Modeling and Theory in Population Biology

Poster Session Abstracts

Tuesday June 2, 2026

2:30 – 3:30 pm

Rethinking admixture mapping in terms of marginal coalescent trees

Junjian (Janis) Liu, University of Southern California

Admixture mapping uses local ancestry to identify trait-associated genetic regions. Admixture mapping was adopted enthusiastically in the early 2000s because it required fewer markers than GWAS given the large extent of admixture linkage disequilibrium in recently admixed populations. There has been speculation that admixture mapping is poised for a return as multi-ancestry biobank resources become more common. However, admixture mapping requires assignment of haplotypes to discrete source populations. This conception of discrete populations has been argued to be a poor representation of much human population structure. Here, we re-cast admixture mapping in a continuous form by framing it in terms of variation in local relatedness. This conception does not require postulated discrete admixture groups. We show that a recent method for performing trait mapping using estimated marginal coalescent trees captures the local-ancestry associations that power admixture-mapping signals while also capturing loci that are more easily detected by GWAS than by admixture mapping. Thus, tree-based locus mapping has the potential to be a powerful overarching framework in the era of multi-ancestry studies of complex traits, one that captures the benefits of both admixture mapping and GWAS without requiring discrete ancestry groupings.

Transmission bottlenecks as emergent stochastic filtering: a branching-process framework for viral lineage loss

Wenjing Zhang, Texas Tech University

Transmission bottlenecks describe the sharp reduction in viral genetic diversity during host-to-host spread. They are a central determinant of viral adaptive potential. Although bottleneck sizes are routinely estimated from sequencing data, the mechanistic basis for this diversity loss remains unclear: is it imposed by a small number of transmitted virions, or does it emerge dynamically from stochastic extinction during early within-host expansion?

We develop a branching-process framework that links early viral dynamics to lineage survival during the establishment phase of infection. Using barcoded influenza data from guinea-pig

transmission experiments (Holmes et al., PLOS Biology 2024), we parameterize the model through the observed early exponential growth rate and compute the probability generating function G(x,t) via a delay differential equation.

We find that even under supercritical growth, stochastic extinction eliminates ∼75% of founder lineages within one day. However, for the majority of recipient animals, this growth-only extinction is insufficient to explain the observed diversity collapse, which is effective lineage counts, L (by Shannon Diversity Index) drop by 100- to 500-fold between the first two days of detectable infection. To account for this gap, we introduce an effective per-virion filter parameter D and show that the empirical contraction ratio satisfies L^_2/L^_1≈1-G(1-D, Δt). Solving for D across recipients yields a median D^≈0.01, indicating that recipient-side processes, such as immune clearance, remove roughly 99% of virions beyond what stochastic growth dynamics already account for.

This framework provides a mechanistic bridge between within-host dynamics and population-level bottleneck estimates, and offers quantitative guidance for designing in vitro serial-transfer experiments that recapitulate in vivo bottleneck strengths.

Ref: Holmes, Katie E., Lucas M. Ferreri, Baptiste Elie, Ketaki Ganti, Chung-Young Lee, David VanInsberghe, and Anice C. Lowen. "Viral expansion after transfer is a primary driver of influenza A virus transmission bottlenecks." PLoS Biology 23, no. 9 (2025): e3003352.

Species-driven microhabitat modifications influence coexistence and productivity under resource competition

Ching-Lin Huang, University of Minnesota, Twin Cities

Species not only respond to habitat conditions but also actively modify them. This modification can influence neighboring species, creating indirect interactions important for shaping community structure. For example, in arid ecosystems, shrubs and trees buffer temperature extremes and increase soil moisture, alleviating stress and boosting understory growth and diversity. Although the role of species-driven habitat modifications has long been recognized, a key challenge remains: we lack a general mechanistic framework to explain when and how such habitat modifications influence community composition, diversity, and productivity.

