BaMBA 12

Agenda

08:00-08:45 Register; coffee and pastries

08:45-09:00 Welcome

09:00-09:50 Lacra Bintu

10:00-10:50 Sean Collins

11:00-11:30 Coffee break, poster viewing

11:30-12:20 BinYu

12:30-02:30 Lunch/coffee/poster viewing

02:30-03:20 Dexter Hadley

03:30-04:20 Andrew Fire

04:30-05:45 Poster viewing

05:45-06:00 Poster removal

Speakers

Lacra Bintu

The effects of chromatin on the dynamics of gene expression: measurements and models

In mammalian cells the state of chromatin highly dynamic and tightly liked to gene expression. It is essential to understand the rules that connect these two dynamics processes if we want to build a predictive model of gene regulation. For instance, how fast can chromatin regulators affect gene expression, how long do their effects last as epigenetic memory, and how far along the chromatin fiber do these effects spread?

In order to quantitatively address these questions, we engineered a set of mammalian cell lines that allow us to precisely control the time of recruitment and release of chromatin regulators at a fluorescent reporter gene. We recruit a set of chromatin regulators associated with various chromatin modifications: histone methylation, histone acetylation, and DNA methylation. We follow the changes in gene expression in

single cells over time using time-lapse microscopy and flow cytometry. For all chromatin regulators studied, their effects can be described as stochastic transitions among three gene states: active, reversibly silent, and irreversibly silent. Mathematically, the only difference between chromatin regulators comes from the values of the stochastic transition rates among these three gene states.

Using a simple Monte Carlo simulation of chromatin modifications dynamics at a nucleosome array, we show that the stochastic nature of gene silencing and activation could arise from spreading of epigenetic modifications along the chromatin fiber. In order to experimentally test this spreading model, we use two reporter genes separated by different spacers, and target the upstream gene with different chromatin regulators. Silencing initiated at the upstream gene can rapidly spread to the downstream gene across different insulator elements and across 5000bp of spacer DNA. Epigenetic memory and reactivation of the two genes is also highly correlated at the single-cell level, suggesting that active gene expression states can also spread.

These results set the basis for a unified mathematical framework of how chromatin regulators operate to control gene expression and epigenetic memory in individual cells over time.

Sean Collins

Understanding a cell's molecular steering wheel

In many contexts, cells need to interpret chemically encoded spatial cues from their environment to survive or carry their physiological role. An extreme example is the chemotaxis of immune cells. These cells sense small differences in the concentration of chemical attractants and use this information to direct their movement. Furthermore, they respond to and prioritize between multiple competing attractant signals, allowing finely tuned regulation of their recruitment and dispersal during immune responses. Such regulation is critical for eliminating infections while avoiding inflammatory diseases. While it is known that the chemotaxis behaviors are driven by a family of G-protein coupled receptors, how signals are processed and integrated downstream of the receptors remains unclear. We are using quantitative live-cell imaging, including new strategies combining precise light-based control of signaling inputs and simultaneous measurement of downstream signaling outputs, to directly interrogate signal processing. Our results provide insights into how the signaling is wired to balance stable cell polarity and persistent movement with directional sensing, and how differences in signaling dynamics may mediate attractant prioritization.

Andrew Fire

Wise, benevolent, and selfish RNAs

Bin Yu

Interactive Random Forests (iRF) and signed iRF (s-iRF)

for predictive and stable high-order interactions

Building on random forests (RFs) and random intersection trees (RITs) and through extensive biologically inspired simulations, we developed the iterative random forest algorithm (iRF) and its enhanced version signed iRF (s-iRF). iRF trains a feature-weighted ensemble of decision trees to detect predictive, stable, and high-order interactions with the same order of computational cost as the RF. s-iRF improves upon iRF by adding signs to the interactions based on new null metrics. We demonstrate the utility of iR and s-iRF for high order interaction discovery in genomics problem including one that concerns enhancer activity in the early Drosophila embryo.

