance on low-resolution genotype data. Using fine-resolution whole-genome data from 140 populations generated by GenomeAsia 100K, we aim to infer the demographic processes and examine genetic adaptations to diverse climatic conditions shaping population diversity across North Eurasia and South America. Our findings reveal genetic adaptations of East Siberian and West Beringian populations against Arctic climate conditions. We found strong signals of positive selection in the gene related to high-fat diets through fatty acid metabolism. Indigenous populations in South America show distinct population structure which has an apparent link with particular environments and climates: Andeans in high altitude, Amazonians in tropical forests, Chaco populations in semi-desert Gran Chaco, and Patagonians in frigid areas.

In the Andes, hypoxic conditions drove adaptation in genes involved in oxygen-sensitive glycosaminoglycan metabolism. Among Amazonian hunter-gatherers, we find environment-mediated adaptive responses in genes involved in energy metabolism, skeletomuscular functions, and immunity. In Gran Chaco, the populations adapted to the arid, hot conditions through selection in genes influencing thermal stress tolerance and kidney water reabsorption, with selection signals originating 3,500 years ago. In Patagonia, adapting to extreme cold involves selection in genes influencing lipogenesis and fatty acid metabolism. Our study underscores climate's pivotal role in shaping human genetic diversity.

S27-P14-Exploring genetic diversity and demographic history of Southeast Asians in warm environment through whole-genome analysis.

Tran The Hung Nguyen

Nanyang Technological University

Southeast Asia (SEA) human populations possess immense genetic diversity shaped by complex demographic history and environmental shifts. Despite its significance in understanding human migration patterns, the genetic history of Southeast Asian populations remains underexplored due to the limited availability of high-resolution genomic data and the lack of sampling of indigenous people in the region. Previous research on population genetics in SEA is mostly limited to mitochondrial DNA and low-resolution SNP analyses, which fail to capture the fine-scale genetic structure and complex admixture patterns of the populations. To address these gaps, the GenomeAsia100K project generated an extensive population genome dataset for the SEA region. This dataset includes 3971 genomes from 231 ethnicities majority from SEA. We performed population genomic analyses to reveal the population structure and diversity of the SEA populations using this dataset which has a particular focus on Myanmar ' s eight major ethnicities. Our study uncovers the complex ancestral structure and historical admixture of Myanmar ' s populations which have been underexplored. Moreover, we seek to address wider questions about how historical climate change may have influenced the region's migration patterns and genetic diversity. Particularly, permafrost and desertification on the Tibetan plateau during the last glacial maximum (LGM) and sea level rise post-LGM in Southeast Asia could influence the migration of ancient populations.

S27-P15-Mobility in East and West Eurasia through the Holocene: a comparison of temporal dynamics through paleogenomics

Mehmet Somel

ODTU

During the Holocene, both West and East Asia experienced similar cultural and economic changes, starting with sedentary life and food production, followed by the adoption of metallurgy and long-distance transport technologies, and the emergence of hierarchical social structures. In recent work, we described mobility and admixture patterns in Southwest Asia over the last 10,000 years as reflected in ancient genomic data. These included several temporal trends: increasing intraregional diversity, regional genetic homogenization during the Neolithic Transition followed by a widening range of mobility after the Bronze Ages, and an increase in relative male-bias in admixture signals. Here we argue that these observed changes in the Southwest Asian gene pool were shaped more by background (continuous, low-level) mobility rather than migration (unidirectional and episodic mass movement events). We further study mobility patterns during the Holocene in East Asia using 1000 published ancient genomes. Interestingly, we do not find the above-mentioned three trends in this region. We propose a number of possible explanations for the observed differences between the West and East Asia, including differences in the scale of mobility events and geographic constraints to regional connections, as well as historical contingency.

S27-P16-Development of a High-Throughput Molecular Inversion Probe Panel for Cost-Effective Genotyping in the Chinese Mitten Crab (Eriocheir sinensis)

Junlong Jiang

Shanghai Ocean University

The Chinese mitten crab (Eriocheir sinensis), a key aquaculture species in China, faces growing challenges from genetic admixture in breeding populations, necessitating rapid and scalable genomic screening tools. While resequencing and SNP arrays provide genetic insights, their high cost and labor-intensive workflows limit large-scale breeding applications. To address this, we developed a Molecular Inversion Probe (MIP) panel tailored for E. sinensis, leveraging its proven utility in medical diagnostics and crop breeding for high-throughput, cost-effective genotyping. Our panel targets 854 SNPs across 50 loci spanning all 40 chromosomes. A refined protocol incorporating dual unique molecular indexes (UMIs) enabled consensus sequence generation and SNP calling via a custom bioinformatic pipeline. From an average of 224,439 paired-end reads per individual, low-quality reads (4%) were filtered, yielding uniform coverage across sites (median proportion: 1.598%) and an average UMI depth of 11.71×. The MIP panel achieved 91% genotyping accuracy, surpassing GATK-based SNP calling at comparable sequencing depths. Validated on 55 individuals from four geographically distinct populations (Yellow River, Liaohe River, Russia, Nanliu River), the panel resolved population structure via PCA and STRUCTURE analysis, clustering individuals by geographic origin. This demonstrates its capacity to deliver high-resolution genetic data with reduced sequencing effort. Our study establishes MIPs as a robust, cost-effective tool for genomic selection in aquaculture breeding programs. By combining high accuracy with scalable workflows, this approach addresses critical gaps in genetic management of E. sinensis and offers a model for other aquatic species.

S27-P17-Mothers on the move - thousands of complete mitochondrial sequences reveal the time-layered formation of the shared maternal gene pool in Estonia and Finland

Anne-Mai Ilumäe

Institute of Genomics, University of Tartu

Phylogeographic reconstructions of European maternal history have mostly highlighted the expansions from glacial refugia. Here we have focused on a large-scale population-based examination of the maternal genepool of northeastern Europe, utilising complete mitochondrial sequences.We collected over 6000 full mtDNA sequences from Estonia, Finland, Sweden, Latvia and Poland with additional modern haploid genomes from published sources. We reconstructed phylogenetic trees for all of these sequences, established monophyletic lineages found in single populations or shared between several focus populations, estimated coalescence times for these lineages and the dynamics of effective population size.As the result, we find almost 40% of Finnish individuals belong to monophyletic clusters common with linguistically related Estonians, whereas below 8% are shared with geographically close Swedish population. In Estonia, roughly 15% of individuals belong to shared Finnish-Estonian monophyletic clusters and less than 2% of Estonian individuals form haplogroups with Swedes. For the shared proportion of the maternal gene pool, effective population size increases in the late Bronze and Iron Age.

The phylogeny informed approach based on a maximally resolved mtDNA tree allows us to detect a geographic pattern previously associated primarily with male lineages. We demonstrate and discuss the importance of Bronze and Iron Age migrations on the formation of the eastern Baltic maternal gene pool by combining multiple line of evidence from modern and ancient mtDNA as well as archaeology.

S27-P18-Temporal dynamics of adaptive evolution: Genomic signatures of selection across millennial-scale environmental adaptation

Liang Lu

Shanghai Ocean University

Understanding how genomic architectures transition from initial environmental adaptation to stabilized evolutionary equilibria remains a central challenge in evolutionary biology. The migratory tapertail anchovy (Coilia nasus) provides a natural model for investigating this dynamic: its ancestral migratory population diverged into two freshwater-resident lineages isolated in Lake Chaohu (~3,000 years old) and Lake Dongting (~4 million years old). Leveraging genome level data of 303 individuals(97 resequencing and 206 gene capture), we integrated population genomic structure analysis, selective sweep detection (SweeD), and multilocus evolutionary metrics (FST, Tajima ' s D, LD decay) to resolve temporal patterns of adaptive evolution. The recently diverged Chaohu population retained elevated genetic diversity (π = 0.0027 vs. 0.0011 in Dongting), while the ancient Dongting lineage exhibited pronounced diversity loss, consistent with progressive relaxation of selection pressures (SweeD top 5% mean likelihood: CH 1.71 vs. DT 0.86). Functional annotation of selected regions revealed divergent selection trajectories: total length of positively selected sequences and LD decay rates scaled with adaptation duration, yet various genome elements (e.g., coding vs. noncoding regions) displayed different temporal patterns in selection intensity. Multivariate analysis uncovered stage-dependent covariation among evolutionary parameters, suggesting context-dependent coevolutionary modes shaped by varying functional constraints. This millennial-scale genomic chronosequence elucidates the continuum of adaptive evolution, providing a framework to reconcile transient selection signatures with enduring evolutionary outcomes in complex traits.

S27-P19-Anthropogenic Events and Their Impact on Species' Genomic Diversity: A Case Study from the Natufian Period to Recent Times

Ksenia Juravel

HUJI

As a biogeographic corridor linking Europe, Africa, and Asia, Israel is a biodiversity hotspot undergoing significant environmental changes due to natural processes and human activities. One of the major challenges in conservation biology is understanding how species respond to these changes over time. Archaeological and historical records indicate that climate fluctuations and human interventions, such as domestication and land-use modifications, have shaped local fauna, leading to behavior, morphology, and genetic diversity shifts. However, the long-term genetic impacts of these processes remain poorly understood.

This study focuses on species inhabiting the Hula Valley, a unique ecosystem that has experienced dramatic anthropogenic changes, from early agricultural settlements in the Natufian period (~13,000 years ago) to extensive land modifications in modern times, including large-scale drainage (~70 years ago) and subsequent reflooding. The Eynan archaeological site serves as a key reference point, preserving evidence of biodiversity before major human disturbances.

Through collecting and analyzing contemporary and ancient specimens, this research examines population dynamics, pangenomic diversity, and the challenges of working with ancient DNA. By studying key species such as the mountain gazelle (Gazella gazella) and the jungle cat (Felis chaus), we aim to reconstruct historical biodiversity patterns and assess long-term changes in genetic diversity. These findings will provide valuable insights into adaptation processes and inform conservation strategies for maintaining ecosystem stability in rapid environmental change.

S27-P20-Paleoclimatic events shaped the genetics, migration and microbial differentiation of two incipient gull species

Chenchang Qian

Kunming Institute of Zoology, CAS

Different climatic and geographic events have shaped bird biodiversity. Migration behavior could reveal the differentiation of genomics. Microbiome biodiversity is related to host species at the phylogenetic level. Utilizing GPS tracking, whole genome resequencing, and microbiome technologies, we can understand genomic adaptation for bird flyway differentiation and evolutionary dynamics. As a naturally species flock, black-headed gull and brown-headed gull, sharing similar characteristics and overlapping habitats, which makes they ideal models for migratory evolution and virus spreading studies.

We obtained a high-quality chromosome-level genome assembly and annotation of black-headed gull, de novo genomes and microbiome libraries of both species. Population genetic analysis shows that the two species separated into independent groups, with brown-headed gull diverging from black-headed gull. Black-headed gull also shows sub-population divergence along geographic distance. Genetic introgression analysis reveals potential gene flow between them. Demographic history inference suggests they diverged about 10,000 years ago, coinciding with the end of the Last Glacial Maximum (LGM). Selective sweep analysis indicates positive selected loci related to high-altitude adaptation, matching the different elevation distributions of the two species.

Our results suggest brown-headed gull is a sub-clade of black-headed gull, forming a single clade, with paleoclimatic events influencing early divergence. Microbiome analysis shows differences in α and β-diversity between the two species, indicating evolutionary forces driving microbiome divergence.

S27-P21-Population genomics and altitudinal adaptation of a spiny crawler mayfly, Drunella ishiyamana (Ephemeroptera: Ephemerellidae), in Mongolia

Ji Hyoun Kang

Korea University

Mayflies play a crucial role in freshwater ecosystems, serving as bioindicators of water quality and ecosystem health. As primary consumers, they function as herbivores and detritivores, contributing to nutrient recycling in streams and rivers. Their sensitivity to temperature fluctuations and hydrological changes makes them valuable indicators of climate change impacts, such as increased water temperatures and altered streamflow patterns. Among them, the spiny crawler mayflies (genus Drunella, family Ephemerellidae) are well known for their altitudinal distribution and specialization in microhabitat partitioning. This study investigates the population genomics of Drunella ishiyamana from eight populations in Mongolia to assess altitudinal population structures and their ecological adaptations. Genetic diversity analysis revealed moderate variation across populations, with observed heterozygosity ranging from 0.298 to 0.371 and inbreeding coefficients (FIS) between -0.097 and 0.055. Pairwise FST values indicated moderate genetic differentiation, suggesting localized adaptation or restricted gene flow. Genotyping was performed using the de novo module of Stacks, identifying 17,592 SNP loci after filtering for no missing data and a minor allele frequency (MAF) $\geq$ 0.05. The STRUCTURE analysis identified K = 4 as the optimal number of genetic clusters, indicating the presence of four distinct populations of D. ishiyamana in Mongolia. This result suggests structured genetic differentiation among populations, likely influenced by altitudinal distribution and localized adaptation. This study provides valuable insights into gene flow, population structure, and local adaptation in D. ishiyamana, emphasizing its role as a key indicator species in freshwater ecosystems.

Symposium 17: Thematic symposium: regulatory genomics and non-coding elements

S17-P01-Identification of RNA binding partners of AtGRP7 using immunoprecipitation-free techniques across Brassicas

Xingyu Yan

The University of New South Wales

RNA-binding proteins (RBPs) interact with RNA molecules and play a key role in regulating gene expression at three stages: translational, post-transcriptional, and pre-mRNA processing. Numerous novel RBPs have been identified in plants in the last decade, particularly in Arabidopsis thaliana. Specifically, the glycine-rich RNA-binding protein 7 (AtGRP7) was found to be involved in circadian rhythm, RNA processing, and the regulation of flowering time due to its binding ability to various RNAs. Hundreds to thousands of AtGRP7 RNA targets have been identified using RNA immunoprecipitation-sequencing (RIP-seq) and individual nucleotide-resolution crosslinking and immunoprecipitation (iCLIP). However, all of the research involving the RNA targets of AtGRP7 is done using immunoprecipitation-based methods, and there is currently no evidence that immunoprecipitation-free methods, which are easier to handle, cost-effective, and use extremely low or even single-cell inputs, have been successfully used to understand their interactions. To further improve the system of methodological applications for plant RNA-protein interaction studies, we applied the immunoprecipitation-free method APOBEC-mediated profiling (STAMP) and the hyperTRIBE approach. The former involves fusing the C-to-U deaminase APOBEC1 to AtGRP7, while the latter requires the fusion of AtGRP7 to the A-to-I deaminase ADAR. Targets are identified by finding modified RNAs with respect to the unfused control. By comparing our targets with previously reported targets of AtGRP7, we have been able to measure the effectiveness of immunoprecipitation-free techniques in A. thaliana. We are working on expanding this technique to other Brassicas to better understand how RNA targets of this RBP change across the phylogeny.

S17-P02-The role of differential methylation in phenotypic divergence between humans and nonhuman African apes

Chen Leibson

Hebrew University of Jerusalem

Human lineages, including anatomically modern and archaic humans such as Neanderthals and Denisovans, exhibit significant morphological and behavioral differences from chimpanzees and gorillas. Because regulatory changes are major drivers of phenotypic differentiation, and as the best proxy of gene expression in ancient samples is DNA methylation, we wished to compare DNA methylation maps across anatomically modern humans, Neanderthal, Denisova and nonhuman African apes.

To account for the skeletal origin of DNA methylation in ancient samples, we measured whole-genome bone DNA methylation in chimpanzees and gorilla, and compared it to reconstructed whole-genome methylation maps of archaic and anatomically modern humans. We found dozens of differentially methylated regions (DMRs) separating human and nonhuman African apes. These DMRs are significantly associated with genes related to the formation and function of the heart, particularly of the primary and secondary pacemakers, in addition to skeletal system morphogenesis, limb development, and cranial nerves formation.

Additionally, we find that in most DMRs, humans exhibit hypermethylation compared to the other African apes. In light of previous findings that DMRs separating modern and archaic humans tend to be hypermethylated in modern humans, this could suggest that the features we recognize today as characterizing modern humans are a continuation of a regulatory process that began much earlier in our lineage ' s evolution.

S17-P03-The genetics of gene expression divergence in wild baboons

Shuyu He

Duke University

Understanding primate evolution requires studying the contribution made by gene regulation. While much progress has been made in understanding the genetic basis of gene regulatory traits in humans, much less is known about how gene regulation in primates evolves on deeper timescales. To address this gap, we integrated genome-wide genotyping and RNA-seq data from 160 wild baboons in the Amboseli ecosystem of Kenya. This population descends from ancient and recent hybridization between yellow baboons (Papio cynocephalus) and anubis baboons (P. anubis; ~1.4 million years diverged). We collected paired control and lipopolysaccharide (LPS)-stimulated blood samples to simulate the response to bacterial infection. We identify cis-acting expression quantitative trait loci (eQTL) for over 20% of blood-expressed genes (FDR<0.1). For ~5.5% of expressed genes, we also identified response expression quantitative trait loci (reQTL), where genotype predicts the magnitude of the within-individual response to LPS stimulation. Compared to all tested variants, (r)eQTL SNPs were enriched for variants that show high between-species divergence but low within-species diversity, consistent with the evolution of selectively relevant species differences in the immune response. Together, our results illustrate how natural hybrid zones facilitate the identification of functionally diverged genes and variants important in the early stages of primate divergence.