Here, I develop a theoretical framework that extends Tilman’s resource competition model by introducing a new state variable: the physical environment. Consumers not only exploit resources but also modify the environment. The altered environment, in turn, feeds back to affect the rates of resource supply, mortality, conversion efficiency, and consumption. I hypothesize that such environmental modification can qualitatively and quantitatively influence competitive outcomes.

In a single-resource system, preliminary results show that such environmental modification can equalize species' R* values, i.e., the minimum resource level required to sustain a population, thereby promoting coexistence and enhancing community productivity. Furthermore, the nonlinear feedback introduced by environmental modification may invalidate the classical resource-ratio hypothesis, requiring stability analysis or simulations to predict competitive outcomes.

This framework provides a general, mechanistic approach to understanding when and how microhabitat modifications influence community structure, offering a guideline for quantifying their effects in empirical systems.

Rare variants drive high variance in human ancestral fitness under a high deleterious mutation rate

Ulises Hernandez, University of Arizona

Identifying genetic variants associated with disease enables the development of new approaches for improving human health. The potential impact of these approaches depends on the number of disease-associated variants segregating within a population. We model such variants as having a deleterious effect in fitness in ancestral human environments. At demographic equilibrium, such variants segregate according to mutations-selection-drift balance models and impose a genetic "load" to individuals. The variation of load within a population depends on the whole genome deleterious mutation Ud and the selection coefficient of deleterious mutations sh. Here we revise human estimates of the distribution of fitness effects (DFE) and Ud to estimate ancestral human variation in load. We model the whole-genome DFE, as a gamma distribution with its mean bounded by empirical estimates. Human estimates of Ud are restricted to strongly deleterious mutations, which are absent from species divergence data. We derive an analytical expression to incorporate weakly selected mutations into Ud. We find that human Ud>5. Using our DFE and Ud, we find that the difference in fitness between two typical ancestral humans was bigger than 20%. We propose that this high variation in load confounds health studies. We also find that most of the variation in load comes from ultra rare variants at frequencies between 0.0001 and 0.01. Our results suggest that estimates of disease genetic variance based on whole genome sequence could resolve the 'missing heritability' problem. We also extended our model to vertebrate-like species and found that high variation in load is widespread across populations with low population sizes. This indicates that our results are also relevant to historically threatened species.

Maintaining diversity in structured populations

David Brewster, Harvard University

Evolution—either by genetic reproduction or by learning—occurs in populations. The structure of a population affects the time scale and outcome of evolutionary processes. The propensity of populations to maintain diversity is of great interest in evolutionary biology, ecology, and social science. Here, we calculate for how long various population structures can maintain diversity under neutral evolution. In this setting, diversity is lost by random drift. We give precise results for a large variety of structures. We find that some structures have higher-order polynomial or even superexponential timescales for maintaining diversity. For realistic population sizes of thousands or millions of individuals, those structures can maintain diversity for times that exceed the lifetime of a universe. Therefore, they protect diversity "forever."

Latitudinal dependence of stability trends in Cenozoic marine plankton

Maike Morrison, Santa Fe Institute

The temporal stability and spatial heterogeneity of global marine ecosystems under changing climates reveal how biodiversity persists or collapses. However, the deep-time evolution of these phenomena remains poorly understood due to limitations of both data and statistical methodologies. We use FAVA, an FST-based Assessment of Variability across vectors of relative Abundances, to reconstruct the stability landscape of pelagic plankton from the Cretaceous-Paleogene extinction to the present. We find that the Cenozoic was a period of punctuated volatility with high turnover post-mass extinction and during the Late Neogene cooling. Our spatially resolved analysis revealed latitude-dependent trends: equatorial regions stabilized over time, whereas polar communities, especially in the Southern Ocean, destabilized. These findings reveal the temporal stability dynamics of an important marine microorganism across geologic timescales, with interesting implications for models of community diversity and stability over time.