Dexter Hadley

From Bits to Bedside: Translating Large-Scale Routine Clinical Datasets into Precision Mammography

We demonstrate how to use deep learning (DL) approaches to translate big data from routine clinical care into medical innovation that directly improves routine clinical care. Typically, large healthcare institutions have sufficient quantities of clinical data to facilitate precision medicine through a DL paradigm. However, this clinical data is hardly translated into direct clinical innovation because computer algorithms cannot readily ingest or reason over it. Using routine mammography screening data for breast cancer as an example, we first downloaded twenty years of free text pathology reports and used natural language processing to infer cancer outcomes for individual patients. We then labeled close to one million mammographic views of breast imaging with our inferred pathology outcomes. Finally, we trained convolutional neural network DL algorithms to directly predict pathology outcomes from breast imaging. With our approach, we demonstrate how to leverage DL to realize precision oncology and significantly improve the interpretation of routine screening mammography for millions of women using routine clinical big data.

Posters

Balanced rate and spiking model for oculomotor integrator
Stella Dong and Mark Goldman
Applied math/neuroscience at UC Davis
The oculomotor neurons integrate a transient saccade into a persistent activity. Since the persistent activity is proportional to the intended position of the eyes in their orbit, integrator neurons are said to

maintain a “memory of eye position”. Mathematically, persistent activity is described in terms of “dynamical attractor” which means any self-sustained and stable state of a dynamical system. The question is that whether such attractor networks can be realized in the brain, biologically constrained models of persistent activity were needed? In this project, we focus on the network of the oculomotor integrator, and we will provide the necessary constraints on the synaptic strengths and on the time constants of the synapses in order for the network to have the persistent activity. Furthermore, the persistent activity is believed to be generated by feedback dynamics in the network. The oculomotor integrator has both positive feedback and negative feedback. In this project, we will introduce the negative-derivative form feedback which occurs when the network has same synaptic strength but offset in time. It will be shown that the negative-derivative feedback counteracts drift in persistent activity, and the negative-derivative feedback networks are robust to common perturbations such as the change or synaptic strength and loss of neurons. We will represent two models: a rate model and a spiking model. Both resulting derivative-feedback oculomotor models are robust to many commonly studied perturbations. In the rate model, we will show the firing rates remain more stable in the negative-derivative feedback models than those are in the positive feedback network against the change of the synaptic strength. In the spiking models, we will approximate the horizontal eye positions using the firing rates. We will show that even though some neurons are silent in the network, the rest neurons will compensate to spike in order to track the eye positions. However, there is a boundary for the optimal compensation for the loss of neurons. When too many neurons are silent, the balance between excitation and inhibition is disrupted, and the eye position representation is lost.

Ribosomes, Traffic Jams and Phase Transitions

Dan Erdmann-Pham, Khanh Dao Duc and Yun Song
UC Berkeley, computer science division

The translation of mRNA into protein is a fundamental biological process, mediated by the flow of ribosomes. This process is well modeled by an interacting particle system which generalizes the Totally Asymmetric Simple Exclusion Process (TASEP). While the TASEP and its variants have been studied for the past several decades through simulations and mean field approximations, a general analytic solution has remained challenging to obtain. By analyzing the so-called hydrodynamic limit, we here obtain exact closed-form expressions for stationary currents and particle densities that agree well with Monte Carlo simulations. In addition, we provide a complete characterization of phase transitions in the system. Surprisingly, phase boundaries depend on only four parameters: the particle size, and the first, last and minimum particle jump rates. Relating these theoretical results to translation, we formulate four design principles that detail how to tune these parameters to optimize translation efficiency in terms of protein production rate and resource usage. We then analyze ribosome profiling data of S. cerevisiae and demonstrate that its translation system is generally efficient, consistent with the design principles we found. We discuss implications of our findings on evolutionary constraints and codon usage bias.