S17-P04-Evolution of cis-regulatory elements in mammalian peripheral immune cells

Yiheng Li

State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-Sen University, Guangzhou, 510275, China

Gene expression varies considerably across species, and this variation is regulated by transcription factors (TFs) - proteins that bind to specific short DNA sequences known as cis-regulatory elements(CREs). These elements are widespread and relatively conserved across different species' genomes. Comparative genomic studies of non-coding regions have revealed that CREs exhibit significant heterogeneity during species evolution, and their spatiotemporal dynamics can lead to diverse phenotypic outcomes. Therefore, CREs play a crucial role in gene expression variation, and we know that immunity-related genes are major targets of positive selection in mammals and evolve rapidly. This raises the question of whether the evolution of cis-regulatory elements in the immune system provides additional insights.

To explore this, we performed single-cell ATAC-seq (scATAC-seq) on peripheral blood mononuclear cells (PBMCs) from seven mammal species: mouse, rat, rabbit, human, macaque, alpaca, and pig. In addition, we examined two mouse strains, C57BL/6NJ and PWK/PhJ, to investigate the evolution of cis-regulatory elements during a short period of divergence.

We developed a convolutional neural network model capable of classifying cis-regulatory elements identified from scATAC-seq data across these species, enabling accurate determination of their functional conservation. Furthermore, we aim to uncover the evolutionary forces shaping these elements by analyzing and comparing the characteristics of cis-regulatory elements with varying levels of functional conservation.

S17-P05-Patterns of extreme outlier RNA expression in population data reveal sporadic over-activation of genes with co-regulated modules in subsets of individuals

Chen Xie

Peking University

Most RNA-seq datasets harbor genes with extreme expression levels in some samples. Such extreme outliers are usually treated as technical errors and are removed from the data before further statistical analysis. Here we focus on the patterns of such outlier expression. Our study is based on multiple datasets, including outbred and inbred mice, humans from the GTEx dataset, different Drosophila species, and single-nuclei sequencing experiments from human brain tissues. All show comparable general patterns of outlier expression. Different individuals can harbor very different numbers of outliers, with some individuals showing extreme numbers in only one out of several organs of the respective individual. A three-generation family analysis in mice was generated and analyzed for the inheritance of outlier patterns. We find that most extreme over-expressions are not inherited, i.e., appear to be sporadically generated in individuals. Still, co-regulated outlier expressions are identifiable in various gene groups, and some correspond with known pathways. Among the co-regulated genes with extreme outlier expression are also the hormone genes prolactin and growth hormone, both in mice and humans, for which we include also protein level data from human cohorts. We show that outlier patterns of gene expression are a biological reality occurring universally across tissues and species. Most of the outlier expressions are spontaneous and not inherited. We discuss the possibility that the outlier patterns reflect edge of chaos effects that are expected for systems of non-linear interactions and feedback loops, such as gene regulatory networks.

S17-P06-Exploring mRNA translation regulation by the Kozak sequence in the 5' UTR

Dake Gao

Peking University, School of Life Sciences

Translation is a crucial step in protein synthesis, and the Kozak consensus sequence, an essential regulatory element within the 5’ untranslated region (5’ UTR) of mRNA, was first identified by Marilyn Kozak. Positioned near the translation initiation site, the Kozak sequence plays a pivotal role in influencing the translation efficiency of downstream coding sequence (CDS). Despite its importance, it remains unclear whether different Kozal sequences exhibit significant variations in translation efficiency and how these variations manifest across tissues and species. In this study, we systematically investigated the regulatory role of various Kozak sequences in translation efficiency using the parallel reporter assay. Our results demonstrated that the most frequently used Kozak sequences may reduce protein yield and revealed notable differences in translation efficiency across tissues and species. These observations were further validated through dual-luciferase experiments. This study highlights the critical role of Kozak sequences in modulating translation efficiency and presents an optimized assay that provides a strategic framework for enhancing protein production under specific conditions. Intriguingly, we found that most transcripts preferentially utilize Kozak sequences with lower translation efficiency, potentially reflecting an evolutionary balance s between survival and productivity.

S17-P07-Promoter-centric gene regulation enables extreme genome scrambling in Oikopleura dioica

Johannes Nicolaus Wibisana

Genomics and Regulatory Systems Unit, Okinawa Institute of Science and Technology, Okinawa, Japan

Oikopleura dioica is a species of pelagic tunicate that is present in all oceans in the world. It is used as a model for developmental studies due to its fast life cycle and ease of culture. We have shown that different cryptic species of O. dioica obtained from different regions of the world have highly scrambled genome, despite having phenotypically indistinguishable traits. Furthermore, gene order is generally not conserved across different species. While the genome is highly scrambled, we have found that gene expression of orthologous genes is for the most part, conserved. Nevertheless, how this is possible is still unknown.

In this research, through experimental and computational approaches, we aimed to describe the genomic and epigenomic landscape of O. dioica, to investigate the reasons for how gene expression can still be conserved despite their extremely scrambled genomes. We found that O. dioica has a very minimal number of enhancers, which might concentrate gene regulation in promoters. This is corroborated with the finding that the promoters tend to be conserved, and that genomic breakpoints between species occur upstream of promoters, suggesting that promoters move together with genes throughout speciation. Therefore, in O. dioica, cryptic speciation and extreme genome scrambling could have been enabled by the conservation of regulatory elements very closely upstream from its genes.

S17-P08-Genetic and Epigenetic Interplay in Salt Adaptation

Xin Mi

Sun Yat-sen University

Epigenetic modifications are known to play a crucial role in environmental adaptation, but their relative contribution compared to genetic variation remains underexplored. To address this question, we used a reciprocal transplant approach to investigates the root transcriptomes and epigenomes of mangrove Kandelia obovata adapted to high and low salinity in coastal marsh. While genetic transcriptomic changes predominated during the adaptation to saline environments, epigenetic modifications, particularly H3K27me3, showed significant plasticity. Additionally, genetic changes often reversed the effects of phenotypic plasticity, with plastic adaptation occurring first in the process of salt adaptation. These findings highlight the interplay between genetic and epigenetic factors in driving K. obovata ' s salt adaptation and underscore the importance of preserving genetic diversity in mangrove species to maintain resilience to environmental stress.

S17-P09-Personalized transcriptional network analysis links age-related loss of gene coordination to individual biological aging

Haotian Wang

Kunming Institute of Zoology, CAS

Background: Aging is characterized by the decline in biological functions, accompanied by changes in gene-to-gene transcriptional coordination, which can be estimated by regulatory interactions in gene transcriptional network. Notably, gene networks and interactions showed interindividual variability, while personalized aging-related gene interaction dynamics in human cohorts have yet to be investigated.

Methods: In this study, we constructed 15,933 personalized transcriptional networks across 26 tissues from 967 donors aged 20 to 80 years old, using the sample-specific network (SSN) framework based on data from Gene-Tissue Expression (GTEx) project.

Results: We identified gene-gene interactions and characterized their age-dependent dynamic trends across tissues, observing a universal trend of increased gene-to-gene interaction loss during aging across tissues. The count of lost interactions is also positively correlated with individual-level aging and senescence-related molecular phenotypes. Notably, we revealed that the lost interactions have potential as biomarkers for individual aging and health status. In addition, we identified gene interaction loss events exhibiting significant positive correlation with age, defined as Aging-Related Lost Interactions (ARLIs), which may be functionally associated with pathways related to proteolytic processes. Finally, we showed that ARLIs may contribute to deleterious effects and increased pathogenicity through gene dosage imbalances.

Conclusions: This study establishes, for the first time, a connection between the loss of gene-to-gene expression coordination and individual-level aging progress. It provides proof-of-principle evidence for using lost gene interaction as biomarkers of healthy aging and highlights the potential risks associated with interaction loss in specific biological pathways during aging.

S17-P10-Identifying osteoarthritis-related genotype-by-environment interactions using cell culture systems

Shuanghui Chen

Max Planck Institute for Evolutionary Anthropology

Complex traits are influenced by the interactions between genetic and environmental factors, but characterizing these interactions can be challenging. We aim to clarify the impact of genotype-by-environment interactions on gene regulation in the context of one complex trait – osteoarthritis (OA), which is a common degenerative joint disease in humans. Here, we investigate OA-related genotype-by-environment interactions using a population-scale cell culture system. We differentiated 32 human induced pluripotent stem cell lines into chondrocytes and exposed these cells to either mechanical stress (cyclic tensile strain, CTS) or inflammatory cytokine (IL-1 β) treatments to induce OA-like phenotypes. We measured gene expression in each treated condition, as compared to a control condition, and identified both shared and treatment-specific regulatory responses. Differentially expressed (DE) genes that are shared across treatments (e.g., COL27A1, COL11A1) are enriched for skeletal development and extracellular matrix organization functions and overlap with other publicly-available OA-related DE datasets. On the other hand, cytokine-specific (e.g., CXCL8) and CTS-specific (e.g., NQO1) DE genes are enriched for inflammatory response and cell redox homeostasis functions, respectively. Incorporating genotypic information reveals several hundred genes with cis expression quantitative trait loci (eQTLs) across conditions and a subset of genes with context-specific eQTLs. Continued exploration of these static and context-specific eQTLs will improve our understanding of OA-relevant genotype-by-environment interactions. Overall, our cell culture system may be a useful model for studying the complex interactions underlying OA pathogenesis and for more broadly clarifying functional genomics landscapes relevant to this complex trait.

S17-P11-The differential alternative splicing pattern in caste of social insects

Yuhang Jia

Tohoku University

Caste differentiation in social insects represents a critical model for understanding phenotypic plasticity through gene regulation. While differential transcription has been extensively documented in caste development, alternative splicing (AS) - a key mechanism generating transcript diversity - remains underexplored at the community level across social insect lineages. To better understand the mechanisms of gene splicing regulation between castes in social insects, we conducted the following analyses: Firstly, we conducted comparative AS analysis using brain RNA-seq data from representative bees, ants, and wasps with varying social complexity. Subsequently, rMATS was used to quantify alternative splicing and genes with caste-differentiated alternative splicing were identified. Lastly, based on the quantification of alternative splicing, PCA and statistical tests were performed to determine whether there are significant overall differences in splicing patterns between castes. This study has demonstrated that alternative splicing exhibits differences between the castes in social insects, and that these differences are related to individual reproductive phenotypic variations. Also, there is a difference of alternative splicing level between eusocial insects and primitively social insects. We will discuss whether alternative splicing affects the evolution of castes in social insects in our presentation.

S17-P12-snRNA maturation dynamics and the role of the Integrator protein in this process

Anastasiia Bolikhova Skolkovo

Institute of Science and Technology, Skoltech

RNA is a huge class of molecules that perform a variety of functions in a living cell. Each type of RNA must undergo a specific sequence of maturation steps to function. In this work, we study one of the main stages of spliceosomal small nuclear RNAs (snRNAs) maturation. These molecules are synthesized as longer precursors, after which the 3′ tails of snRNAs are gradually cut off, first in the nucleus and then in the cytoplasm. However, the exact length of snRNA precursors at each of these stages is poorly understood. The Integrator protein plays crucial role in the nuclear maturation of snRNA.

The aim of this study is to determine the 3′-end maturation dynamics of various snRNAs. Using the method of isolating newly synthesized RNA via 5-Ethynyluridine followed by high-throughput sequencing, we suggested the precursor length of each snRNA and calculated the rate of 3′-end shortening. The effect of the knockdown of the genes encoding the Integrator protein subunits on the snRNA maturation dynamics was also investigated. snRNAs have many copies in the genome, only some of which serve as a template for the synthesis of functional molecules. By focusing on the most expressed copies, we showed that 3′ region transformation dynamics differ not only between snRNA types, but also between different copies of the same type.

Differences in the masturation of the 3′-ends of snRNAs expands the understanding of the dynamic regulation of splicing.

The work was carried out with the financial support of the Russian Science Foundation #21-64-00006P.

S17-P13-Unravelling the evolution of intrinsic terminators

Stepan Denisov

The University of Manchester

Transcriptional termination allows isolation of operons, enabling tight regulation of gene expression in bacteria. An intrinsic terminator (IT) is an RNA structure that folds on nascent mRNA to terminate transcription. The key functional characteristic of an IT is its efficiency, defined as the proportion of terminated, as opposed to read-through, transcripts.

Here, we adopted a multi-level approach that accounts for IT sequence, structure, surrounding gene context, and function (termination efficiency) to investigate the evolution of ITs between the genomes of Escherichia coli and Salmonella enterica.

The majority of ITs lack sequence homologs between the two genomes, in contrast to coding regions. While most ITs with a homolog have only one between the two species, we identified one large and several medium-sized sequence clusters of ITs. The large cluster overlaps with the mobile element IS186, which propagated both before and after the divergence of E. coli and S. enterica.

The presence of an IT between two homologous operons is more conserved than the IT sequence itself, meaning that the same operons often possess ITs that are unrelated by sequence. The link between an IT and its upstream operon is more conserved than its link with the downstream operon. Both sequence-related and context-only related ITs exhibit structural similarities; however, structures are significantly more conserved among sequence homologs, and only the latter have conserved function.

Thus, we demonstrate how a multi-level approach is necessary to gain insights into the evolution of short and rapidly-changing sequences, such as intrinsic terminators.

S17-P14-Mapping Functional Impact of Alternative Splicing Coupled Nonsense-Mediated Decay

Dylan Stermer

University of Chicago

Alternative splicing (AS) is widely studied for its role in increasing protein diversity. By comparison, its regulatory role is much less recognized. Recently, we demonstrated that comparing AS between nuclear and cytoplasmic mRNA enables the mapping and quantification of splicing events that influence steady-state mRNA expression. In human cell lines, we estimated that 15-20% of all mRNA transcripts are rapidly degraded by nonsense-mediated decay (NMD), suggesting that the regulatory impact of AS-NMD coupling has been significantly underestimated. However, it remains unclear to what extent AS-NMD events arise from stochastic splicing errors versus functionally regulated processes. To address this, we profiled nuclear and cytoplasmic mRNA across multiple vertebrate species and an early-branching metazoan (Nematostella). We developed a method to identify NMD-inducing AS events and leveraged publicly available RNA-seq data to assess their tissue-specific patterns. In total, we mapped 1,818 high-confidence NMD-inducing AS events, which were strongly enriched (>4-fold) in nuclear compared to cytoplasmic mRNA. Tracking the conservation of these events revealed that 10-15% exhibit strong sequence conservation at their splice sites, while 18% show significant differential usage across tissues (ΔPSI > 10%). We identified deeply conserved examples, persisting through Nematostella, indicative of essential regulatory roles. However, the majority of AS-NMD events identified lack both conservation and tissue specificity. In conclusion, our findings reveal hundreds of functionally relevant AS-NMD events while highlighting that most may arise from stochastic splicing.

Symposium 19: Thematic symposium: conservation genomics

S19-P01-Phased genome assemblies reveal haplotype-specific genetic load in the critically endangered Chinese Bahaba (Teleostei, Sciaenidae)

Rongfeng Cui

Sun Yat-sen University

While haplotype-specific genetic load shapes the evolutionary trajectory of natural and captive populations, mixed-haplotype assembly and genotyping hindered its characterization in diploids. Herein we produced two phased genome assemblies of the critically endangered fish Chinese Bahaba (Bahaba taipingensis, Sciaenidae, Teleostei) and resequenced 20 whole genomes to quantify population genetic load at a haplotype level. We identified frame-shifting variants as the most deleterious type, followed by mutations in the 5'-UTR, 3'-UTR and missense mutations at conserved amino acids. Phased haplotypes revealed gene deletions and high-impact deleterious variants. We estimated ~1.12% of genes missing or interrupted per haplotype, with a significant overlap of disrupted genes (30.35%) between haplotype sets. Relative proportions of deleterious variant categories differed significantly between haplotypes. Simulations suggested that purifying selection struggled to purge slightly deleterious genetic load in captive breeding compared to genotyping interventions, and that higher inter-haplotypic variance of genetic load predicted more efficient purging by artificial selection. Combining the knowledge of haplotype-resolved genetic load with predictive modelling will be immensely useful for understanding the evolution of deleterious variants and guiding conservation planning.

S19-P02-Spatial variation in Toll-like receptor diversity in koala populations across their geographic distribution

Jian Cui

USYD

Habitat loss, disease, and road deaths have driven the koala (Phascolarctos cinereus), an iconic Australian species, to be listed as endangered in the northern regions of its range. Among the key threats to koala populations is disease, primarily chlamydia and koala retrovirus. The distribution of these diseases across the species' range is uneven. Toll-like receptors (TLRs) are critical components of the innate immune system, responsible for recognising and responding to various pathogens. Genetic variations in TLR genes can impact an individual's susceptibility or resistance to infectious diseases. Advances in genomic technology and bioinformatics now allow us to explore this variation between populations. In this study, we analysed a dataset of 413 re-sequenced koala genomes to investigate TLR diversity across the east coast of Australia. Previous research identified five primary genetic clusters, with genetic diversity decreasing along a gradient from northern Queensland to Victoria. We identified 45 single nucleotide polymorphisms (SNP) resulting in 51 alleles across ten TLR genes. Our findings reveal that koala TLR diversity forms four distinct genetic groups, aligning with patterns observed in the koala's major histocompatibility complex (MHC), another key immune gene family. The bioinformatics approach presented in this study has broad applicability to other threatened species with existing genomic resources.