The Evolution of Phenotypic Heterogeneity in the Bacterial Flagellar Network: Biophysical Constraints and Mechanistic Epistasis

Murat Tugrul, Humboldt University of Berlin

Classical approaches in evolutionary theory often rely on additive and deterministic fitness models. However, recent work suggests that both epistatic interactions and non-genetic

heterogeneity play a critical role in determining the response to selection. In microbial populations, for instance, evolutionary rescue under antimicrobial stress is increasingly linked to persistent phenotypes regulated by complex metabolic and gene regulatory networks. Understanding how molecular biophysical constraints, network rewiring, and stochasticity dictate these evolutionary trajectories is therefore essential.

In my poster, I present a theoretical framework bridging biophysical models of regulation with a mutation-selection-drift evolutionary process. Using the three-tiered hierarchical flagellar network of Salmonella enterica as a model system, we utilize coupled stochastic differential equations to characterize how transcription factor promoter binding kinetics, specificity, and cooperativity dictate cell-to-cell variance. Using a generalized 3-gene network template, we demonstrate how biophysical constraints and stochasticity drive the evolution of bimodal "switches" for bet-hedging and determine the resulting network architectures.

By treating phenotypic states (i.e., the number of flagella) as coordinates in a high-dimensional probability space, we quantify how this non-genetic heterogeneity expands the temporal window for evolutionary rescue under environmental stress. Our results demonstrate that a non-motile reservoir is a stable evolutionary outcome and provide a mechanistic explanation for the divergent rewiring and architectures observed in Salmonella and E. coli. We conclude by discussing how this integration of biophysics and population genetics provides a quantitative framework for predicting the accessibility of fitness peaks and the predictability of resistance evolution.

Same trait, different variants: How demography and selection jointly shape polygenic trait architecture

Prothama Manna, Clemson University

Populations that disperse into new environments often experience highly complex demographic histories that dramatically reshape patterns of genetic diversity, including at loci that underlie complex trait variation. Frequently, range expansion also involves shifts in phenotypic optima, creating novel selection pressures and driving adaptation in polygenic traits. In this context, fundamental assumptions of many traditional population genetics methods are violated, making them ill-suited to disentangling the joint effects of demography and selection on complex trait architecture. To address this gap, we use forward-time simulations implemented in SLiM 5 to model the evolution of a polygenic trait in two recently diverged populations under no selection, stabilizing selection, or directional selection, and either a constant population size or a bottleneck followed by expansion. Our models explore a range of parameters that broadly reflect human population history, including varying degrees of gene flow. This framework allows us to track the full set of causal variants over time and to decompose additive genetic variance across the

allele-frequency spectrum under different evolutionary scenarios. Broadly, we find that stabilizing selection concentrates additive variance in rare, large-effect variants, while directional selection redistributes variance to intermediate and common alleles, especially in populations that retain more ancestral variation. Following divergence, additive variance becomes dominated by population-specific variants that accumulate independently within each lineage. As a result, differences in evolutionary history limit how well genetic predictions generalize across populations, even when populations evolve toward the same phenotypic optimum. This limitation not only reflects allele frequency divergence at shared variants but also a progressive shift in genetic architecture, as private variants account for an increasing fraction of additive genetic variance. Our findings contribute to a broader understanding of how multiple evolutionary processes jointly shape polygenic trait architecture and provide a process-based neutral model for inference and trait prediction across genetically diverse populations.

Effects of Spatial Structure on Selective Sweep Inference

Aalhad Bhatt, Emory University

Many natural populations are spread out over 2-dimensional space, with individuals interacting locally. Sweep detection methods are effectively hitchhiking detection methods. They rely on being able to find sections of the genome that have markedly lower levels of diversity in comparison to neutrality. The sizes of these regions and the exact patterns that they show (like the SFS) are determined by parameters like the location of the sweeping locus and the strength of selection that we do not know a priori. Sweep detection methods estimate these parameters using theoretical expectations of how a selective sweep looks like. However, these expectations are based on models of well-mixed populations.
We use PySLIM recapitation to simulate sweeps in spatially extended populations to show that even modest amounts of spatial structure, such as would commonly be found in real populations, can change the signature of selective sweeps. Sweep times in spatially structured populations grow linearly in the spatial extent instead of logarithmically in the population size, which can substantially slow sweeps down. Slower sweeps leave weaker signatures. These changes can act as confounding factors for inferring the existence, location, and strength of the sweep.