Evolutionary deep scanning reveals prevalent trade-offs in yeast adaptation

Yuping Li; Dmitri Petrov; Gavin Sherlock
Department of Biology, Stanford

The idea of trade-offs has been a fundamental premise of biological and evolutionary thought. In general, if a trade-off exists between the two traits, the two traits will not be simultaneously maximized. Thus, trade-offs constrain the evolutionarily accessible space and generate a nontrivial evolutionary “front”. As individuals can improve both performances without trade-offs until they reach the front, one needs to investigate the evolutionary front to demonstrate the existence of trade-offs. However, to identify the evolutionary front, one needs to assay a large number of adaptive mutants and to precisely measure their performances in several key traits, which, until recently, was challenging to achieve. Here we used the

yeast growth cycle as a model, which includes a sequence of growth phases, to study trade-offs between performances in these phases. Using experimental evolution, we isolated thousands of adaptive clones with improved performances in different phases. A precise performance quantification for these adaptive clones in each growth phase enabled us to depict the front of the evolutionarily accessible space. With the evolutionary front, we observed prevalent trade-offs between performances in the yeast growth cycle. In addition, genome-wide sequencing of these adaptive clones uncovered the genetic basis underlying observed trade-offs.

Kinetoplast DNA--The Chain-Mail Hockey Puck

Ryan Polischuk, Dr. Rajiv Singh, Dr. Javier Arsuaga
Department of Physics, UC Davis

Trypanosomes are a group of disease-causing parasites that happen to possess a fascinating form of mitochondrial DNA. Called “kinetoplast DNA” (kDNA), this subset of the organism’s overall genetic information encodes for important energy-production activities, and—uniquely in trypanosomes—is housed in a single highly structured and topologically ordered object, which is shaped like a hockey puck. The kDNA comprising this structure consists mainly in small (kb-scale) circular molecules (plasmids) called minicircles, which connect pairwise via single topological links (Hopf links) to form an object reminiscent of a sheet of medieval chain mail. Mathematical biologists have developed theoretical models for key topological properties of kDNA minicircles, such as the average number of minicircles that link to a given circle—the mean valence (Chen et al. 1995, Diao et al. 2012). Here, we report our work to extend Diao and colleagues’ minicircle model by focusing on the orientations of the minicircles. By analogy with magnetic systems, we have chosen an energy function called the 2-D anisotropic Lebwohl-Lasher model, whose low energy states mirror the parallel arrangement of kDNA minicircles observed under electron microscopy. We have solved this system analytically, via an approximation method called mean-field theory (MFT), and exactly, using Monte-Carlo (MC) importance sampling. We get a close agreement between our predicted crossover-type transition in the MFT and our MC results for the rank-2 quadrupole order parameter, Q_0, which measures out-of-plane orientational order in the minicircle system.

Linking probability of minicircles in kinetoplast DNA (kDNA)

Aparna Komarla, Maxime Pouokam, Michelle Flanner, Javier Arsuaga
Department of Mathematics, UC Davis

Trypanosomes are single celled protozoans with a unique mitochondrial DNA structure called kinetoplast DNA (kDNA). kDNA is partitioned into minicircles and maxicircles that are topologically linked forming a network. This network is often regarded as the most structurally complex mitochondrial DNA in nature (Lukes, 2002). Different biophysical models can be used to simulate minicircles in kDNA as polymer chains. Three polymer chain models that have used are: geometrical circles, freely jointed chains (FJC) and worm-like chains (WLC). In this work we aim at determining whether rigid circles are a valid representation of minicircles. In particular, we test whether the linking probability of two geometrical minicircles of radius R is comparable to the linking probability of two chains (FJC or WLC) with radius of gyration, R. We develop a function that compares the results from these models and provides insight into the different simulation models.

Computational Discovery of Heterogeneous Binding Specificity of Transcription Factors

Luis Chumpitaz, Md. Abul Hassan Samee, and Katherine S. Pollard
Gladstone Institutes, Gladstone Institute of Data Science and Biotechnology.