S19-P03-Linking chromosomal fusions to adaptive evolution in an invasive copepod species complex

Zhenyong Du

University of Wisconsin–Madison

Genome architecture evolution is a rapidly advancing field in evolutionary biology and conservation genomics, driven by the influx of genomic data revealing unexpected patterns and mechanisms of adaptation. Chromosomal rearrangements are increasingly recognized as contributors to evolutionary adaptation, yet direct empirical evidence linking chromosomal evolution with selection response remains scarce. This study investigates the role of chromosomal fusions in adaptive evolution within the copepod Eurytemora affinis species complex (Crustacea), where certain clades repeatedly give rise to invasive populations that cross habitat boundaries. High-quality chromosome-level genome assemblies from three clades revealed remarkable karyotypic evolution driven by up to eleven independent chromosomal fusion events within this species complex. The Atlantic clade (E. carolleeae), the most invasive, exhibited the fewest chromosomes (four) due to numerous fusion events. Comparative analyses between saline and freshwater populations of E. carolleeae uncovered significant enrichment of selection signatures at the chromosomal fusion sites, particularly at the centromeres, coinciding with regions of reduced recombination. At these fusion sites, collections of ion transport-related genes form potential “ supergenes. ” These particular sets of ion transporter genes repeatedly show population genomic signatures of selection in both wild populations and laboratory selection lines during salinity transitions. This work provides novel empirical support for the hypothesis that chromosomal fusions might facilitate adaptation by clustering coadapted alleles, emphasizing the potential role of chromosomal variation in driving evolutionary novelty and speciation. This study offers insights into how genome architecture evolution influences species invasiveness and resilience, with implications for understanding and conserving genetic diversity in changing environments.

S19-P04-Identifying and understanding new meiobenthos from the Deep Pacific: making up a taxonomic shortfall

James Fleming

JAMSTEC

The oceans represent 71% of the Earth's surface, the largest biome on the planet. At the same time, estimates suggest that as many as 80% of marine species may yet be unknown, or undescribed. One area where this shortfall is particularly noticeable is in ocean meiobenthos, organisms between 0.045 and 1mm. Here, difficulties in sampling, sample retention and identification of cryptic morphology have posed great challenges to researchers in the past, in particular making it difficult to directly understand the key roles that these small organisms likely play in seafloor ecosystems.

The need for taxonomic identification and species discovery in these environments presents a challenge to conservation biology that genomics might help solve. We cannot protect and conserve what we do not know or understand. How can we understand how habitats might change if we are unsure of the roles played by those living there?

Directly studying the seafloor in situ is challenging, considering the size of these organisms, but genomics might present a way forward to understanding the processes that are important to survive at 1mm in oceanic sediment. Annotation and identification of ecologically relevant pathways, and comparative genomics with closely related coastal species may help elucidate how these organisms are distinct, and help us understand the role they play in life at depth. We present an ongoing workflow for the formal extraction, identification and sequencing of previously challenging marine taxa, with the first high-quality genome of Kinorhyncha alongside our previous genome data of related taxa.

S19-P05-Whole-Genome Resequencing Reveals the Genetic Health of Reintroduced California Red-Legged Frog (Rana draytonii) Populations in Southern California

Hanyu Yang

University of California, Los Angeles (UCLA), Department of Ecology and Evolutionary Biology

The California red-legged frog (Rana draytonii) is federally listed as threatened and state-listed as a California Species of Special Concern. Southern California populations have experienced pronounced declines, including the near extirpation of the species from Ventura, Los Angeles, Riverside, and San Diego Counties. As a result, reintroduction programs are underway to reintroduce the species from the Simi Hills in Ventura/Los Angeles Counties, CA, to streams that they historically occupied in the nearby Santa Monica Mountains. Demographically, these reintroductions seem to be succeeding. However, these populations have never been evaluated for their genetic health using modern genomic tools, in part because of their large, approximately 10 GB genome size. Therefore, the extent of deleterious variation and the frequency and impact of inbreeding remain unknown aspects of this large reintroduction initiative. Previous studies using reduced marker sets have demonstrated a genetic bottleneck in the Simi Hills reintroduction source population. Our goal is to use whole-genome resequencing to explore the genomic consequences of relying solely on a single bottlenecked population for reintroduction. This study utilizes our newly generated whole-genome resequencing data to evaluate the levels of inbreeding based on runs of homozygosity, genome-wide heterozygosity, and genetic load within the source population, the established reintroduction sites, and several of the remaining Southern California populations. The findings from this genomic study will help determine whether current reintroduction strategies are effective and determine whether alternative approaches, such as sourcing genetic material from other, more distant populations, may better support the long-term survival of this species.

S19-P06-Species-specific molecular responses of subtropical corals to natural bleaching events: Insights from coral microbiome, metagenome, and metatranscriptome analyses

Jack Chi-Ho Ip

Lingnan University

Climate change-induced marine heatwaves pose significant threats to corals through increased frequency and severity of bleaching events. To effectively conserve coral reef ecosystems, it is crucial to understand the variations in thermal susceptibility among coral species. Our previous studies highlighted the diverse responses of bacterial and zooxanthellae communities, as well as holobiont transcriptomic changes, in three coral species — Montipora peltiformis, Pavona decussata, and Platygyra carnosa—during the natural bleaching event of 2017. However, variations in sampling sites and times limited our capacity to conduct a comprehensive comparison of the molecular mechanisms contributing to species-specific bleaching resilience. To address these knowledge gaps, we investigated the microbiome and metagenomics of five scleractinian corals — Acropora solitaryensis, Duncanopsammia peltata, M. peltiformis, P. decussata, and P. carnosa — using samples collected during the 2022 bleaching event at Bluff Island, Hong Kong. By assessing coral physiological parameters and characterizing bacterial and zooxanthellae communities, this study aims to elucidate the roles of microbiome communities in coral survival and recovery from bleaching. Additionally, we conducted functional analyses of metagenomes and metatranscriptomes to identify the adaptive molecular responses of both coral hosts and symbionts during recovery with different bleaching susceptibility. Our findings provide insights into the molecular mechanisms underlying holobiont resilience, which may determine the fate of subtropical coral species in response to global and regional environmental changes.

S19-P07-Genetic diversity loss is functionally irreversible by natural mutation alone

Toby Kovacs

The University of Sydney

Maintaining genetic diversity is essential for species conservation, as it preserves the ability of species to adapt to novel environmental pressures. Unfortunately, widespread genetic diversity loss has been reported across the tree of life. This valuable genetic diversity can potentially recover, even after a bottleneck event, but this recovery depends on the accumulation of new mutations over generations. However, the exact timescale for recovery has remained unclear because of the limited characterization of mutation rates in non-model species. Using recent estimates of mutation rates across the tree of life, we modelled the recovery of genetic diversity following a population bottleneck. Our simulations and theoretical models demonstrate that mutation rates have little impact on genetic recovery rates in threatened species, which are instead determined primarily by population size. For example, an isolated population with an effective size of 500 that experiences a 50% reduction in heterozygosity would require approximately 1,600 generations to restore heterozygosity to 90% of pre-disturbance levels. Based on these findings, we propose that a 50% loss of heterozygosity in a threatened species equates to a loss of at least 1,600 generations of evolutionary history. Our results highlight that the loss of genetic diversity can be effectively irreversible within the timeframe of conservation management programs, which underscores the importance of preventing genetic bottlenecks to preserve biodiversity and adaptive capacity.

S19-P08-Increased genetic load in purebred chickens and purging of coding sequences in long ROHs

Ruoshi Huang

Westlake University

Domestication involves series of bottlenecks that reduce the effective population size (Ne) and genetic diversity, and artificial selection for desired traits can elevate the frequencies of harmful mutations linked to favored genotypes. Domestic chicken, being the most abundant birds on Earth, have experienced multiple rounds of selection events and population bottlenecks that radiated thousands of modern breeds around the world, making them an excellent group to study how different demographic events influence the distribution of deletious mutations. Our study used representative domestic chicken populations and their wild progenitor Red Jungle Fowl (RJF) to investigate the effect of varied strengths of artificial selection and inbreeding on the accumulation of genetic load. We found that the initial domestication exerted minimal impact on chicken genetic load. However, the much more intense breeding practices reduced Ne of commercial breeds and the efficiency of natural selection in removing deleterious mutations from genomes. The realized load of both coding and non-coding variants is significantly increased in the heavily selected breeds compared to village chickens and RJF populations, mainly contributed by inbreedng. Additionally, recent inbreeding events have led to the purging of coding sequences while enriching non-coding deleterious variants in long runs of homozygosity, particularly in commercial breeds. We also found that the relative allele frequencies of deleterious mutations are drastically elevated through breed-specific artificial selection. Our results highlight the importance of both coding and non-coding regions when assessing genetic load and provide new insights for detecting signals of purging in ROHs.

S19-P09-Chromosome-scale genome assembly in the orchid Epidendrum fulgens reveals insights for environmental adaptation

Fabio Pinheiro

University of Campinas

Widely distributed species can span various habitats and local environmental conditions, making them great study models for adaptation and selection in the wild, although the specific targets of selection have not been assessed for many species. To better understand these diversifying patterns on a microevolutionary scale, we can assess adaptive patterns through the lens of genomics, studying how the genomic variation among populations is shaped and how selection is acting on a whole genome level. To investigate adaptation through a latitudinal environmental gradient in southern Brazil, we sequenced, assembled, and analyzed the first genome of a neotropical orchid: the Epidendrum fulgens, a perennial species that inhabits sand dunes in the highly deforested Brazilian Atlantic Forest. We used this newly-assembled genome as a reference to call variants from RNA-seq data, with information from 80 individuals collected from 10 populations along the species geographical range. We analyzed 75,973 single-nucleotide polymorphisms (SNPs) and performed population genomics and differential gene expression analyses. Our final primary assembly comprised a total of 12 chromosome-scale scaffolds, containing 97.3% of plant BUSCO genes, with an N50 of 88.6 Mb and L50 of 5 scaffolds. Similar assembly values were found also for both phased haplotypes. We detected three genetic clusters following the latitudinal and environmental gradient of temperature and precipitation. We also found 447 SNPs potentially involved in climate adaptation. Our results provide insights into the adaptive processes shaping the E. fulgens' distribution across a latitudinal environmental gradient in a deforested and extremely biodiverse neotropical biome.

Funding: FAPESP (2020/02150-3).

S19-P10-Understanding bonobo evolution and diversity from genomic data

Sojung Han

University of Vienna

Climate change and species extinction rates are accelerating, making conservation a bigger issue than ever. Since evolutionary research requires biological material, there is a dilemma: it is important to understand the evolutionary history of the species, in order to help the species conservation based on what the populations experienced. At the same time, it is very difficult to get biological materials due to multiple obstacles. Bonobos are a very good example for this situation: an endangered great ape species rare both in captivity and in the wild. Due to political unrest and international regulations, biological materials for high-coverage whole genomes of bonobos are very limited.Here, we present results obtained from re-analysis of previously published data sets, including 20 exomes and mtDNA from wild-born bonobos, 14 whole-genomes of bonobos, and partial mtDNA of 136 wild individuals. Data from mitochondria previously suggested population substructure between bonobo groups on the eastern, central and western part of their geographical distribution. Using exome and whole-genome data, we identify genetically distinct bonobo groups with associated geographic origins. Our estimates indicate that the split time within bonobos is at least ~145,000 years ago, which is comparable to that between eastern and central chimpanzees. The estimated long-term effective population sizes for the far-west population is among the lowest among great ape lineages, strengthening the need for preserving these populations.Our work highlights the value of reusing existing data for exploring unanswered questions, with implications for conservation efforts, such as awareness of local population structure for reintroduction

S19-P11-Insights from Australian government conservation genomics projects

Eilish McMaster

University of Sydney

For the past five years, the New South Wales (NSW) state government has applied conservation genomics to support the protection of endangered plant species. To date, over 100 species have been assessed. Although we had initial expectations regarding the outcomes of genomic studies on endangered species, our findings have revealed diverse and often unexpected patterns across the species studied, with significant implications for their management. This poster will present the key elements of our workflow, including costs, analyses, and outcomes, along with specific examples from our work. We advocate for the global adoption of this accessible genomic method to inform and enhance conservation management practices worldwide.

S19-P12-Mixed-Ploidy Evolution and Speciation in Cyclocarya (Juglandaceae): Taxonomy, Adaptation, and Polyploidy Origins

Yan-Feng Song

Beijing Normal University

Polyploidy, arising from whole-genome duplication (WGD), is a key mechanism driving speciation and evolutionary diversification in eukaryotic species. A compelling example of this phenomenon is found in Cyclocarya Iljinsk. (Juglandaceae), a genus widely distributed across the subtropical regions of East Asia. It naturally exists as both diploid and autotetraploid cytotypes, offering a unique and valuable germplasm resource. Traditionally regarded as a monotypic genus, this study —through ploidy identification, morphometric analysis, and population genomics of 179 individuals from 34 populations — uncovers a complex evolutionary history within Cyclocarya, revealing the presence of distinct taxa. The diploid group in central China is identified as the basal lineage, having diverged in the late Pliocene. It exhibits significant morphometric differences compared to individuals from other populations, supporting its re-instatement. In contrast, the diploid group from southeastern Yunnan clusters with autotetraploid individuals, suggesting closer phylogenetic relationships. This finding implies that the autotetraploid originated from a single polyploidization event, likely driven by Quaternary climate fluctuations, with the diploid ancestor undergoing polyploidization. Positive selection analysis reveals that genes involved in hydrogen peroxide metabolism in autotetraploids are under positive selection, potentially providing an adaptive advantage for their widespread distribution. This study clarifies the taxonomic status and evolutionary history of Cyclocarya species through comparisons of morphometric traits, population genetic analysis, and phylogenetic reconstruction across different ploidy groups. It also offers new insights into the origin and expansion of polyploidy and its adaptive basis, presenting a valuable model for the study of complex mixed-ploidy taxa.

S19-P13-Epistatic regulation of an adaptive phenotype increases population resilience to fluctuating environments

Genis Garcia-Erill

Bioinfromatics Research Center, Department of Molecular Biology and Genetics, Aarhus University

Populations are locally adapted to their environment that can vary across space and across time. Understanding the genetic architecture of phenotypic variation that enable natural populations to track temporal changes of the environment is crucial in the context of the ongoing climate change. One successful example of ongoing climate adaptation is that of chemotype diversity in common thyme (Thymus vulgaris). Thyme harbors genetically determined individual chemotypes varying 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 adapted to different local habitats. In recent years, ecotypes are shifting local frequency as they adaptively respond to ongoing climate change. To understand the genomic architecture of these phenotypes, we sequenced 1000 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 associate genetic and phenotypic diversity to environmental variation across local and regional spatial scales. Using a newly assembled genome, we mapped genetic variants determining chemotypes across populations. These include copy number variation at genes encoding terpene synthases in the biosynthesis of the monoterpenes that define the chemotypes. We characterize epistatic interactions across these loci and how this results in the maintenance of the two adaptive ecotypes. Finally, we use population genetic forward simulations to explore if and why the inferred epistatic and oligogenic genetic system provides an advantage over alternative genetic architectures for tracking environmental variation and climate change.

S19-P14-Determinants of local adaptations in arthropods in fragmented Denmark grasslands

Jilong Ma

Aarhus University

Arthropod populations are declining worldwide, with habitat loss from human land-use changes and rapid environmental shifts exacerbating the threat. Local extinctions may occur when populations either fail to adapt to these changes or suffer critical size declines due to ecological constraints or direct human impact. Survival under such pressures may depend largely on the capacity for local adaptation; however, the spectrum of adaptive abilities among arthropods and their underlying determinants remain poorly understood. In Denmark, intensive agricultural practices have fragmented natural landscapes, providing an ideal setting to explore these dynamics. We collected and pool-sequenced populations from 30 arthropod species—including representatives of 10 collembola, 10 beetles, and 10 spiders—across 55 grassland sites with detailed local environmental data. This species assemblage, which spans a range of dispersal abilities and effective population sizes, experiences comparable levels of habitat fragmentation and environmental change. By linking local adaptation signals with population genomic metrics — such as effective population size trajectories, gene flow, and genetic diversity—our study integrates ecological and genomic perspectives to elucidate factors driving population resilience. Moreover, by quantifying the degree of local adaptation, we assess its role in buffering against extinction risk. Although preliminary, our approach promises to enhance our understanding of the adaptive mechanisms that enable arthropods to persist in rapidly changing, human-altered landscapes.

S19-P15-High quality genome assembly of the Beluga Whale: unraveling Arctic adaptation, longevity, and conservation challenges

ALINA DEMKINA

Biotech Campus LLS, Moscow, Russia

The beluga whale (Delphinapterus leucas) is a unique Arctic and sub-Arctic cetacean species known for its adaptability to extreme cold environments. Although globally classified as "Least Concern" with an estimated population of over 200,000 individuals, belugas play a critical ecological role in the Russian Arctic, where they are integral to the marine food web and culturally significant to indigenous communities. Understanding the genomic underpinnings of this species is crucial for conservation genomics, particularly given the extreme difficulty of studying beluga whales in the wild and the challenges of collecting biological samples.