We quantify the effects of this change on algorithms for sweep-finding. We establish parameter regimes for when spatial structure can be safely neglected in sweep inference and when it must be taken into account. We also discuss the potential reinterpretations of conclusions about the strength and frequency of past sweeps in natural populations.

Visualization of genomic variation and admixture through graph embeddings and Fermat distance.

María Inés Fariello, Universidad ORT Uruguay/Universidad de la República

One of the challenges in population genetics data analysis is high dimensionality. Dimensionality reduction via principal components analysis (PCA) is commonly used for genetic data visualization and to infer population structure. However, PCA primarily captures global linear trends, while local genetic variation often appears in higher components. As a result, standard two-dimensional projections may fail to represent key structure, particularly in admixed populations.

Non-linear approaches such as t-distributed Stochastic Neighbor Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP) emphasize local neighborhoods but typically distort global geometry, making distances between clusters difficult to interpret.

We propose a graph-based framework for genomic data visualization and analysis. Individuals are represented as nodes in a weighted graph, where edge weights are computed using Fermat distance, a density-sensitive geodesic metric that captures intrinsic structure in the data. To obtain low dimensional embeddings, we explore several graph-based representations, including node2vec, spectral embeddings, and force-directed (spring) layouts, each providing complementary perspectives on population structure.

Constructing the graph using Fermat distance yields visualizations that preserve both local and global patterns, enhancing the separation of small or isolated populations and, crucially, improving the representation of admixed individuals such as those in the Uruguayan population. Beyond visualization, the graph structure enables extended analyses, including the extraction of ancestry proportions via community detection. This makes the approach a flexible and powerful tool for exploring genetic diversity and admixture.

Quantifying turnover in microbial communities using the traveling salesperson problem

Chloe Shiff, Stanford University

Many microbiome studies yield time-series data with species composition proportions at each of a series of points in time. The temporal turnover in species composition can provide important information about ecological dynamics. In this study, we present a novel statistic to quantify the amount of turnover present in time-series data that records species compositions. Our proposed statistic, Turnover Across Compositional Orderings (TACO), is equal to the sum of pairwise distances between composition vectors that correspond to adjacent timepoints, normalized by the range that the sum can take across all possible temporal reorderings of the vectors. Because the

problems of finding the minimum and maximum distances over all possible reorderings are analogous to traveling salesperson problems, to obtain these distances, we use established computational algorithms for such problems. We demonstrate TACO in two applications. First, we use TACO to understand the longitudinal dynamics of barcoded yeast strains in fluctuating experimental environments. Second, we apply TACO to time-series data of human microbiomes that have been perturbed by antibiotics. TACO provides numerical values that are comparable across many settings, and it can be applied to data arranged in ways other than temporally, such as data arranged in a single spatial dimension.

Leveraging linkage disequilibrium variation to refine functionally informed estimates of trait heritability

Hannah Snell, Brown University

Among the complexities of understanding the heritability of diseases in humans, the variation of linkage disequilibrium (LD) across the human genome and between populations remains underexplored. Previous methods, such as LD Score Regression (LDSC, Bulik-Sullivan et al., 2015) and its functionally stratified version (S-LDSC, Finucane et al., 2015), use summary statistics from genome-wide association studies and LD to estimate heritability enrichment in various diseases. Both methods, however, collapse information from the LD matrix, so that only the amount of LD is modeled per SNP, thereby missing other LD-related patterns in the full LD matrix that may explain variation in disease heritability (such as SNP correlations with many weak independent signals versus one strong signal). As an extension of these methods, LD Eigenvalue Regression (LDER, Song et al., 2022) uses eigenvalue decomposition of the LD matrix to represent LD across spatial genomic regions. Building on these approaches, we present a conceptual advance to LDER that incorporates functional information to refine heritability estimates. We model LD structure using a latent representation derived from a graph convolutional neural network (GCNN). We train our framework on simulations derived from real genotype and functional information and validate these in the S-LDSC model. We are applying our framework to quantitative traits and molecular biomarkers represented in the Genotype Tissue Expression Project (GTEx). In the future, we will assess its performance on more diverse genetic backgrounds and developmental insights at the single-cell level.