BaMBA 12

Transcription factors (TFs) are regulatory proteins that bind the genomic DNA in a sequence-specific manner. Mutations in these specific sequences are known to be associated with developmental disorders and diseases. Understanding the DNA-binding specificity of TFs is therefore a central goal of computational biology. A fundamental assumption in this realm is that a TF binds to a single “canonical” sequence (typically 10 – 12 bases long) and its minor variants (i.e., different from the canonical sequence in a few bases). In other words, the specificity of a TF is assumed to be homogeneous, and this homogenous specificity is modeled by a formalism called “motifs” that captures the canonical sequence and its variants where the TF binds. However, a few recent studies have shown that a TF may bind to two distinct sequences – implicating the existence of heterogeneous binding specificity of TFs. To understand the existence and extent of this interesting phenomenon, we are analyzing the largest in vitro HT-SELEX (High Throughput Systematic Evolution of Ligands by Exponential Enrichment) library of TF-DNA binding assay. This library has DNA binding specificity data of more than 500 TFs that are expressed in human and/or mouse. State of the art methods to analyze HT-SELEX data are biased toward the known motifs. We therefore developed an unbiased pipeline to quantify the enrichment of 12-mers (short sequences of 12 nucleotides) in these libraries and cluster the 12-mers with varying degrees of flexibility. In our pilot analysis, this pipeline discovered the known motifs for every TF. However, for many TFs, the known motifs cannot explain a considerable fraction of bound 12-mers. Instead of the known motifs, our pipeline discovers secondary motifs for these highly enriched but unexplained sequences. Our preliminary results thus suggest that the phenomenon of heterogeneous binding may be more common than currently assumed. This can significantly improve our current understanding of the mechanisms of TF-DNA binding and shed light on many aspects of TF-DNA binding that we have failed to explain so far under the assumption of homogeneous binding motifs.

Computer simulation of site-specific recombination in simple cubic lattice

Diwen Lu, Michelle Flanner, Allison Moore, Mariel Vazquez
Department of Mathematics, UCDavis

Chromosome unlinking can be accomplished by multiple rounds of site-specific recombination. But we do not have a clear understanding of how site-specific recombinases can identify where to make two cuts and recombine the broken DNA chains to convert a complicated knotted structure into a simpler knot or link, or into an unknot. We here study unknotting and unlinking by site-specific recombination. We generalize an existing mathematical model and use computer simulations to perform local recombination leading to changes in DNA topology.

Wang Landau Algorithm Applied to Knots in Cubic Lattice

Zihao Zhu, Shawn Witte, Michelle Flanner, Javier Arsuaga, Mariel Vazquez
Mathematics Department

We are interested in uniform sampling from the space of self-avoiding polygon in $\mathbb{Z}$^3. Wang Landau is a generalized Markov Chain algorithm that trains to sample uniformly from a space of energy states. Here we use the BFACF algorithm with Wang Landau sampling. The energy states are the number of edges of a lattice polygon. The Wang Landau process consists of two phases: training and sampling. In the training phase, we store two main objects, a histogram and a weight list. First, the weight list is initialized to some small unit and the histogram is empty. Next, we begin sampling using a Metropolis Hastings implementation of the BFACF algorithm. We record the occurrence of the samples in the histogram and increment the weight for that energy state by a small amount simultaneously. After the histogram is almost flat according to the flat check criterion, we reduce the amount we increment the weight list (this is denoted as flatter on), reset the histogram and rerun the process until the weight increment gets appropriately small. Finally, output the weight list and sample by using the weight list.

Traversing the fitness landscape of lung adenocarcinoma

Christopher D McFarland, Zoë N Rogers, Ian P Winters, Wen-Yang Lin, Dmitri A Petrov, and Monte M Winslow
Stanford University, Department of Biology