In this study, we present a high-quality genome assembly of the beluga whale, using a combination of MGI deep short-read sequencing, long and ultra-long ONT reads. Our assembly achieves the highest quality among those currently available on the NCBI Assembly database. We also performed genome annotation and phylogenetic analysis, revealing insights into the evolutionary relationships of belugas and their unique adaptations to the harsh Arctic environment.

Furthermore, we have sequenced the genome of the Baird’s beaked whale (Berardius bairdii) and plan to expand our research by sequencing other long-lived cetaceans, such as the fin whale (Balaenoptera physalus). These efforts aim to facilitate conservation initiatives, explore the genetic factors contributing to their exceptional longevity, and provide insights into the broader evolutionary strategies that enable a long lifespan and survival in extreme Arctic conditions.

Symposium 21: Thematic symposium: virus evolution and pathogen evolution

S21-P01-The Evolutionary Dynamics of Usutu Virus: Emergence, persistence, and spread of a rapidly evolving orthoflavivirus in Europe

James Baxter

Roslin Institute, The University of Edinburgh

Usutu virus (USUV) is an Orthofl avivirus frequently associated with sudden mass mortality events amongst blackbirds (Turdus merula), and birds of prey. Transmitted by Culex spp. mosquitoes, USUV infects a broad host-range, including ‘dead-end’ infection of mammals. Human infection is infrequent and typically asymptomatic but can cause encephalitis and meningoencephalitis. The incidence of USUV in Europe has dramatically increased since first detection in the late 1990s; however, a detailed analysis of the factors driving the increases in incidence remains forthcoming. In this study, we compile and analyse the most comprehensive dataset of USUV genetic sequences to date, combining 1417 publicly available sequences with 237 recently generated sequences shared as part of the Versatile Emerging Infectious Disease Observatory. We leverage these data to unveil the pan-European evolution and dispersal history of key USUV lineages, revealing distinct introductions and persistence pathways that correlate with major flyways in the east-Atlantic and central Mediterranean. Crucially, we implement and further develop methods to incorporate partial sequences alongside our full genomes, allowing us to maximise the spatiotemporal coverage of our analysis. Next, we apply spatially explicit phylogeographic models to determine key drivers of USUV di ff usion. We demonstrate a role for crop density and Passeriformes spp. density in propagating viral dispersal across the continent, while dispersal velocity is lower in high-altitude and urban zones. Finally, we compare the ecological drivers of USUV and West Nile Virus, as we seek to understand why such similar viruses appear to follow inherently diff erent epidemiology across Europe.

S21-P02-Functional Evolution of SARS-CoV-2 Conferred by Non-spike Mutations

Zhenglin Zhu

Chongqing University

Understanding the functional evolution of SARS-CoV-2 and associated mechanism is important both for the control of future pandemic coronavirus-induced disease and the development of viral evolutionary theory. For these, our previous work constructed an online analysis platform of coronavirus on a historical and global scale, uncovered frequent recombination events across coronavirus attributed to different hosts, and predicted fast-increasing mutations just after the pandemic of SARS-CoV-2. Aside from the spike mutations (e.g. D614G), our in-silico surveys predicted that the non-spike mutations (such as N R203K/G204R and NSP4 T492I) are capable of promoting viral infectivity and altering the severity of the disease. Through a combination of comprehensive epidemiological analyses and experiments, we demonstrated that R203K/G204R confer a replication advantage, increase virus fitness and virulence, incur positive selection and associate with the emergence of Alpha (Cell Host & Microbe, 2021), and NSP4 T492I associates with the increased infectivity, transmissibility, and attenuated pathogenicity observed in Delta and Omicron, through altering non-structural protein cleavage, which changes viral properties (Cell Host & Microbe, 2023). These studies highlight the importance of non-spike mutations in the evolutionary dynamics of SARS-CoV-2. Moreover, our recent unpublished work demonstrates that a single non-spike mutation is capable of driving a rapid evolution by elevating mutation rates, impacting deaminases and exerting strong epistasis. These findings we present should contribute significantly both to the theory and research of viral evolution.

S21-P03-Recombination and Selection Trajectory of the Monkeypox Virus During Its Adaptation in the Human Population

Jialu Zheng

Sun Yat-sen University

Monkeypox, caused by the monkeypox virus (MPXV), was historically confined to West and Central Africa but has now spread globally. Recombination and selection play crucial roles in the evolutionary adaptation of MPXV; however, the evolution of MPXV and its relationship with the recent, ground-breaking monkeypox epidemic remains poorly understood. To gain insights into the evolutionary dynamics of MPXV, comprehensive in silico recombination and selection analyses were conducted based on MPXV whole genome sequence data. Three types of recombination were identified: five ancestor-sharing interspecies recombination events, six specific interspecies recombination events and four intraspecies recombination events. The results highlight the prevalent occurrence of recombination in MPXV, with 73.3% occurring in variable regions of the genome. Selection analysis was performed from three dimensions: proteins around recombination regions, proteins from recombinant ancestors and MPXV branches, and whole-genome gene analysis. Results revealed 2 and 7 proteins under positive selection in the first two dimensions, respectively. These proteins are mainly involved in infection immunity, apoptosis regulation and viral virulence. Whole-genome analysis detected 25 genes under positive selection, mainly associated with immune response and viral regulation. Understanding their evolutionary patterns will help predict and prevent cross-species transmission, zoonotic outbreaks and potential human epidemics.

S21-P04-A slight mismatch between a gene′s codon usage and the cellular tRNA supply is beneficial

Feng Chen

Sun Yat-Sen University

Each amino acid except two is encoded by multiple synonymous codons, but at unequal frequencies. Such codon usage bias (CUB) is observable in almost all species, and commonly assumed as the result of natural selection towards an optimal CUB that matches the cellular tRNA supply. Here we hypothesized instead that the optimal CUB of a gene should slightly mismatch the tRNA supply to avoid excessive translational costs, while ensuring adequate functional payoff. By modifying the CUB of a resistance gene expressed in bacteria under antibiotic selection, we demonstrate that a small mismatch with the tRNA supply confers faster bacterial growth than those with minimized or large CUB-tRNA mismatches. Intriguingly, the optimal degree of CUB-tRNA mismatch increases as the resistance gene becomes less important in media with lower antibiotic concentrations, which is explainable by our model as a shift in the balance between the gene's functional payoff and translational cost. Furthermore, genomic analyses in model organisms suggested that the optimal degree of CUB-tRNA mismatch is larger for endogenous genes with lower functional importance and higher mRNA abundance, respectively supporting the impact of functional payoff and translational cost. Finally, we found that mutations increasing or decreasing the CUB-tRNA mismatch of native genes are both predominantly deleterious, such that the CUB-tRNA mismatch is likely selectively maintained rather than minimized to that achievable in the presence of genetic drift and mutational bias. These results challenge the commonly assumed unidirectional selection on CUB and highlight the CUB-modulated balance between functional payoff and translational cost.

S21-P05-HLA class I escape drives the evolution of SARS-CoV-2 in human population

Ekaterina Riumina

Skolkovo Institute of Science and Technology

SARS-CoV-2 evolution is shaped by human adaptive immunity. Contrary to antibody escape (escape from the B-cell response), the role of escape from cytotoxic T cellular (CTL) response is controversial. Here, we study the origin and spread of SARS--CoV--2 variants that allow escape from presentation by the HLA class I alleles and show the presence of positive selection favoring such mutations. We find that ~35% of mutations that are characteristic of the variants of concern or that have reached high (>5%) frequencies in viral populations facilitate CTL escape. The presence of selection is evidenced by two key observations. First, mutations associated with escape from common HLA alleles reach higher frequencies than those that allow escape from rarer HLA alleles. Second, escape mutations reach higher frequencies in those countries where the corresponding HLA alleles are more frequent, indicating that viral escape is driven by the local genetic composition of the human host population. Moreover, mutations with such correlations and no HLA binding disruption effect were experimentally proved to disrupt binding with TCRs.

We show that the observed escape cannot be due to selection only in immunosuppressed individuals, and instead indicates selection over the course of acute SARS-CoV-2 infections.

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. The work was financially supported by the Russian Science Foundation grant 20-15-00395

S21-P06-Molecular Evolution of Protein Sequences and Codon Usage in Monkeypox Viruses

Ke-Jia Shan

Peking University

The monkeypox virus (mpox virus, MPXV) epidemic in 2022 has posed a significant public health risk. Yet, the evolutionary principles of MPXV remain largely unknown. Here, we examined the evolutionary patterns of protein sequences and codon usage in MPXV. We first demonstrated the signal of positive selection in OPG027, specifically in the Clade I lineage of MPXV. Subsequently, we discovered accelerated protein sequence evolution over time in the variants responsible for the 2022 outbreak. Furthermore, we showed strong epistasis between amino acid substitutions located in different genes. The codon adaptation index (CAI) analysis revealed that MPXV genes tended to use more non-preferred codons compared to human genes, and the CAI decreased over time and diverged between clades, with Clade I > IIa and IIb-A > IIb-B. While the decrease in fatality rate among the three groups aligned with the CAI pattern, it remains unclear whether this correlation was coincidental or if the deoptimization of codon usage in MPXV led to a reduction in fatality rates. Using dual-luciferase assays, we further discovered that the deoptimization of codons in the viral sequence might weaken protein expression during virus evolution, indicating that codon usage might play an important role in virus fitness. This study sheds new light on the mechanisms that govern the evolution of MPXV in human populations and highlights the importance of codon usage in virus evolution and provides guidelines for codon optimization in mRNA and DNA vaccine development.

S21-P07-The lifecycle of bacterial species.

Daniel Falush

The Institute currently known as Shanghai Institute of Immunity and Infection

How do bacteria speciate? Many bacterial species have enormous population sizes, disperse readily and recombine frequently. These characteristics should militate against speciation. I provide an overview of how adaptation to specific niches can lead to genome-wide cessation of gene flow and hence speciation. So far, we have established two routes. In the first route, a distinct ecological strategy arises, which involves adaptation at multiple genome regions. This process leads to the evolution of ecospecies as identified in Vibrio parahaemolyticus or Helicobacter pylori. The next stage in this route is that differentiation expands to becomes genome-wide, as seen for example in some Streptococcus species. Finally, there will be selection against gene flow and speciation. A second route is via the evolution of an ecocline, which we propose has happened in Klebsiella pneumoniae. Here, rather than starting with two distinct types, differentiation occurs at a continuous-valued trait, which is influenced by many loci. Diversifying selection can lead to the distribution of trait values becoming bimodal, with the trait values at the two peaks progressively diverging, causing a diminution of gene flow and eventual speciation. The general importance of these processes can be explored based on the genetic structure and diversity levels of well-sequenced bacterial species.

S21-P08-Inference of fitness landscape dynamics in SARS-CoV-2 amino acid sites

Galya Klink

HSE, International laboratory of statistical and computational genomics

It is said that every cloud has a silver lining. SARS-CoV2 pandemic provided us with an unprecedented amount of data and increased motivation to its analysis. Shifts in amino acid propensities in sites of viral proteins that occur during virus evolution with time and phylogenetic space is something that on one hand make virus evolution less predictable but on the other hand can help us to explore virus fitness landscape in several directions that potentially can be important during future epidemics. We show that changes in site-specific amino acid preferences between SARS-CoV2 lineages can be inferred with simple non-parametric test using information that is provided by UShER mutation paths. Using this approach, we evaluate whether the increase in number of mutations on an amino acid on the leaves of the tree can serve as a predictor of its importance in the next time period. We also show that having enough sequence data we are able to distinguish two causes of changes in site-specific fitness landscape that are epistatic r environmental shifts.

S21-P09-Deep Mutational Scanning of Bat-SC2r-CoV ZC45 Receptor Binding Domain Revealing Bat-SC2r-CoV Evolution and Host Shifting

Feiyu Chen

ZJU-UoE Institute, Zhejiang University

Bat SARS-CoV-2 related coronavirus (bat-SC2r-CoV) is a cluster of bat-origin coronavirus that is phylogenetically close to SARS-CoV-2. Studies have identified several mutations on SARS-CoV-2 receptor binding domain (RBD) to individually enhance viral pathogenicity and infectivity. Similar mutational effects have also been found in some bat-SC2r-CoVs that exhibit same strategy of hACE2 usage as SARS-CoV-2. Yet the potential effects of single RBD mutations on host/receptor shifting of bat-SC2r-CoV lineages independent of hACE2-usage is largely overlooked. This study focuses on the ACE2-independent ZC45 from R. pusillus and investigates how single mutations across its RBD would shape viral spike functionalities, leading to potential host shifting to human. Utilizing Deep Mutational Scanning, a library was designed to include all possible amino acid substitutions of ZC45 RBD and consequently displayed via lentivirus to mimic real viral transduction. Potential mutations enabling hACE2 binding and cell entry were identified with the help of next-generation sequencing. Additionally, the ZC45 spike library was introduced into AlphaFold2 to predict mutational effects on spike conformation and alteration of possible binding pockets. Preliminary findings indicate that mutations of amino acid residues critical for ACE2 interaction in SARS-CoV-2 genome effectively diversify ZC45 spike structure from WT prediction. This work contributes to a better understanding of viral evolution and host-shift, with potential implications for zoonosis.

S21-P10-Historical cattle movements have driven global Mycobacterium bovis transmission

Andries van Tonder

University of Cambridge

Mycobacterium bovis is the main causative agent of bovine tuberculosis (bTB), a chronic disease primarily infecting cattle but also capable of infecting a wide range of other mammals including humans and wildlife. Whilst previous work has highlighted the role of historical cattle movements in the global distribution of M. bovis, this has not yet been related to genomic changes using large-scale analyses. We aimed to characterise the global population structure of M. bovis, infer key dates around global transmission and use information about historical global cattle movements to provide context to genomic analyses. Ancestral state reconstruction of two lineages, La1.8.1 (Eu1) and La1.7.1 (Eu2), using sampling location inferred an early transmission of La1.8.1 to the Americas, supported by historical evidence for cattle movements initially from the Canary Islands to Spanish colonies in Central and South America from 1493 onwards, followed by multiple subsequent re-introductions to Europe and onward spread to other parts of the world particularly, British colonies. Later reintroductions to the UK could be explained by exports of live cattle to Britain and Europe from the USA in the late 19th century to supply meat following major European cattle losses to Rinderpest outbreaks. Conversely, the inferred historic transmission of La1.7.1 shows an early emergence in Northern Europe, spread to Southern Europe and subsequent transmission to North and South America during periods of Spanish and Portuguese colonial expansion. The integration of genomics with historical evidence of cattle movements thus allows the global spread of M. bovis to be better understood.

S21-P11-Evolution of marine contagious cancers in cockles characterized by genomic instability

Alicia L Bruzos

Université de Caen Normandie

Transmissible cancers are malignant cell clones that spread among individuals through transfer of living cancer cells. Several such cancers, collectively known as bivalve transmissible neoplasia (BTN), are known to cause leukaemia-like disease in marine molluscs such as the common cockle which inhabits the Atlantic coasts of Europe and Africa. To investigate the origin and evolution of contagious cancers in common cockles, we collected 6,854 specimens and diagnosed 390 cases of BTN. We then generated a reference genome and assessed genomic variation in the genomes of 61 BTN tumours. Tumour-specific variants confirmed the existence of 2 cockle BTN lineages with independent clonal origins, and gene expression patterns supported their status as haemocyte-derived blood cancers. Mitogenomes revealed several mitochondrial capture events in BTN, as well as co-infection of cockles by different tumour lineages. Cytogenetic and copy number analyses uncovered genomes marked by pervasive instability and karyotypic plasticity. Whole-genome duplication, amplification of oncogenes CCND3, MDM2 and MYC, and deletion gene MGMT, are likely drivers of BTN evolution. Characterization of satellite DNA identified elements that are absent from tumours despite vast expansions in the cockle germ line, suggesting ancient BTN clonal origins. Our study illuminates the evolution of transmissible cancers under the sea and reveals indefinite tolerance of extreme instability in neoplastic genomes.

S21-P12-Antigenic Evolution Landscape and Vaccine Composition Recommendation Strategy for SARS-CoV-2

Jingru Feng

Sun Yat-Sen University

Background: The antigenic properties of SARS-CoV-2 have changed multiple times during the pandemic, resulting in a decrease in vaccine efficacy. To address this challenge, we rapidly developed a predictive model to assess antigenic similarity, mapped the antigenic evolution landscape, and proposed a strategy for optimized vaccine composition.

Methods: We identified the conformational epitopes of spike proteins by evaluating antigen-antibody binding probability and the relative solvent-accessible surface area. By integrating epitopes, physicochemical properties, and receptor-binding features, machine learning models to distinguish antigenically similar and dissimilar strains were developed based on neutralization experiment data. Using the model, an antigenic correlation network (ACNet) was constructed, and Markov clustering algorithm was used to identify antigenic clusters. Finally, a vaccine composition recommendation strategy was proposed according to the prevalence of antigenic clusters.

Findings: Our predictive model achieved a cross-validation accuracy of $0 \cdot 86$, with epitope mutations identified as the most critical feature in the antigenic shift. From the constructed ACNet, comprising thousands of representative strains, seven antigenic clusters were identified. We observed that the antigenic clusters shifted rapidly, occurring roughly six months. Our vaccine composition recommendation strategy was effective in recommending strains that closely matched circulating strains under nearly all conditions.