When sexual selection meets drift: the coevolution of traits and preferences is decoupled from overall trait-preference genetic correlations during Fisherian sexual selection

Kuangyi Xu, University of Toronto

Fisher's process is central to sexual selection theory, whereby mate choice generates a genetic correlation between male traits and female preferences, driving their coevolution through positive feedback. However, empirical studies often fail to detect strong trait–preference genetic correlations, leaving the effectiveness of Fisher's process debated. One overlooked factor is that trait–preference genetic correlations can be influenced by interactions between selection and genetic drift, whereas previous models have assumed infinitely large populations. Using population genetic models, I show that interactions between Fisherian sexual selection and genetic drift increase trait–preference genetic correlations when the male trait is rare, but decrease them as the trait becomes common, while generally facilitating the evolution of female preferences. As a result, in finite populations, the temporal average of trait–preference genetic correlations is often close to zero or even negative, but stronger average correlations do not predict greater coevolution of either male traits or female preferences. These results suggest that genetic correlation may be an unreliable indicator of the operation of Fisher's process, offering a potential resolution to this dilemma in sexual selection theory.

Conditional Models of the Ancestral Recombination Graph

Jinmin Li, Cornell University

Conditional models that describe the probability of sampling new sequences given existing data are fundamental in population genetics. Their applications range from the Li-Stephens model for haplotypes to the "threading" algorithm used in the Ancestral Recombination Graph (ARG) inference. Despite its importance, the conditional model for ARGs has not been rigorously characterized, leading to ambiguity and misuse. Here we present a principled derivation of conditional ARG models, quantify existing biases, and provide a possible correction form. Through theoretical analysis, we demonstrate the connection between properties of the underlying ARG model and the well-definedness of its conditional process. We show in particular that the sequentially Markovian coalescent (SMC) and SMC' models are incompatible with a conditional framework. We further studied the projection bias introduced by this inconsistency in SMC and SMC'. Building on this formulation, we extend the conditional ARG model to a more general form, providing clear rules of conditioning within the process. Our results demonstrated the complexity and subtlety of conditional models of ARG. We expect our formalization and generalization of the conditional model to inspire the development of inference methods and downstream applications of ARG

Increasing group size can cause catastrophic collapse of cooperation

Rokeya Rahman, University of Kentucky

In many biological and social systems, cooperation depends on collective actions that generate benefits only when a minimal number of individuals coordinate to contribute. These interactions are often modeled using threshold public goods games, in which public goods are produced only if participation exceeds a critical threshold. Cooperative groups vary greatly in size across species, with some very large groups allowing sophisticated types of cooperation, such as division of labor. But how increasing group size influences the evolutionary stability of cooperation remains poorly understood. In this study, we investigate how increasing group size affects the evolutionary dynamics of cooperation, especially when cooperation depends on reaching a certain threshold. Previous work has shown that cooperation can collapse due to increasing costs (Peña et al. 2014). Our model reveals that a similar collapse can arise even when costs and benefits remain fixed. Specifically, increasing group size alone can alter the stability of cooperative equilibria and can cause a saddle-node bifurcation that eliminates the stable cooperative equilibrium states. As group size exceeds a critical value, cooperation can disappear abruptly, leading to an abrupt transition to full defection. These findings emphasize how demographic or ecological changes that lead to larger group sizes could disrupt cooperation systems. Our findings provide a theoretical understanding of how cooperation breaks down in biological and social systems, especially where collective benefits depend on coordinated effort.