Cancers exhibit a highly diverse landscape of somatic alterations whose effects on tumor evolution, both alone and in combination, remains largely unknown. To interrogate the evolutionary consequences of multiple tumor genotypes in an autochthonous tumor environment, we developed a method that integrates DNA barcoding with CRISPR/Cas9-mediated genome editing in mice. First, via ultra-deep sequencing of barcoded tumors, we track the size of hundreds of early tumors within a single mouse. Isogenic tumors initiated at the same time within the same mouse quickly diverge in size. Different cancer genotypes exhibited categorically-different tumor size distributions, which suggested alternative modes of cancer growth. Thus, we devised two alternative Markov models of tumor evolution, which we then tested in a time series that tracked the size of thousands of tumors for 2 – 32 weeks. Next, we paired our barcoding technology with CRISPR/Cas9-mediated genome editing to measure the fitness of thirty-one of the most frequent human lung adenocarcinoma genotypes. We observe that most gene losses confer context-dependent growth, i.e. they are adaptive only in the presence (or absence) of other genetic events. Our fitness measurements correlated well with mutational co-occurrence frequencies in human lung adenocarcinoma. Our new tumor-barcoding and CRISPR/Cas9-mediated genome-editing approach measures tumor growth with unprecedented parallelism and precision, reveals tremendous heterogeneity in isogenic tumor growth that informs cancer's mode of evolution, and identifies a rugged fitness landscape of tumor evolution that should help personalize cancer therapeutics.

Statistics of Topological Analysis of Genome in Three Dimension

Javier Arsuaga, Brian Cruz, Sean Burgess, Mark Segal, Mariel Vaquez, Javier Arsuaga
Biostatistics

Recent developments have improved our understanding of the shape of the genome and enabled the determination of three-dimensional (3D) structures based on chromosome conformation capture contact (CCC) data. Theoretical studies suggest that the model structures obtained should be highly knotted due to the confinement. Understanding the ensuing geometry of the obtained structures depends on being able to capture such entanglement in quantitative terms, hence there is a need to develop reliable methods measuring model structures entanglements. Here we propose a novel statistical based approach to quantify the geometrical entanglement between two open chains using closure methods, as those proposed for measuring the entanglement of proteins. This method is applied to diverse contexts: reproducibility of genome, topological simplicity, and others. Our main results identify the Rabl-like configuration as a regulator of topology simplicity of chromatin in budding yeast.

Formation of a meiotic loop array via extrusion in S. Cerevisiae

Stephanie Schalbetter, Geoffrey Fudenberg, John Baxter, Katie Pollard, Matthew Neale
Gladstone Institutes, UCSF

During meiosis, chromosomes undergo extensive conformational changes to both allow recombination and faithfully transmit genetic information to the next generation. An iconic structural feature of meiotic chromosomes is the emergence of an array of loops, emanating from a proteinaceous axis, in prophase I. Over several decades, factors associated with this axis have been identified, including the meiosis-specific cohesin subunit Rec8. Still, how these factors influence chromosome conformation and the mechanism whereby this axis assembles remain at large. Here we combine chromosome conformation analysis by Hi-C with yeast genetics and 3D in silico modelling to dissect the pathway

whereby meiotic chromosomes are assembled. Using Hi-C, we find that meiotic chromosomes not only become linearly compacted, but also gain focal interactions between pairs of rec8 sites, reminiscent of those seen between CTCF sites in mammalian interphase. Synchronizing cells in pachytene, we find these focal interactions are entirely dependent on Rec8 but emerge independently of both homologous recombination and homolog pairing. Using polymer simulations, we show that meiotic chromosome organization is inconsitent with direct bridging models. Instead, we find the patterns of interactions that emerge in meiosis are consistent with the mechanism of loop extrusion limited by boundary elements. Together our results present a likely pathway for the first stage of meiotic chromosome assembly, where loop extrusion mediated by meiotic cohesins becomes halted at sites of active transcription.

Prediction in cancer genomics using topological signatures illustrated with logistic regression

Georgina Gonzalez, Arina Ushakova, Radmila Sazdanovic, and Javier Arsuaga
University of California, Davis