Interpretation: Traditional antigenic mapping methods are time-consuming and limited in strain scope for characterizing the antigenicity of SARS-CoV-2 and recommending vaccine compositions. Our data-driven approach provided a rapid solution for understanding SARS-CoV-2 antigenic evolution and recommending vaccine compositions. This strategy may contribute to the prevention and control of SARS-CoV-2 epidemics and inform future vaccine development.

S21-P13-Evolutionary dynamics and variation characteristics of SARS-CoV-2

Xiaolu Tang

Peking 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 "S lineage" (U8,782 and C28,144) and the derived "L lineage" (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.

S21-P14-Understanding the role of Anseriformes in avian flu transmission dynamics in China

Ying Zeng

Roslin Institue, UoE

Avian influenza, especially highly pathogenic avian influenza (HPAI), poses severe risks to wild birds and poultry, causing significant economic losses, and can be transmitted to mammals. Wild Anseriformes (e.g., duck, geese, or swan) are natural hosts of avian flu viruses. Their long-distance asymptomatic migration enables virus dispersal, and their interaction with domestic ducks increases the risk of poultry infections. Understanding the role of domestic and wild Anseriformes in virus transmission is helpful for effective surveillance and outbreak prediction.

In China, different avian flu subtypes have different spread and persistent patterns in bird host populations in the past 10 years. For instance, H5N6 and H6N6 were dominant in ducks during 2014 – 2016. From 2016 onwards, H9N2 became the dominant subtype in domestic ducks. While recently circulated panzootic H5N1 viruses are mainly found in wild ducks and geese but sporadically found in domestic ducks. Using simulation and phylodynamic methods, we explored the patterns of evolutionary and transmission dynamics of different avian flu subtypes/lineages to understand the role of Anseriformes in the virus's emergence and amplification. We also integrated ecological, agricultural, and human intervention data to understand potential drivers of the spatial dispersal of viruses in different periods. For those significant drivers identified, we further assessed virus risks using spatial estimation models. In all, our findings map key factors influencing virus spread by Anseriformes in China, contributing to targeted avian flu control strategies.

S21-P15-Geographical variation and selective equilibrium of the P. falciparum sickle-associated mutations

Gavin Band

University of Oxford

The Plasmodium falciparum genome harbours genetic mutations that enable it to overcome the protective effects of sickle haemoglobin (HbS), raising deep questions about the underlying evolutionary biology. Here, we study this interaction from a geographical viewpoint. We first use available survey data to build a global map of sickle haemoglobin frequencies, extending previous work and allowing us to account for estimation uncertainty in our analysis. We then incorporate parasite population genetic data to estimate the relationship between the local frequencies of HbS and the P. falciparum sickle-associated (Pfsa) allleles. We find evidence of a strong positive relationship between HbS and Pfsa+ both across the continent of Africa and at local scales. These results therefore suggest that Pfsa+ alleles are positively selected for by HbS, but a major puzzle is why the alleles have not reached fixation, nor appear to be undergoing selective sweeps in most populations. To elucidate this, we explore population-genetic models combining positive selection by HbS with a weak opposing cost incurred by Pfsa+ parasites in HbAA individuals. We find that the spatial variation in these models is sufficient to generate stable equilibrium patterns qualitatively similar to observed data. The opposing force against Pfsa+ alleles is not known, but could relate to the high observed linkage disequilibrium between loci. These results therefore raise the possibility that the Pfsa+ mutations are evolving under a form of balancing selection driven by geographical variation in HbS frequencies. Functional studies will be needed to determine the biological mechanisms and selective forces involved.

S21-P16-Genomic data from a medieval case of Yersinia ruckeri infection sheds light on host adaptation in Yersinia sp.

Cameron R Ferguson

UCL Genetics Institute, Department of Genetics Evolution & Environment, University College London, UK.

Yersinia ruckeri is a gram-negative, rod-shaped, facultative anaerobe primarily associated with enteric redmouth disease in salmonids. Initially isolated from rainbow trout, Y. ruckeri was believed to have a narrow aquatic host range. However, its detection in a growing number of species, including a range of mammals, suggests a broader host range with potential epizootic and zoonotic implications.To investigate the evolutionary history and determinants of host-range in Y. ruckeri, we analyzed 132 genomes, incorporating publicly available data, an unpublished isolate from an infected Harbour porpoise, and an exceptionally high-coverage (44X) infection from the medieval remains of a human in Trondheim, Norway. The latter provides rare evidence of a past opportunistic infection in an individual likely exposed to marine environments.We reconstructed the Y. ruckeri pan-genome and assessed gene enrichment in mammalian versus aquatic hosts. Our findings reveal minimal evidence of host-specific adaptation for mammalian infection, instead supporting Y. ruckeri's broad host range, with at least three independent spillover events into mammals. Bayesian tip-dating analysis, incorporating our ancient sample, allowed estimation of an extremely low evolutionary rate for Y. ruckeri, placing its most recent common ancestor between 76,280 and 11,132 BCE.Together our work supports an ancient emergence of Y. ruckeri and expands current understanding of its transmission potential between aquatic and mammalian hosts.

S21-P17-Understanding drivers of seasonal influenza variant fitness to improve evolutionary forecasting

Philippa Steinberg

University of Washington, Seattle

Seasonal influenza is a rapidly evolving respiratory virus. The fixation of advantageous mutations in the hemagglutinin (HA) surface protein segment drives the spread of new variants and results in antigenic drift in the viral population. Antigenic variants escape from a fraction of neutralizing antibodies, resulting in frequent reinfection, as seen by annual flu season epidemics. This process of variant emergence and rapid fixation allows for the study of real-time evolution, where detailed data exists on viral genetic makeup, phenotype characteristics as well as the landscape of human population immunity that drives variant fitness. Influenza lineages A/H3N2, A/H1N1pdm, B/Victoria, and B/Yamagata (extinct) exhibit highly variable circulation patterns. The Multinomial Linear Regression (MLR) model has been successfully applied by the Bedford lab to estimate SARS-CoV-2 variant frequencies and relative fitness, but it has yet to be systematically applied to seasonal influenza. My project expands the MLR model to seasonal influenza to improve estimates of variant fitness across time and geography. I estimate seasonal influenza variant fitness changes through time and space and forecast future variant frequencies. Understanding of influenza's long-term antigenic evolution will guide future research on respiratory viruses including avian influenza (e.g., H5N1) and respiratory syncytial virus (RSV).

S21-P18-Diana: A WebAssembly-Powered Tool for TraDIS Data Exploration and Analysis

Wytamma Wirth

The Peter Doherty Institute For Infection And Immunity

Understanding the molecular basis of phenotypes such as persistence in the human gut or antimicrobial resistance requires advanced genetic tools. Transposon-directed insertion-site sequencing (TraDIS) is widely used for functional genomic studies. Here, we present Diana, a WebAssembly-powered tool designed for the exploration and analysis of TraDIS data. We cover the design, implementation and validation of the tool. Diana enables users to safely and rapidly analyse their data, and its integration into our comprehensive platform streamlines TraDIS experiments from start to finish. To demonstrate our platform's performance, we assessed the impact of vancomycin exposure on a library of 48,458 unique transposon mutants, confirming the crucial role of the vanB operon in VREfm vancomycin resistance. Diana thus facilitates accessible, genome-wide functional studies to unravel the molecular mechanisms underlying persistence and pathogenesis. The tool is open-source and freely available at https://diana.cpg.org.au.

S21-P19-Spatial Model of Self-Limiting Genetic Biocontrol Strategies

Zehua Chen

Imperial College London

Malaria transmission in Africa is intricately linked to mosquito population dynamics. This study utilises a two-locus, two-stage, two-sex model, formulated through a system of partial differential equations (PDEs), to simulate the population dynamics of mosquitoes. We will analyse multiple self-limiting genetic control strategies such as Sterile Male Technique, (Fs-)RIDL, Split/Dominant (female) sterile homing, X-Shredder, and Y-Linked Editor (YLE), assessing their effectiveness in suppressing mosquito populations within targeted areas. By integrating the spatial distribution of different sexes and genotypes of mosquitoes across their life stages, this model will predict and compare the outcomes of each genetic alteration strategy. The simulations focus on key metrics such as suppression efficacy and localisation.

Symposium 22: Thematic symposium: evolution theory and recent methodological advances

S22-P01-Feasibility and Limitations of Using Deep Learning to Detect Mass Extinctions on Fossilized Phylogenetic Trees

Minghao Du

Central South University

Inferring diversification rates on phylogenetic trees can reveal patterns in biological evolution. Throughout geological history, multiple mass extinction events have significantly influenced biodiversity, posing challenges for accurately inferring diversification rates. While extensions of (fossilized) birth-death process have incorporated mass extinctions in Bayesian phylogenetic inference, precisely determining the timing and magnitude of these events remains challenging. This limitation likely arises from insufficient signal preserved within reconstructed phylogenetic trees. Recently, deep learning methods have been employed as efficient alternatives to likelihood-based approaches for inferring diversification rates.

In this study, we simulated phylogenetic trees with time-varying diversification rates, fossil sampling rates, and mass extinction events. We trained a neural network model that integrates graph attention neural networks and long short-term memory architectures, in order to estimate the presence, timing and magnitude of mass extinctions. Our results demonstrate that deep learning-based methods can detect mass extinctions with accuracy comparable to likelihood-based methods. Using the trained model, we examined how detection accuracy is affected by tree size, mass extinction timing and magnitude, as well as other factors. Additionally, we assessed the robustness of the deep learning model by introducing slight perturbations to tree topology and branch lengths. Finally, we tested the model on phylogenetic trees outside the training range to establish the limitations of the deep learning approach.

S22-P02-Generative AI models for evolutionary genomic inferences

Matteo Fumagalli

Queen Mary University of London

Evolutionary genomics is swiftly transitioning into a data-driven discipline thanks to the availability of large-scale genomic data and the need to study increasingly complex evolutionary scenarios. Machine learning, and in particular deep learning, algorithms have emrged as popular techniques for population genetic inferences. Typically trained with synthetic data generated by efficient simulators, deep neural networks have proved to be useful in the inference of neutral and adaptive processes in model and non-model species by maximising the amount of retained genomic information.

Here, I will showcase latest advances in using generative adversarial networks to infer evolutionary parameters. This technology uses antagonistic training of two competing neural networks to produce synthetic genomic data indistinguishable from real observations. By reverse-engineering the trained simulator, it is possible to estimate demographic and adaptive parameters.I will illustrate how we used this technology to estimate the recent evolutionary history of malaria vectors in sub-Saharan Africa. Understanding the genetic dynamics and migration patterns of mosquito populations from the Anopheles genus is crucial for effective control strategies against vector-borne diseases. We were able to accurately test for presence or absence of genetic migration among sub-Saharan mosquito populations, and their temporal changes in population size. We demonstrated how this approach is superior to common strategies based on compressing the information into summary statistics.I will finally present applications to other non-model species in light of varying data quality, and discuss future directions in the field.

S22-P03-COL: a method for identifying putatively functional circular RNAs

Zheng Li

Shanghai Jiao Tong University (SJTU)

Circular RNAs (circRNAs) are a class of generally non-coding RNAs produced by back-splicing. Although the vast majority of circRNAs are likely to be products of splicing error and thereby confer no bene fi ts to organisms, a small number of circRNAs have been found to be functional. Identifying other functional circRNAs from the sea of mostly non-functional circRNAs is an important but dif fi cult task. Because available experimental methods for this purpose are of low throughput or versality and existing computational methods have limited reliability or applicability, new methods are needed. We hypothesize that functional back-splicing events that generate functional circRNAs (i) exhibit substantially higher back-splicing rates than expected from the total splicing amounts, (ii) have conserved splicing motifs, and (iii) show unusually high back-splicing levels. We confirm these features in back-splicing shared among human, macaque, and mouse, which should enrich functional back-splicing. Integrating the three features, we design a computational pipeline named COL for identifying putatively functional back-splicing. Different from the methods that require multiple samples, COL can predict functional back-splicing using a single sample. Under the same data requirement, COL has a lower false positive rate than that of the commonly used method that is based on the back-splicing level alone. We conclude that COL is an efficient and versatile method for rapid identification of putatively functional back-splicing and circRNAs that can be experimentally validated. COL is available at https://github.com/XuLabSJTU/COL.

S22-P04-DIVERGE v4: a web server for large-scale analysis of functional divergence across multi-gene families

Yichang Chen

College of Pharmaceutical Sciences, Zhejiang University

Functional divergence plays a crucial role in understanding the evolution of protein functions and the emergence of functional specificity. Identifying sites associated with functional divergence is essential for defining specific protein functions, providing insights into the atomic basis of functional specificity and advancing our understanding of disease mechanisms, drug discovery, therapeutic targeting, and protein engineering. In this study, we present a webserver version of DIVERGE, a tool initially published as standalone software (Mol Biol Evol,2013,30,1713−9). The software and associated algorithms have since accumulated 1,539 citations across Google Scholar. DIVERGE was developed to analyze functional divergence and has maintained a strong, active user base since its last update.

Over the past three years, the software has received more than 20 citations each year, highlighting its wide use and importance. Compared to other programs, DIVERGE can accurately predicts functional divergence using statistical models. It calculates functional divergence coefficients and provides a site-specific divergence profile by processing multiple sequence alignments and user-defined subfamily divisions. With the rapid growth of multi-gene family data and improved computational resources, earlier versions of the software were limited in handling large-scale computations and lacked integration with modern computational frameworks. To overcome these limitations, DIVERGE v4 was developed with the following improvements: a) A robust new algorithm for large-scale functional divergence analysis in multi-gene families; b) A user-friendly webserver (https://pgx.zju.edu.cn/diverge) and an intuitive Python API; c) An analysis of 4,540 human protein families, providing a valuable resource for future research on protein functional specificity.

S22-P05-Macrogenetic atlas of prokaryotes

Chao Yang

Shanghai Institute of Immunity and Infection, Chinese Academy of Sciences

Macrogenetics investigates the patterns and predictors of intraspecific genetic variation across diverse taxa, offering a framework to address longstanding evolutionary hypotheses. Here, we present a comprehensive macrogenetic atlas of prokaryotes, integrating genomic and population genetic data (21 types, 50 parameters) from over 15,000 bacterial and archaeal species with phylogenetic, phenotypic, and ecological datasets. We uncover extensive correlations across data categories and explore their potential causal relationships. We find that the combined effects of effective population size (Ne) and mutation rate (u) serve as key drivers of genome size variation, but that within-species diversity instead depends on the interplay between Ne and the recombination rate. We show that as within-species diversity increases, selection plays an increasingly important role in structuring that diversity, for example creating convergent ecospecies structures in Streptococcus mitis and S. oralis. Our work implies that the most diverse species are champions of both evolutionary flexibility and exquisite adaptation, assembling extensive sets of coadaptated genes within a single recombining gene pool but are likely to be unstable due to high speciation rates. Overall, our atlas represents a widely applicable resource for prokaryotic research and offers novel insights into the drivers of macroevolution.

S22-P06-A new approach to integrating complex substitution models into molecular clock analysis and its implications for dating ancient symbiont evolution

Sishuo Wang

The Chinese University of Hong Kong

Establishing the evolutionary timeline of organisms using molecular clocks is crucial in evolutionary biology. Traditional substitution models assume the same substitution patterns across sites, which could lead to inaccurate results in phylogenetic reconstruction. This limitation can be addressed by complex substitution models which account for across-site substitution heterogeneity. However, current molecular clock methodologies face significant problems in accommodating complex substitution models, which have been demonstrated critical for studying the phylogenies of fast-evolving or deep evolutionary lineages. Hereby, we introduce phyloHessian, a Julia-based software designed to facilitate molecular dating under complex substitution models. phyloHessian employs various techniques, including an extension of CODEML's finite difference algorithm, to improve numerical accuracy and stability in calculating the phylogenetic Hessian matrix under diverse substitution models. The Hessian matrix is then integrated into PAML-MCMCTree's approximate likelihood framework to complete dating analyses. Extensive simulations demonstrate that the use of profile mixture models, such as C60, significantly enhances the accuracy of date and rate estimates, particularly in the presence of substitutional saturation and compositional heterogeneity, making it especially valuable for dating deep phylogenies. Additionally, profile mixture models display greater robustness to model and calibration specifications compared to their homogeneous counterparts. Further, Iweinvestigate the implications of molecular clock analyses using different substitution models on estimating the origin times of ancient symbiotic lineages. The methodologies developed in this study are readily applicable to a wide range of organisms, providing reliable molecular dating tools that meet the demands of molecular clock analyses under more realistic substitution models.

S22-P07-Serpentinization as the source of energy, electrons, organics, catalysts, nutrients and pH gradients for the origin of LUCA and life

Loraine Schwander

Heinrich Heine University Düsseldorf

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. CO2 was made widely available in Earth's oceans around 4.5 – 4.35 Ga years ago, after the Moon-forming impact caused CO2 to first degas into the atmosphere, where rain dissolved it into the oceans after the young Earth cooled down again. 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. 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 Saturn's icy moon Enceladus, providing a perspective on serpentinization as a potential source for life on other worlds.

S22-P08-A new information criterion to compare partition versus mixture models and its application to empirical datasets

Huaiyan Ren

Australian National University

In phylogenetics, partition and mixture models are widely employed to analyse large and heterogeneous genomic datasets using Maximum Likelihood methods. A common question in the field is: Which approach better fits the data — partition models or mixture models? Traditional model selection methods tend to show undesirable biases for such comparisons. Here, we introduce a new measure, marginal Akaike information criterion (mAIC), which adjusts the likelihood of partition models in a way that is comparable to mixture models.

We benchmark mAIC on a diverse set of empirical datasets and demonstrate that mixture models are generally favoured, but that model selection outcomes vary on sequence type (nucleotide vs. amino acid) and the partitioning scheme applied.

mAIC calculation is incorporated directly in IQ-TREE version 2.3.6.7.mf in the default output when using partition models.

S22-P09-Towards a Universal Model Specification Language for Phylogenetics

Walter Xie

University of Auckland

As phylogenetic models grow in complexity and datasets expand, there is increasing demand for a simple, precise language to specify models and analyze data. We introduce LinguaPhylo (LPhy), a lightweight, human- and machine-readable language designed to clearly define Bayesian phylogenetic models. LPhy includes a graphical interface for model building, data simulation, and generating natural language narratives for manuscripts. Additionally, the command-line tool LPhyBEAST converts LPhy models into BEAST2-compatible XML files for inference, enhancing collaboration and ensuring reproducibility.

LPhy extends beyond traditional phylogenetic tools, supporting large-scale calibrated studies, single-cell sequencing simulations for cancer research, read count substitution modeling, and protein structure evolution. It also facilitates new extensions, such as custom population size functions for demographic inference. By simplifying model specification and improving reproducibility, LPhy provides a powerful framework for computational evolutionary research.

S22-P10-Retriever: an imputation pipeline designed for non-model species

Meikun Zhou

UQ

Missing genomic data presents a significant challenge in biological analysis, particularly affecting reliability and interpretation of sequencing results. While current imputation tools effectively handle missing data in model organisms, their requirement for external reference panels - constructed through complete-coverage sequencing of large cohorts of individuals - creates substantial barriers for research in non-model species. To address these challenges, we developed Retriever, a novel imputation framework that constructs a chimeric reference panel directly from user-provided samples. This approach eliminates the need for external reference panels and is compatible with state-of-the-art software like Beagle, as well as machine learning (ML) algorithms, offering potential for improved accuracy through ML advancements. The chimeric reference panel, unique to Retriever, minimises resource demands while preserving crucial population structure. I will show how Retriever outperforms traditional methods, particularly for non-model organisms and populations with limited external data. Extensive benchmarking shows that Retriever achieves consistent accuracy exceeding 96% when tested in Arabidopsis, Saccharomyces and Gallus, with the option to integrate the chimeric reference panel into Beagle4 for further improvements. Future updates will focus on iterative refinement of imputed genotypes, further enhancing accuracy and reducing errors from initial approximations. Retriever provides an accessible, cost-effective, and accurate solution for genotype imputation, particularly for non-model species.

S22-P11-Spatial granularity and the advance of alleles under frequency-dependent selection

Florence Débarre

CNRS / Sorbonne Université

Some methods for the management of populations, like artificial gene drive, involve the spread of positive frequency-dependent alleles. Whether the allele spread spatially or not depends on its introduction threshold, but also on the granularity of the enviroment. In a landmark study, Barton (1979) showed that in a spatially discrete environment, too wide steps could halt the spread of a wave that would otherwise travel in a spatially continuous environment. The identification of the step width threshold however involved approximations. Here, we revisit this study with a detailed numerical exploration, and we show that the step width threshold is actually larger than Barton's approximation, meaning that in some conditions, an allele will spread while it was predicted to be stalled. Our findings have implications for the careful use of artificial gene drives for the control of natural populations.

S22-P12-A case study in improving the computational performance of Bayesian phylogenetic inference

Joëlle Barido-Sottani

CNRS

Bayesian phylogenetic inference is a powerful tool for reconstructing phylogenies and their associated diversification dynamics, giving a full picture of the process and the uncertainty present in the data. However, this type of inference is known to be more computationally expensive than alternative techniques such as Maximum Likelihood inference. Computational cost can strongly limit the use of Bayesian inference on datasets of realistic size, especially for more complex models. Here, I present my work on improving the performance of the package ClaDS, part of the Bayesian phylogenetic inference software BEAST2. The package implements the Cladogenetic Diversification rate Shifts model, a diversification model with progressive changes in speciation and extinction rates across lineages of a phylogeny. This model provides a realistic representation of gradual shifts in diversification rates, for instance due to environmental conditions. However, its initial implementation in BEAST2 showed very poor performance, especially when co-estimating the phylogeny and the model. To address this issue, I used several different strategies, focusing both on the underlying inference algorithm and its actual implementation into software, and evaluated their impact on the performance. Finally, I provide some general guidance for developers of similar methods when addressing performance concerns, based on my experience.

S22-P13-Automatizing high-throughput screens for positive selection and relaxation across hundreds of genomes

Bernhard Bein

Senckenberg Society for Nature Research

Long-read sequencing and scalable annotation and orthology inference methods like TOGA (Tool to infer Orthology from Genome Alignments) enable researchers to create reference-quality genome assemblies and high-quality annotations with ease. However, coupling divergent phenotypes with differences in gene sequences caused by episodic diversifying (positive) selection or relaxation of selection still poses a challenge, since slight differences in how codon alignments are cleaned, which assemblies are included in a screen for selection and which parameters are modelled can lead to stark differences in the results, increasing the risk of false positive calls. Here, we present sleasy (SeLection EASY), a fully-automated pipeline written in snakemake that can run all steps in a genome-wide selection screen, scaling to tens of thousands of input genes from hundreds of assemblies. Sleasy takes orthologous gene sequences, aligns these sequences with MACSE2 and computes gene trees with IQtree2. It then feeds alignments and trees to tools of the HyPhy suite (aBSREL, BUSTED, MEME, RELAX) and creates result tables and visualizations. The pipeline is highly customizable and comes with TOGA integration, but can be applied with any type of input annotation. Using a single command, we recreate results from two different datasets seamlessly. Moreover, we leverage sleasy's ability to run the same dataset with different parameter settings on one existing and one new dataset to benchmark the impact of these parameters on results, and develop guidelines for how to run robust, reproduceable selection screens.

S22-P14-Structural phylogenetics: tracing evolution through protein structure

Caroline Puente-Lelievre

University of Auckland

Protein structure represents a rich and largely untapped resource for uncovering evolutionary relationships. Protein structural phylogenetics, an interdisciplinary branch of molecular evolution, leverages three-dimensional structural data to trace evolutionary histories and explore the diversity of protein structures and functions. Historically, the field has been constrained by the limited availability of high-resolution structural data. Recent breakthroughs in artificial intelligence have revolutionised the field by making protein structural data widely accessible. Structural alphabets offer a promising approach, building upon decades of progress in sequence and character-based phylogenetics. At the intersection of structural biology, biochemistry, evolutionary biology, and computational biology, structural phylogenetics requires collaborative efforts to further develop this exciting and promising field towards more translational evolutionary research. As molecular phylogenetics continues to shift towards protein structure, we enter the third generation of phylogenetics, where deep learning and structural insights redefine our understanding and applications of evolutionary history.

S22-P15-Positive Selection Inference by Deep Learning

Han Liu

Institute of Zoology, Chinese Academy of Sciences

The detection of positive selection signal in coding sequence (CDS) data is a routine task in molecular evolution analysis. Existing methods based on mechanistic models and likelihood inference rely on simplified assumptions, often overlooking various heterogeneity factors in sequence evolution that may significantly impact the prediction performance. Meanwhile, complicated mechanistic models suffer from low statistical power due to inherently limited data size. To solve this dilemma, we propose the approach of Positive Selection Inference by Deep Learning (PSIDL). We constructed a transformer model with CDS alignment as input to predict if the sequences of designated foreground taxa underwent positive selection. The model was trained by large simulated datasets incorporating various heterogeneity factors such as synonymous rate variation, dinucleotide substitution, and codon usage bias. We then tested the performance of our trained models and compared with existing methods, by simulated test datasets spanning diverse parameter subspaces within and outside those used in the training process. Thus, we found that our models achieved superior performance in virtually all tests, even when generalized to test dataset with novel parameter combinations. In addition, we applied our models to multiple real-world cases, and obtained comparable or more positive selection signals when expected. Overall, PSIDL exhibited promising performance in detecting positive selection pattern of coding sequences particularly regarding heterogeneous sequence evolution scenarios.

Symposium 23: Thematic symposium: Symbiosis

S23-P01-Rapid evolution of mitochondrion-related genes in haplodiploid arthropods

Yiyuan Li

Institute of Plant Virology, Ningbo University, Ningbo, China

Mitochondrial genes and nuclear genes cooperate closely to maintain the functions of mitochondria, especially in the oxidative phosphorylation (OXPHOS) pathway. However, mitochondrial genes among arthropod lineages have dramatic evolutionary rate differences. Haplodiploid arthropods often show fast-evolving mitochondrial genes. One hypothesis predicts that the small effective population size of haplodiploid species could enhance the effect of genetic drift leading to higher substitution rates in mitochondrial and nuclear genes. Alternatively, positive selection or compensatory changes in nuclear OXPHOS genes could lead to the fast-evolving mitochondrial genes. However, due to the limited number of arthropod genomes, the rates of evolution for nuclear genes in haplodiploid species, besides hymenopterans, are largely unknown. To test these hypotheses, we used data from 76 arthropod genomes, including 5 independently evolved haplodiploid lineages, to estimate the evolutionary rates and patterns of gene family turnover of mitochondrial and nuclear genes.

We show that five haplodiploid lineages tested here have fast-evolving mitochondrial genes and fast-evolving nuclear genes related to mitochondrial functions, while nuclear genes not related to mitochondrion showed no significant evolutionary rate differences. Among hymenopterans, bees and ants show faster rates of molecular evolution in mitochondrial genes and mitochondrion-related nuclear genes than sawflies and wasps.

Our results reject the small population size hypothesis in haplodiploid species. A combination of positive selection and compensatory changes could lead to the observed patterns in haplodiploid species. The elevated evolutionary rates in OXPHOS complex 2 of bees and ants suggest a unique evolutionary history of social hymenopterans.

S23-P02-Metatranscriptomic and metagenomic analyses reveal the roles of coral holobionts in recovery phase during natural bleaching event: A case study of subtropical Acropora solitaryensis

Yingxin Li

Lingnan University

Ocean warming has led to widespread coral bleaching over the past half-century. While many studies have explored potential bleaching mechanisms under controlled experiments, little is known about the responses and dynamics of coral holobionts that contribute to recovery during natural bleaching events. This study investigates the recovery process by analyzing physiological parameters, microbiome structures, metatranscriptomic, and metagenomic data in subtropical Acropora solitaryensis during the 2022 natural bleaching event. Bleached and non-bleached Acropora colonies were tagged and sampled across four recovery phases: 0, 1, 4, and 11 weeks, in Hong Kong waters. Our results showed that coral pigmentation and Fv/Fm increased after one month, returning to non-bleached levels by week 11. This recovery trajectory was reflected in the diversity and compositional patterns of both symbiotic zooxanthellae and bacterial communities. Specifically, zooxanthellae diversity increased, while bacterial diversity declined after four weeks. Metagenomic analysis revealed persistent differences in bacterial metabolic processes, including carbohydrate, amino acid, cofactor, vitamin, and energy metabolism, which contribute to nutrient cycling within the holobiont. Furthermore, metatranscriptomic analysis indicated significant transcriptional changes in the coral host, but not in the symbiotic zooxanthellae, with Acropora genes associated with stress response, immunity, embryonic and system development, and morphogenesis and tissue formation potentially facilitating cellular and physiological recovery of the coral host. Overall, this study enhances our understanding of coral recovery from natural bleaching, highlighting that visibly recovered corals still exhibit impairments in cellular metabolic and immune responses. This insight contributes to improved predictions of coral survival and resilience under climate change.

S23-P03-A genomic approach to investigate the specificity of symbiont-conferred defence in the tripartite system aphid – Hamiltonella - parasitoid wasp

Giacomo Moggioli

EAWAG

Parasitoid wasps exploit many insect species as hosts by injecting their eggs inside them, allowing the developing larvae to feed on and ultimately kill their hosts. Aphids, however, have evolved a unique strategy to counter parasitoid infections through a facultative symbiosis with the bacteria Hamiltonella defensa. The resistance against parasitoid wasps, provided by Hamiltonella, is hypothesized to be mediated by toxins, which arrest the development of wasp larvae and are encoded in the bacteriophage APSE associated with this endosymbiont. In Aphis fabae, Hamiltonella-conferred resistance to parasitoid wasps is highly specific: different combinations of Hamiltonella strains and parasitoid wasp genotypes result in higher or lower levels of protection to the aphid. Despite previous efforts, the genomic basis of Hamiltonella-mediated toxicity and its specificity in interactions with different parasitoids is not fully understood. To unravel the genetic mechanisms involved in this tripartite system, we generated 14 high-quality genome assemblies of Hamiltonella strains isolated from diverse aphid populations and introduced them into a single clonal genotype of A. fabae. Phenotypic data and comparative genomic analyses of Hamiltonella and APSE aim to identify key genes and pathways involved in toxicity and in the resulting spectrum of protective phenotypes. Additionally, we examine the co-evolutionary dynamics between Hamiltonella and parasitoid wasps, investigating whether specific genomic features of Hamiltonella and APSE correlate with resistance to different parasitoid genotypes. This work sheds light on how genetic variation in Hamiltonella influences its defensive capabilities, advancing our understanding of the ecological and evolutionary pressures driving this intricate host-symbiont-parasitoid system.

S23-P04-Parasites strategies of coinfection: the evolution of within-host parasite-parasite interactions in Plasmodium falciparum

Ellen M. Leffler

Department of Human Genetics. The University of Utah

Coinfections occur when genetically distinct parasites infect a single host, where they can interact directly or indirectly with each other. The most studied interaction between them is competition for resources, which leads to increased virulence. Plasmodium species, the causal agent of malaria, are an excellent model to study the evolution of parasite interactions, as coinfections are common in regions with high transmission. Yet, little is known about the role of parasite-parasite interactions in Plasmodium evolution, nor how they may be impacted during eradication efforts. Using publicly available data from P. falciparum populations across nine countries, we extracted the frequency at which each allele was found in coinfected hosts and fit a regression model to identify alleles that co-occur together more often than expected (cooperation) or less often than expected (competition), given their overall circulating frequency. Several of our candidate coinfection-associated alleles map to genes involved in parasite invasion and resource acquisition processes, suggesting a role in niche differentiation. We hypothesize that some of these may be under balancing selection due to tradeoffs in fitness in single vs. multiple infections. Indeed, we find approximately a quarter of our candidate alleles are under balancing method using both Betascan1, and an alternative metric comparing shared and private synonymous and nonsynonymous polymorphisms. Our work points to candidate polymorphisms that may help parasites thrive in coinfections and could have functional consequences when coinfection rates change during the malaria elimination measurements.

S23-P05-Symbiosis shapes microbial spatial organization and functional activity in cnidarian holobionts

Guoxin Cui

King Abdullah University of Science and Technology, Saudi Arabia

Microbial communities are integral to cnidarian holobiont function, yet their tissue-specific roles and interactions with symbiotic algae remain poorly understood. Here, we combine tissue-specific metatranscriptomics and stable isotope labeling to investigate bacterial localization, gene expression, and contributions to nitrogen assimilation in the model cnidarian Aiptasia. We reveal distinct microbial communities across tissue layers, with bacterial composition and activity significantly influenced by symbiotic state. Nitrite reductase was strongly upregulated in the symbiotic gastrodermis, suggesting enhanced bacterial nitrogen processing in regions rich in photosynthetically derived energy. Stable isotope tracing further demonstrated direct microbial contributions to essential amino acid biosynthesis, with bacteria and algal symbionts synergistically enhancing nitrogen assimilation. These findings highlight the spatial and functional integration of microbial metabolism within cnidarian tissues, advancing our understanding of holobiont nitrogen cycling and its implications for cnidarian resilience in changing environments.

S23-P06-Evolutionary turnover of symbiotic bacteria and their metabolic genes in termite gut microbiomes

Ales Bucek

Biology Centre

Gut microbiota function is crucial to termites' ecological success, yet the evolutionary turnover of symbiotic bacteria and their metabolic genes remains poorly understood. To address this, we analyzed gut microbial metagenomes of 145 termite species. Our findings reveal that termite gut prokaryotic communities have remained remarkably conserved since termites first appeared~150 million years ago (Arora et al. 2022). Cophylogenetic patterns between termites and bacteria suggest vertical transmission plays a major role, though horizontal transfer can disrupt these associations. By constructing bacterial phylogenies using marker genes, we identified 18 bacterial lineages showing cophylogenetic patterns with termites over millions of years, some dating backto their common ancestor (Arora et al. 2023). Additionally, we discovered 131 termite-specific carbohydrate-active enzyme (CAZyme) gene clusters involved incellulose digestion, with bacterial CAZyme sequences linked to Cryptocercus or Mastotermes, the sister group of all termites. This suggests termites rely upon many bacterial CAZymes endemic to their guts rather than repeatedly acquired from non-termite gut microbial reservoirs (Berankova et al. 2024). This finding might be surprising, given the generally prevalent horizontal gene transfer (HGT) in prokaryotes but is supported by our analysis of 14 complete bacterial genomes from Nasutitermes takasagoensis gut bacteria indicating that HGT is infrequent. However, it significantly impacts key mutualistic functions such as amino acid synthesis and nitrogen metabolism (Prokhorova et al., in prep). These findings provide preliminary insights into how high termite-bacteria fidelity and low HGT may jointly maintain reservoirs of endemic gut microbiota and enzymes in termite guts.