Copy Number Aberrations, gains and losses of genomic regions, are a hallmark of cancer and can be experimentally detected using microarray comparative genomic hybridization (aCGH). In previous works, we have developed a topological based method to analyze aCGH data whose output are regions of the genome where copy number is altered in patients with a predetermined cancer phenotype. We call this method Topological Analysis of array CGH (TAaCGH). Here we combine TAaCGH with traditional statistical or machine classification algorithms to build classifiers using copy number aberrations. We chose logistic regression on two different binary phenotypes related to breast cancer to illustrate this approach. The first case consists of patients with over-expression of the ERBB2 gene. Over-expression of ERBB2 is commonly regulated by a copy number gain. TAaCGH found the region 17q11-q22 associated with the phenotype and using logistic regression we reduced this region to 17q12-q22. The classification rate of this region was 78% of the ERBB2 positive individuals in a validation data set. We also analyzed over-expression in Estrogen Receptor (ER), a second phenotype commonly observed in breast cancer patients and found that the region 5p14.3-12 together with six full arms were associated with the phenotype. Our method identified 4p, 6p and 16q as the strongest predictors correctly classifying 76% of ER positives in our validation data set. Although for this set there was a significant increase in the false positive discovery rate. We suggest that topological and machine learning methods can be combined for prediction of phenotypes using genetic data.

Mathematical Equilibrium Modeling Reveals a Key Role for Negative Superhelicity in Driving R-loop Formation

Robert Stolz, Shaheen Sultana, Craig Benham, Frederic Chedin
MCB, UC Davis

R-loops are three-stranded nucleic acid structures that form during transcription upon reannealing of the nascent RNA to the template DNA strand, creating a DNA:RNA hybrid and causing the non-template strand to loop out in a single- stranded state. Studies by the Chedin lab and others indicate R-loops are prevalent features in the genomes of many organisms including mammals and occur in conserved hotspots. Despite their prevalence, our understanding of the fundamental physico-chemical, thermodynamic, and topological parameters that dictate R-loop formation still lags. We have developed an R-loop equilibrium energetic model that incorporates DNA sequence and DNA superhelicity as well as torsional and junctional energies. This model reveals that R-loops are energetically favorable over a range of DNA sequences and predicted that negative DNA superhelicity plays a critical role in driving R-loop formation. Subsequent biochemical experiments in which R-loop formation was monitored using single- molecule assays verified that prediction. Our work forms the basis for an emerging quantitative understanding of R-loops as topologically
driven nucleic acid structural transitions.

Juicy Details of a Synthesized Chromosome

Diana Sernas, Keith Fraga, Javier Arsuaga
Mathematics Dept, University of California - Santa Cruz

The human genome is two meters long and is condensed inside the cell nucleus that is only six micrometers in diameter. How cells overcome this structural problem remains an open question. The study of genome architecture has benefited from new sequencing tools called chromosome conformation capture (CCC) technologies that can estimate the 3D proximity of regions of the genome by their “contact frequency”. These methods explicitly illustrate how two DNA loci may be several kilobases away but remain proximal in 3D space. Saccharomyces cerevisiae (yeast) provides a model system to investigate nuclear architecture. In particular they are an example of the Rabl configuration, a chromosome organization that clusters centromeres and telomeres in different regions of the nucleus. The Arsuaga group has proposed that the Rabl configuration prevents the entanglement of chromatin fibers. To further test this hypothesis, I am generating 3D reconstructions of yeast genomes that contain large scale rearrangements. To achieve this goal, I am adapting existing software Juicer 1 for the analysis of these data sets. Juicer is the first step of a bioinformatics pipeline meant to processes raw files from Hi-C experiments capturing genome-wide contacts. In this work I present the implementation of Juicer to analyze yeast genomes.

Modeling Longitudinal Changes in Child Vocalization Behavior in infants with Autism Spectrum Disorder

Fahad Paryani, Gordon Ramsay
Emory University School of Medicine Department of Pediatrics and Marcus Autism Center

Current research in Autism Spectrum Disorder (ASD) is often focused on finding biomarkers for risk since early detection has been shown to significantly increases standards of living by providing early treatment for speech development. Since ASD is known to cause speech deficit and sleep disruption identifying these issues are the focus of this project. We acquired the data from day long recordings of over 100 infants through the first year of life. The LENA and in house software counts the number of cries, babbles, and vocal interaction with the caregiver. Using the vocalization data, we created a longitudinal map that displays the trajectory of child vocalization throughout the day and across the first year of life using cubic spline interpolation. The results revealed sever hindrance in vocal development for infants who are at risk for ASD. To assess the sleep deficit, we used the Two-Process Model of sleep and circadian rhythm to link the behavioral element, child vocalization, to the biological. Although this linkage is not confirmed the computational model we developed shows promising result in acting as a biomarker for risk. The linkage can provide large implication on advancing our understanding the biological changes that occur in ASD.