Symposium 24: Thematic symposium: colonizing new environments and invasiveness

S24-P01-The macroecological impact of thermal tolerance evolution in Phrynosomatid lizards

Kiran Masroor

Yale University Department of Ecology & Evolutionary Biology

The capacity to tolerate diverse thermal environments is crucial for species ' geographic range and survival in a changing climate. In this study, we examine the evolution of thermal tolerance breadth (TTB) in phrynosomatid lizards, investigating its relationship with environmental variability and intrinsic traits. We found that TTB is more strongly correlated with regional rather than local temperature variation, underscoring the role of macroevolutionary adaptations in geographic expansion. Contrary to expectations, traits like parity mode and body size did not significantly predict TTB, suggesting that environmental factors primarily shape thermal adaptations. Our analyses challenge the paradigm that cold tolerance (CTmin) evolves more readily than heat tolerance (CTmax). Instead, we found more frequent shifts in CTmax during phrynosomatid radiation compared to CTmin. Surprisingly, viviparity, often linked to adaptation in cold environments, was not a determinant of TTB. Similar thermal tolerances were observed in both viviparous and oviparous species inhabiting comparable climates, highlighting the complexity of intrinsic-extrinsic interactions in thermal evolution. These findings advance our understanding of biodiversity patterns and emphasize the importance of physiological adaptability in navigating climate change. Incorporating TTB into conservation frameworks could help identify species with narrow thermal tolerance ranges that are more vulnerable to rising temperatures. Extending this research across broader taxonomic groups promises further insights into the evolutionary dynamics shaping species distributions and resilience.

S24-P02-Evolutionary History of the Na+/H+ Antiporter Gene Family in Arthropoda

Patricia Zito

University of Wisconsin-Madison

Populations of the copepod Eurytemora affinis species complex have invaded freshwater environments multiple times independently from saline habitats. During these invasions, the Na+/H+-antiporter (NHA) gene family has repeatedly been under selection. To explore evolutionary patterns of the NHA family, our goals were to (1) reconstruct the first arthropod NHA phylogeny, (2) analyze signatures of selection (dN/dS) across the phylogeny, (3) map SNPs under selection onto predicted NHA-7 protein models of E. affinis complex, and (4) infer a co-expression gene network for genes showing parallel differential expression between saline and freshwater populations.

We curated 165 NHA sequences from 60 species to reconstruct phylogenies using Maximum Likelihood and Bayesian approaches. We determined phylogenetic branches and sites under selection using HyPhy and predicted protein models using Alphafold2. Finally, we inferred a co-expression network using RNA-Seq data with DESeq2 and EdgeR.

The evolutionary history of NHA was unique among cation exchangers, with taxon-specific paralogs forming predominantly after major arthropod lineages diverged. Interestingly, the SNP (R283H), which had shown signatures of selection following freshwater invasions, mapped to a coiled insertion between transmembranes 3 and 4 in the dimerization domain of the protein model. Interestingly, another SNP (V310C) associated with salinity tolerance mapped near the ligand binding site, potentially altering protein conformation. Finally, ion transporter genes formed a co-expression network showing coordinated gene regulation modules during invasions. This study uncovers a dynamic evolutionary history of the NHA family and links sites under selection during invasions to potential functional sites within the NHA protein.

S24-P03-Convergent Relaxed Selection in Independent Freshwater Adaptation of Landlocked Rhinogobius

Wanying Xu

Sun Yat-sen University

Similar environmental changes can repeatedly drive the loss of previously adaptive traits through relaxed selection, but this process has not been extensively studied. During multiple independent transitions from seawater to freshwater, different lineages within Rhinogobius may exhibit convergent relaxed selection in adapting to freshwater while losing traits associated with seawater adaptation. In this study, we assembled a high-quality chromosome-level genome of Rhinogobius formosanus and performed comparative genomic analyses using whole-genome data from 51 Rhinogobius individuals collected from 41 sites across 18 river systems in China. Additionally, transcriptomic analyses of the amphidromous R.leavelli were conducted under varying salinity conditions to identify gene expression changes. Our results indicate that landlocked Rhinogobius predominantly experienced relaxed selection, with significant overlap of relaxed genes between three landlocked clades. These genes are mainly associated with ion channel regulation, neural function and development. We also find that genes that exhibit both relaxed selection and significant differential expression under different salinity conditions were also repeatable among the three clades in all tests. Notably, the signals of convergent relaxed selection were stronger compared to those of intensified selection and positive selection. Overall, our findings provide strong evidence of convergent evolution driven by relaxed selection in Rhinogobius, offering insights into potential non-adaptive convergent evolutionary mechanisms.

S24-P04-Parallel evolution of terrestrial egg deposition in Old-World Pila and New-World Pomacea snails (Ampullariidae)

Yufei Zhou

Lingnan university

Freshwater apple snails (Ampullariidae) consist of members that exhibit both underwater and aerial oviposition, which provide insight into for key step for terrestrialization. Specifically, with all sister genera laying egg underwater, the Old-World Pila and New-World Pomacea have separately evolved the out-of-water oviposition. Adaptation to aerial oviposition is often accompanied by convergent changes across Pila and Pomacea, epitomized by the calcareous eggshell and specific egg proteins of the perivitelline fluid (PVF). Here we generated chromosome-level genome of Pila celebensisand scaffold-level assembly of Pila pesmei, which have double genome size of ~1Gb compared with other ampullariids (~500Mb). Through comparative genomic and phylogenetic analyses, we investigate genome-wide gene evolutionary dynamics and explore the evolutionary mechanisms between these two aerial oviposition ampullariid families. Our results identified and analysied parallel and convergent evolution at the orthogroup level in Pila and Pomacea, the expanded functions primarily involve metabolic, response to stimulus and UV radiation, indicating that parallel expansions of gene families played essential roles in facilitating adaption to aerial oviposition. Additionally, PVF gene evolutionary dynamics was explored to assess difference between underwater and aerial oviposition species, such as massive trandem duplicaton of PV1 genes and intra-chromosomal duplication of two PV2 subunits (MACPF and tachylectin) in Pila and Pomacea. Overall, this study enhances our understanding of the parallel evolution of terrestrial reproduction in Ampullariidae, paving the way for a deeper comprehension of terrestriality in gastropods.

S24-P05-Comparison of cold adaptation of Pomacea canaliculata from lakes and paddy fields and the underlying molecular mechanisms

Honglei Shi

College of Life Sciences, China Jiliang University

The apple snail Pomacea canaliculata, native to South America, is among the world's worst invasive species, which has caused serious damage to freshwater biodiversity and rice production in Asia. The suitable habitats for P. canaliculata were believed to be in tropic and subtropic areas; however, it has recently spread north to the temperate zone in China. In this study, we collected samples from lake and paddy fields to investigate the molecular mechanisms in cold adaption of P. canaliculata. After raising the snails in the lab (25 ± 1 ℃, L:D = 14 h:10 h) for two weeks, we decreased the temperature by 5℃ every 5 days from 25℃ to 0℃. The mortality rates of females and males were 32.26% and 92% for the lake and 10.75% and 36.36% for the paddy, respectively, indicating that P. canaliculata was not tolerant to freezing, with the lake population less tolerant to cold than the paddy population. We sequenced the transcriptome of the foot tissues from female snails at 10℃ and 0℃ to reveal the indirect and direct cold effects, respectively. The snails from the lake upregulated pathways related to regulation of actin cytoskeleton and adipocytokine signaling pathway, while the snails from the paddy activated immune and adipocyte differentiation pathways at 10℃. The differentially expression genes at 0℃ from both habitat types were focused on the pathways related to autophagy. Our results provide important information for understanding invasion mechanism of the apple snails.

S24-P06-Invasion genomics trace recent human-mediated expansion events of the swarming fly Plecia longiforceps in East Asia

Donghee Kim

School of Biological Sciences, Seoul National University, Seoul 08826, Republic of Korea

As human activities increasingly disturb ecological systems, range-expanding organisms provide valuable insights into anthropogenic impacts on biodiversity. Plecia longiforceps, previously documented only in southeastern China and Taiwan, has been recently reported in the Ryukyu Islands of Japan and the Seoul Metropolitan Area of South Korea. In Korea, its yearly mass emergence and mating swarms are attracting major public attention. Using whole-genome resequencing of 155 specimens across six regions, we found P. longiforceps populations in Taiwan-Ryukyu and Korea-China to be highly diverged. Korean populations likely originated from coastal northern China in or around the Shandong Peninsula, while Ryukyu populations likely stemmed from Taiwan. Northern P. longiforceps populations in Korea and Shandong have experienced gradual population declines after divergence from populations in southeastern China and recently serious bottlenecks, resulting in reduction of genetic diversity down to a half of that in southeastern China. We found genes linked to protein glycosylation and amyloid metabolic processes to be highly differentiated between northern and southern populations, while genes linked to fatty acid synthesis and lipid metabolism are uniquely differentiated in the Korean population, which represents the northernmost range of P. longiforceps. We speculate these pathways to be related with adaptation to colder climates during northward expansions. This study illuminates the demographic history and genetic mechanisms underpinning the successful invasion of P. longiforceps into new environments.

S24-P07-Evolutionary History and Molecular Evolution of Na+/K+-ATPase alpha-subunit Gene Family in the copepod Eurytemora affinis species complex

Jinhui Wang

University of Wisconsin-Madison

Populations of the copepod Eurytemora affinis species complex have recently invaded from saline to freshwater habitats. Notably, gene paralogs of the Na+/K+-ATPase a-subunit (NKA-a) gene family have repeatedly shown signatures of selection during these invasions. We aimed to explore phylogenetic relationships and patterns of molecular evolution for this gene family in the E. affinis complex and across the phylum Arthropoda. A phylogeny based on NKA-a gene sequences from high-quality genomes of 42 arthropod species revealed two major groups (I and II) within the NKA-a gene family. The E. affinis species complex, Daphnia, and four species of chelicerates showed NKA-a gene family expansions, with 4 to 9 of paralogs. Six sites under positive selection (using HyPhy's MEME) were located in the A, N, and M7 domains and in the inter-domain linker region. Using the branch-site aBSREL in HyPhy, two branches (ancestral nodes of NKA-a-1,4,5 and NKA-a-2,6) showed signatures of positive selection within the copepod clade. In previous research, NKA-a-1,2,4,5 also showed signatures of selection during independent saline-to-freshwater invasions and in laboratory salinity selection lines. When we mapped the SNPs under selection from previous studies onto the NKA-a protein model, two amino acid substitutions in A domain and M2 helices alter hydrophobicity and potentially impact stabilization of conformation of these domains. This change could influence the opening and locking the cytoplasmic gate. Our results contribute insights into the evolutionary history of the NKA-α gene family in Arthropoda and molecular mechanisms during range expansions of invasive populations in aquatic habitats.

S24-P08-Cyclical parthenogenesis helps expand into marginal habitats in some fitness landscapes

Zimo Yang

Auburn University

Previous theoretical work has shown that for quantitative traits, rates of adaptive evolution are similar in cyclical parthenogens and obligate sexual populations, but this assumes that populations are under directional selection on a smooth and simple adaptive landscape. In obligate bisexuals, high rates of recombination steadily churn up genetic diversity, and selection pushes a population towards a new phenotypic optimum with frequent small steps. In contrast, in cyclical parthenogenetic populations, over the course of several asexual generations, cryptic genetic diversity can accumulate via complementarity between otherwise mildly deleterious mutations. Then, when sex happens, recombination exposes much of that cryptic diversity to selection. Consequently, selection pushes a cyclically parthenogenetic population towards a new phenotypic optimum with more intermittent, larger adaptive steps. On a smooth adaptive landscape, the net rate of adaptation depends little on the frequency of recombination. On the other hand, we predict that in heterogeneous environments, in which demes have distinct phenotypic optima and can be classified as "core" (in which populations are well adapted) and "margin" (in which populations are mal adapted), cyclical parthenogens will be more likely to expand into marginal habitats. With individual-based simulation models allowing for potentially pleiotropic and epistatic mutation, we demonstrate that cyclic parthenogenetic populations do indeed tend to adapt to marginal habitats more rapidly than obligated bisexual populations on certain fitness landscapes. We are trying to describe and explain this pattern by analyzing the evolution of genetic structure, mutation effect, and standing genetic variation.

S24-P09-Evidence of positive selection shaping polar adaptation in fishes

Tina Cai

University of California, Santa Cruz

Polar environments are extremely oxygen-rich, with constant or seasonally subfreezing temperatures. Multiple lineages of fishes from temperate waters have independently colonized polar oceans, where they were presented with new selective challenges. Suborders Zoarcoidei and Notothenioidei, which contain most of the polar fish species, convergently evolved a suite of novel traits to live in these extreme habitats, including the production of ice-binding antifreeze proteins. This environmental transition also led to the relaxation of selection on genes that became less essential, including those that respond to temperature and oxygen variability. Despite a growing understanding of adaptive traits in polar fishes, the breadth of molecular pathways and mechanisms driving polar adaptation remains unclear. Here, we analyze 25 high-quality, long-read reference genomes of Zoarcoidei and Nototeniodei species to assess convergent and lineage-specific patterns of polar adaptation. We identify protein-coding genes that experienced convergent patterns of positive selection across species and assess large-scale patterns of gene expansion/contraction to track gene families' evolutionary dynamics in extreme cold. We hypothesize that genes involved in mitigating cold-induced stresses, such as antifreeze protein and cellular organization-related genes, will likely be subject to expansion and positive selection. We also examine evidence of relaxed selection on genes and pathways that became less critical following adaptation to polar regions. Genes associated with fitness costs, such as heat shock protein and hypoxia-related genes, will likely undergo relaxed selection. Our findings provide insights into how genome evolution underpins cold adaptation and highlight the evolution of pathways that promoted the performance of polar fishes.

S24-P10-Genomic insights into evolutionary history and local adaptation of Drosophila suzukii

Qi Fu

Westlake University

Climate change and globalization are accelerating biological invasion, underscoring the urgent need to comprehend invasion history and adaptation mechanism of invasive species to make effective management strategies. D. suzukii, an invasive pest species originating from East Asia, has rapidly extended its range globally over the past 15 years, presenting an opportune model organism for studying adaptation during invasions. While previous research has uncovered its invasion history, a comprehensive understanding of global genetic diversity and local adaptation remains elusive, primarily due to insufficient sampling from its native region. In this study, we sequenced 178 D. suzukii lines collected from diverse environment gradients in China. Our findings indicate that: (i) worldwide suzukii populations can be separated into four major genetically distinct groups, with a moderate differentiation observed within the China populations; (ii) native populations have undergone population expansion, whereas invasive populations have experienced a bottleneck; (iii) the invasive histories in North America and Europe are independent, and recurrent gene flow between China and invaded populations is limited in recent time; (iv) multiple environmental variables such as altitude, temperature, precipitation and human land usage drive local adaptation, and candidate genes under positive selection are enriched in processes including antennal development, olfactory behavior, response to ethanol; (v) chromosomal inversions facilitate local adaptation and shape population structure of western and eastern China populations in the presence of gene flow. Our study thus offers insights into the invasion success of this agricultural pest, providing a foundation for future research aimed at developing targeted management strategies.

Symposium 25: Advancing Inclusion, Diversity, Equity, and Accessibility in Molecular Biology and Evolution

S25-P01-An updated solutions manual and e-learning platform for Ziheng Yang's books: "Computational Molecular Evolution" and "Molecular Evolution: A Statistical Approach"

Jianhao Lv

The Chinese University of Hong Kong

Molecular evolution and phylogenetics are crucial domains in biology, yet pose a high entry barrier due to their interdisciplinary nature. Ziheng Yang ' s books, "Computational Molecular Evolution" [1] and "Molecular Evolution: A Statistical Approach" [2], are widely acknowledged as highly impactful textbooks in the field, offering not only comprehensive content but exceptional exercises covering proofs/computations, programming, and software operation. However, the absence of solutions to these exercises presents great obstacles for biologists aiming to deepen their understanding of the knowledge in molecular evolution.

Hereby, we present an updated solutions manual for all exercises included in the two books. Specifically, on the basis of our previous solutions manual [3], the updated solutions greatly focus on alternative solutions, more detailed codes, and errata. In addition, we provide several more exercises to cover the sections that do not have exercises in the original books. We have also updated the online learning platform (https://github.com/sishuowang/Solutions_Manual_CME2006_MESA2014). We believe that our work will foster a community-driven effort to assist anyone interested in delving deeper into computational molecular evolution and phylogenetics.

S25-P02-Visualization of ancient humans from the Allen Ancient DNA Resource using ArcGIS: applications of interactive web maps for ancient DNA studies

Woojung Yi

Department of Biology, University of Richmond, Richmond, VA

The Allen Ancient DNA Resource (AADR) is a widely used resource that collects ancient DNA from humans across the globe. With over 10,000 published ancient humans included to date, the AADR has greatly facilitated the study of human population history and adaptation. However, one barrier to effectively utilizing this resource is the lack of a graphic user interface to assess the availability of ancient human samples for a region or time period of interest. Here, we developed a web-based interactive map to visualize ancient humans from the AADR.. With this freely available, public-facing, constantly-updated, AADR visualizer, researchers are now able to quickly and easily explore the available ancient human DNA for subsequent population genetic analyses. Moreover, this tool is ideal for educational purposes, reducing the barriers to entry for researchers new to human evolutionary genomics.

S25-P03-Genomic history and ancestral origins of Afro-Mexican populations

María Fernanda García Rodríguez

LIIGH

Three main continental ancestries have shaped the genetic diversity of the present-day Mexican population: Indigenous American, European, and African. Although the African contribution is proportionally smaller at the population level, it holds significant importance in understanding the country's genetic diversity and history. However, few studies have explored the genetic contributions and population history of Afro-descendant communities in Mexico. Our research aims to characterize and analyze the genetic diversity of Afro-Mexican populations, tracing their genetic origins at the subcontinental level, both within Africa and Mexico. We conduct genomic analyses of volunteer participants (N= 381) from Veracruz and the Costa Chica regions of Guerrero and Oaxaca, alongside implementing methods for detailed local ancestry inference analyses. Preliminary results suggest that the African ancestries in Afro-Mexicans can be traced to two different sources within sub-Saharan Africa, while the Indigenous American component clusters with reference populations near their geographical location in Mexico. Furthermore, including ancient African individuals from the Colonial Period may provide additional insights into the origins of African ancestries in Afro-Mexican populations. The main findings of this research highlight the fundamental contribution of Mexico's Afro-descendant populations to the cultural and genetic diversity of Mexico.

S25-P04-snputils: A Python library for lowering barriers to genetic variation and population structure processing

Marçal Comajoan Cara

UC Santa Cruz; Stanford University

The growing fields of genomics and population genetics present significant technical barriers, with researchers often needing to develop custom software to handle complex data analysis workflows. We present snputils, an open-source Python library designed to lower these barriers by providing streamlined tools for SNP and ancestry data analysis. By simplifying essential operations like input/output operations, data manipulation, file format conversion, and visualization, snputils reduces the technical expertise required to conduct genomic analyses. The library supports a wide range of commonly used file formats, including VCF (.vcf, .vcf.gz), PLINK1 (.bed, .bim, .fam), PLINK2 (.pgen, .pvar, .psam), local ancestry inference formats (.msp), and global ancestry files (.Q, .P), making it easier for researchers to work with genomic data from any population. The library includes advanced dimensionality reduction methods such as ancestry-specific PCA and MDS, along with comprehensive admixture mapping functionalities. These tools are particularly valuable for analyzing admixed populations and populations with complex ancestry patterns that may be poorly captured by standard methods. The development of snputils represents a community-driven effort to lower technical barriers in population genetics and promote more inclusive research practices. snputils is freely available on GitHub, with comprehensive documentation and tutorials available at https://snputils.org. Through this open-source tool, we aim to contribute to a more accessible and inclusive environment for genomic research and facilitate broader participation in the field.

Symposium 30: Open symposium session B

S30-P01-Unveiling the Genetic and Molecular Basis of Seasonal Camouflage in Hares

Huijuan Bi

Stockholm University

Seasonal camouflage, characterized by color changes in fur or plumage, enables species to blend into seasonally snow-covered environments. While previous work has linked genetic variation in pigmentation gene ASIP with alternative winter coat colors in several hare species, the molecular mechanisms by which environmental cues drive these phenotypic changes remain elusive. Here, we utilize the mountain hare, which transitions from summer brown to winter grey, as a model to uncover the regulatory mechanisms behind color change. Using whole-genome sequencing across multiple hare species, we identified two structural variants in intronic regions of ASIP. Both variants are exclusively present in species with a winter-white phenotype, exhibiting complete correlation across species. Screening transcription factor binding sites within these regions revealed two candidates: NPAS2, a circadian regulator significantly negatively correlated with ASIP expression and potentially involved in regulating molt timing, and CTCF, a genome architecture regulator co-localizing a regulatory element within the ASIP intron and promoter. To deepen our understanding of the roles of different skin cell types in seasonal camouflage, we constructed a comprehensive molecular map of hare skin using single-cell RNA sequencing. We analyzed 19,669 cells obtained from two skin patches: brown skin and grey skin. Differentially expressed genes and known markers allowed us to identify six main cell types. Notably, brown skin exhibited a higher proportion of immature melanocytes, while grey skin showed a greater percentage of mature melanocytes. This study sheds light on the potential regulatory mechanisms underlying seasonal camouflage, advancing our understanding of environmental adaptation.

S30-P02-Evidence for large intraspecific genome size diversity in arthropods

MOHAMMED M.TAWFEEQ

1. Evolutionary Biology & Ecology, Université libre de Bruxelles (ULB). 2. Interuniversity Institute of Bioinformatics in Brussels – (IB)2, Brussels.

The term "genome size" refers to the total amount of DNA present in a haploid cell, an important biological trait known to vary significantly across eukaryotic species. Despite numerous studies about interspecific differences in genome size, its intraspecific diversity remains scarcely studied to date. Using arthropods as a model system, we investigated intraspecific genome size variations in three selected groups of arthropods: i) niphargid amphipods; ii) Nepa insects; iii) Asellus isopods. Mitochondrial sequences COI and nuclear 28S markers were used to construct haplotype networks to delineate species boundaries. Using a refined Feulgen Image Analysis Densitometry protocol (notably based on three standards instead of one to return precise genome size estimates), we found large, 2- to 3-fold intraspecific differences in genome sizes between closely related individuals in each of these groups. Sequencing the genomes of conspecific individuals of contrasted genome sizes will be required to find out whether the observed differences are due to whole-genome duplications, copy-number variations, supernumerary chromosomes and/or variations in the relative amount of non-coding sequences versus gene-coding ones.

S30-P03-Computer Simulation and Mathematical Modeling of the Interactions between Ecological and Sexual Selection to Reveal the Mechanism of Sympatric Speciation

John Lin

Evolution Research, Inc.

Sympatric speciation is an unresolved mystery in evolutionary biology. Ecological niche specialization followed by premating reproductive isolation (RI) is considered a key initial step, but the causal link between them remains unclear. This study aims to develop a simplified yet comprehensive model ecosystem to uncover the mechanisms driving sympatric speciation. We first used individual-based computer simulations to investigate how ecological and sexual selection interact to produce premating RI in sympatric populations under disruptive ecological selection. We then derived mathematical models to describe the nonlinear dynamics of the system and displayed their numerical solutions as phase portraits. Our findings reveal that under favorable conditions, fixed points may emerge in the phase portraits, signifying stable mating-bias-allele polymorphisms and varying degrees of premating RI between niche ecotypes. The models allow us to identify the precise parametric conditions necessary to produce such fixed points and enable high-mating-bias alleles to invade and establish premating RI. Restricting migration between niches tends to impede mutant invasion, while multi-locus mating traits produce hybrids that can hinder speciation. Based on these results, we propose a five-stage mechanism of sympatric speciation: Disruptive ecological selection initially drives the invasion of high-mating-bias alleles, quickly establishing premating RI. This facilitates the subsequent recruitment of late-stage mechanisms, such as adaptive coupling and postzygotic isolation, to complete the speciation process. Applying nonlinear dynamics theories to model sympatric speciation yields novel insights that may help address theoretical objections to sympatric speciation and establish it as a prevalent mode of speciation in nature.

S30-P04-Computer Simulations and Analyses of Coupling Among Reproductive Barriers in Late-Stage Sympatric Speciation

Lauren Gieck

Evolution Research Inc.

The mechanisms of sympatric speciation remain a challenge in evolutionary biology. Closely related species often exhibit multiple barriers that generate strong and irreversible reproductive isolation (RI). Theorists suggest that early-stage mechanisms establishing initial RI differ from late-stage mechanisms that couple additional barriers. We previously developed a two-allele mating-bias model to show how initial premating RI can arise under disruptive ecological selection. Here, we extend this model to examine how pre- and post-mating barriers can couple with initial barriers to achieve robust RI. We used computer simulations to analyze the invasion and coupling dynamics of various barrier mechanisms. Early-stage, adaptive premating barriers, driven by maladaptive hybrid loss, effectively establish initial RI and couple with other barriers but are easily reversed. In contrast, late-stage barriers often rely on reduced gene flow for successful invasion and coupling. Mutations reducing hybrid viability are underdominant and require hitchhiking with advantageous barriers to invade. Chromosomal inversions, an adaptive late-stage mechanism, can combine barrier properties into a supergene and gain a net fitness advantage to invade, resulting in strong, less reversible RI. Late-stage, nonadaptive Bateson–Dobzhansky–Muller barriers evolve slowly but confer the least reversible RI. Our findings reveal a positive feedback loop in which early-stage barriers facilitate the establishment of late-stage barriers, and late-stage barriers strengthen and secure early-stage barriers. This feedback progressively reinforces RI until it becomes irreversible. By examining the properties and invasion dynamics of various barrier mechanisms, this study proposes a comprehensive process of how barriers emerge and couple in sympatric speciation.

S30-P05-Phylogenetic and taxonomic classification of the Microtus (Terricola) grafi based on samples from the Balkan Peninsula using ancient mtDNA

Gabriela Damentka

University of Warsaw, Centre of New Technologies, Warsaw, Poland

Within the genus Microtus (Arvicolinae, Rodentia), we can distinguish the subgenus Terricola. It includes more than 13 species, whose distribution extends between the Atlantic, Caspian, Baltic, and Mediterranean Seas. The most widespread species of the Terricola subgenus is the European pine vole [M. (T.) subterraneus]. Owing to the low morphological variation within the genus Microtus, unambiguous systematic classification can be a problem. In 1992, in the Late Pleistocene materials from the Bacho Kiro cave (Bulgaria), a distinct species of Microtus (Terricola) grafi was identified and described. Samples taken from the site showed the original morphological features of the first molar (M1) (Brunet-Lecomte et al., 1992).

To resolve the taxonomic position of M. (T.) grafi, we obtained mitochondrial DNA sequences from 19 ancient and 8 modern specimens from the Balkans. We reconstructed the phylogeny based on mitochondrial DNA cytochrome b and included all the Terricola species. Samples identified by morphology as M. (T.) grafi were located on the branch formed by the modern specimens of M. (T.) subterraneus. We suggest that individuals previously classified as M. (T.) grafi should be classified within the morphological variation of M. (T.) subterraneus species.

S30-P06-Meta-analysis of antibiotic resistance genes in aquatic environments

Wei Wu

The University of Southern Mississippi

As antibiotic resistance genes (ARGs) in aquatic environments continue to rise globally, impacting water quality and public health, numerous studies have documented ARG concentrations, with some exploring their potential drivers. However, systematic and quantitative reviews linking ARGs to both anthropogenic and environmental factors remain limited. To address this gap, we conducted a global meta-analysis of ARG concentrations using Bayesian inference to investigate climatic and socio-economic factors as potential drivers. Our findings indicate that local-scale climatic variables play a more significant role in predicting ARG concentrations than country-scale socio-economic variables. Specifically, ARG concentrations increased with precipitation, and we identified an optimal temperature at which ARG levels peaked. These insights are crucial for developing location-specific regulatory policies aimed at controlling antibiotic use and mitigating the risks associated with antibiotic resistance.

S30-P07-Investigating Neuropeptide Reception in Mnemiopsis leidyi: A Bioinformatics Approach to GPCR Identification

Aidan Pugh

University of Bristol

Understanding of neuropeptide function in Ctenophora remains limited. Previous research has identified and mapped the expression of putative neuropeptides in Mnemiopsis leidyi, however little is known about the receptors these molecules bind to. G protein-coupled receptors (GPCRs) are known to bind neuropeptides and offer a promising approach to investigate neuropeptide reception mechanisms in Mnemiopsis. The conserved structure and established role in signalling of GPCRs across animals makes them more informative compared to the short, fast evolving and harder to track neuropeptides with regards to evolutionary history. However, identifying GPCR function in Ctenophora is challenging due to the uniqueness of this phylum, the limitations of homology-based tools like BLAST and the scarcity of high-quality genomes. We address this challenge using phylogenetic analyses, protein domain searches and hidden Markov models (HMMs), to identify candidate neuropeptide GPCRs. We use FoldTree to help identify neuropeptide GPCRs in Ctenophora using 3D protein models based on known invertebrate neuropeptide receptors. AlphaFold docking simulations are used to predict ligand-receptor binding of the neuropeptide GPCRs identified. In the future, we plan to experimentally validate tissue-specific expression of candidate GPCR genes in Mnemiopsis. By exploring ligand-receptor mechanisms, we will gain insight into the origins of the nervous system and address evolutionary questions about how neuronal components emerged.

S30-P08-Sperm competition intensifies purifying selection on spermatogenesis-relevant genes in primates

Vasili Pankratov

Aarhus University

Primates include more than 500 species that vary vastly in many aspects of their behavior, including mating patterns. Mating systems characterised by a single female mating with multiple males in a short period of time, namely polygynandry and polyandry, are associated with high levels of sperm competition. Sperm competition, in turn, is thought to result in strong selective pressures on traits that maximize sperm production, the classical example being the relative testes size. However, less is understood about how sperm competition affects molecular evolution.

Here, we address this question using data on genetic diversity within species generated by the Primate Genome Diversity Project. For this, we focused on ~520 genes associated with ‘spermatogenesis’ GO term. We then calculated the ratio of nucleotide diversity at nonsynonymous and synonymous sites (piN/piS) in the concatenated sequence of those genes in each species. We adjusted it for the piN/piS in a concatenated sequence of ~19,000 genes. Next, we tested if the interspecies differences in this value can be explained in a model with a) the mating system as a binary variable (‘polygynandrous’ vs ‘other’; 155 species) or b) the relative testes mass (52 species) as predictors. The effects of the polygynandrous mating system in model a and of relative testes size in model b are -0.015 (95%CI: -0.023; -0.007) and -0.007 (95%CI: -0.014; -0.0014) suggesting lower piN/piS and hence stronger purifying selection on spermatogenesis-relevant genes in species with more intense sperm competition.

S30-P09-Evolutionary dynamics under near neutrality: Codon family-specific reversals of natural selection in the Drosophila melanogaster lineage

Haruka Yamashita

National Institute of Genetics

Nearly neutral evolutionary dynamics may be prevalent in genome evolution but the functional effects of weakly selected mutations generally lie outside the range of direct measurement. Synonymous codon usage bias is well-suited for population genetic inference; under major codon preference, translationally superior "major" codons confer fitness benefits relative to their less efficiently and/or accurately decoded synonymous counterparts. The fitness benefits of major codons are likely vanishingly small, but common selective forces among genes justify pooling of mutations with predicted fitness effects to confer statistical power to detect footprints of directional forces. We employed publicly available genome-scale data for African populations of Drosophila melanogaster and Drosophila simulans. In D. simulans, minor to major codon mutations segregate at higher frequencies within populations than major to minor changes. In contrast, in D. melanogaster, codon family-specific polymorphism patterns reveal a reduced efficacy of natural selection in most synonymous families, but surprisingly, support reversals of favored states in the four codon families encoded by NAY. These codon families show accelerated synonymous fixations in favor of NAT codons deeper in the gene tree, within the ancestral D. melanogaster lineage. This study demonstrates how fine-scale sequence pattern analyses can detect weak fitness effects which may not be detectable by experimental approaches and reveal ongoing genome-wide adaptation. The correspondence between reversed selection and wobble position queuosine modification in NAY cognate tRNAs is intriguing, but the phenotypic basis of reversals remains unclear.

S30-P10-Genomic evolution of ocular horn polymorphism in desert vipers

Theo Busschau

New York University Abu Dhabi

The functional and evolutionary significance of horns in snakes is poorly understood. Using a comparative phylogenetic approach, we showed that horns are evolutionarily labile with frequent fluctuations in the extent of character states as lineages diversified. This trend is further supported by the presence of several polymorphic species where horns are not always present. To understand the genomic basis of horns, we use the desert vipers (Cerastes sp.) as our primary model system. The genus comprises three species, of which Cerastes cerastes and C. gasperettii are polymorphic while C. vipera is considered hornless. Genome-wide association on C. gasperettii show that horn polymorphism is significantly associated with a single ≈20kb genomic region among genes associated with head phenotypes in other organisms. Multiple species- and phenotype-specific structural variants are observed in this region, where hornless phenotypes are linked to a tandem duplication in a conserved non-coding element. The evolutionary history of this region supports the independent origin of horn polymorphism through independent de novo mutations across species. We extended these analyses to the ecologically similar Bitis caudalis species complex in which horns are mostly present but occasionally absent. Here, hornless B. caudalis individuals are concentrated in two geographically proximate populations with one population showing evidence of hybridization with the hornless Bitis peringueyi. Long-read data and scans for signatures of selection further shed light on the evolutionary mechanisms underpinning horn polymorphism in these genera.