Data Science in Sugarcane Production

Edith Mugehu, Ijeoma Ezika, Jacqueline Scoggins, Daniela Ushizima
Berkeley Institute for Data Sciences

Sugarcane ethanol is an expandable green alternative to crude oil use as it provides a scalable solution to reduce carbon dioxide emissions. According to Jaiswal et al in Nature Climate Change volume 7, pages 788–792 (2017), sugarcane ethanol may offset 86% of CO_2 emissions compared to oil use, but sugarcane production expansion often competes with worldwide demands for food production. Therefore, new strategies in AgTech, which combines crop science, production and improvement, plant biotechnology, and data analytics, have been explored in a collaboration between American and African researchers from Zimbabwe and Nigeria as part of a scientific diplomacy program, sponsored by the U.S. Dept. of State IEE TechWomen. This poster shows preliminary results using data science to better

understand data collected at the Zimbabwe Sugar Association Experiment Station, and present data-driven models regarding production of different sugarcane species.

A mathematical approach to understanding the interplay of natural and sexual selection in speciation
Megha Suswaram, Dr. Justin Yeakel, Dr. Danielle Edwards
Quantitative Systems Biology, University of California, Merced
Many mathematical models have been used to study evolution in spatial systems, including population genetics, quantitative genetics, and individual based models over various time spans. Several empirical tests and mathematical models have explored how speciation occurs, individually by divergent NS and SS. Despite recent interest in how natural selection (NS) and sexual selection (SS) combinatorially contribute to advance speciation, there are still no mathematical models or empirical tests that describe how these two mechanisms could theoretically drive the accumulation of reproductive isolation – from incipient divergence through to complete reproductive isolation and speciation.
This study presents a theoretical model that outlines the action of NS and SS in different strengths and directions to drive speciation. We extend the study of evolutionary dynamics to eco-evolutionary models based on adaptive dynamics. We derive expressions for the strength of selection that apply both in continuous space and time within a single population using a single locus model. Designed for selection on a quantitative trait mean, this presents a way to integrate local directional selection across space and determine whether the trait value will increase or decrease over time. Our preliminary results show that when NS and SS act together in the same direction, even at low strengths, it pushes trait divergences coincident with speciation rapidly as opposed to when these processes act in opposing directions.

Putting Relicanthus in its place: the impact of mixture model choice on phylogenetic reconstruction
Madelyne Xiao, Mercer Brugler, Estefania Rodriguez
Department of Invertebrate Zoology, American Museum of Natural History
First described in 2006, Relicanthus daphneae is a deep-sea coral that lives on the ocean floor near hydrothermal vents in the East Pacific. It was originally classified as an anemone until a phylogenetic analysis in 2014 called this classification into question. The tree resulting from a maximum likelihood analysis for the family of Actiniarians (anemones) placed Relicanthus outside of Actiniaria; a recent analysis of Relicanthus's gene order, however, suggests its membership among the anemones. An ongoing study seeks to relate the choice of mixture model (e.g., maximum likelihood, maximum parsimony, Bayesian inference) to the resulting phylogenetic tree, taking into account the robustness of the data set in question (number of genes, specimens, etc). In particular, we are interested in the impact of mixture model choice on the placement of Relicanthus with respect to the Actiniarians.

Topological Analysis of Correlation Networks
Benjamin Roycraft, Wolfgang Polonik, Dietmar Kueltz
UC Davis Statistics
Several methods are proposed for the analysis of correlation networks. Correlation networks are modeled as weighted undirected graphs. Persistence homology is used to summarize essential structure, and via network peeling effective functional summary statistics are constructed. Techniques from functional data analysis are applied, including functional principal components (fpca). Error estimation is given by the bootstrap. Simulations are provided to establish the discriminating power of the network statistics. Finally, an application is made to the protein interaction networks of stickleback fish.