Computational Analysis of Massively Parallel Sequencing Data to Identify Low Frequency Mutations and Rearrangements
Karthik Karnik
Mentors: David Kwiatkowski and Paul W. Sternberg

Massively parallel sequencing (MPS) is a form of sequencing technology for high-throughput DNA sequencing projects, including the analysis of human genetic variants. In the Kwiatkowski Lab, MPS is used to detect sequence variants and mutations in patients who have the genetic disorder Tuberous Sclerosis Complex, and in several types of tumors and cancers. Using MPS, the output received in the lab is a binary alignment map (BAM) file containing reference genome-aligned sequencing data, as well as a collection of reads which do not align correctly. However, BAM files alone are not feasible for identifying variation found in the input DNA. Rather, they must be interpreted and analyzed by other programs to obtain data that indicates DNA sequence variants. We have developed a robust, user-friendly program written in the Unix command line, Python, and MATLAB that specifically searches the BAM file data for sequence variants across a given gene at a user-specified allele frequency and determines potential sites of large-scale deletions. This program has been tested with various projects and will be used on the lab’s entire collection of BAM files from various genomic and cell-free DNA samples.

Using Network Methods to Identify Sexually Dimorphic Regulatory Pathways in Non-Small Cell Lung Cancer
Tina Wang
Mentors: John Quackenbush and Matt Thomson

Non-small cell lung carcinoma (NSCLC), the most common subtype of lung cancer, has differing natural history, progression, and response to therapy in men and women. To explore whether this sexual dimorphism was due to complex gene regulatory interactions involving sex chromosomes and hormones reflected in the action of disease-associated transcription factors, we used two gene regulatory network inference algorithms, PANDA (Passing Attributes between Networks for Data Assimilation) and LIONESS (Linear Interpolation to Obtain Network Estimates for Single Samples), developed at the Quackenbush Lab at Dana-Farber Cancer Institute and the Harvard T.H. Chan School of Public Health, to analyze squamous cell carcinoma and lung adenocarcinoma data from The Cancer Genome Atlas. We found substantial differences in the overall network topology implicating distinct roles for transcription factor regulatory processes in men and women from comparing gene regulatory networks. Among the most differentially targeting transcription factors between the sexes were ARID3a and FOXL1, which play critical roles in colorectal cancer and ovarian cancer, respectively. We also found many other transcription factors that have not been previously studied in NSCLC but are involved in processes relevant to the disease and its progression. Preliminary results provide potential avenues for further research and clinical drug development.

Tools for Visualization and Analysis of Behavioral and Neuroimaging Data
Ayan Bandyopadhyay
Mentors: Vinod Menon and Dean Mobbs

The objective of this project was to provide tools to supplement the Stanford MET study, which tested the effectiveness of in-person and tablet-based cognitive tutoring in increasing the brain plasticity of children with Mathematical Learning Disorders. I developed a web application to be used by experiment administrators to view the behavioral data of each student enrolled in the study as they complete a set of tablet tasks each week. This web application will be used to ensure that the students are completing their tasks, and to monitor changes in response time and accuracy in real-time. Furthermore I adapted a neuroimaging pipeline to the fMRI data in this study by incorporating a watershed skull-stripping workflow. By testing this adapted pipeline against existing pipelines, I determined the efficacy of watershed skull-stripping in improving fMRI preprocessing.

Investigating MORC-1 Phosphorylation in the C. elegans Germline
Gloria Ha
Mentors: John Kim, Nicita Mehta, and Rob B. Phillips

Endogenous small interfering RNAs (endo-siRNAs) in the nematode Caenorhabditis elegans are deposited from the maternal germline into the embryo. This transgenerational inheritance of endo-siRNAs is critical for both germline maintenance and for heterochromatin formation at target genes. We previously found that morc-1, the C. elegans homolog of the highly conserved Microrchidia family of chromatin-binding proteins, is a crucial link between endo-
siRNAs and multigenerational chromatin organization. MORC-1 is comprised of an ATPase domain and a zinc finger domain. Based on conservation analysis, an “E39A” mutation in the ATPase domain of MORC-1 was found to exhibit much higher levels of phosphorylation than wild type MORC-1. Mass spectrometry yielded two putative phosphorylation sites of MORC-1, both serines. A “WWAA” mutation in the zinc finger domain of MORC-1 has been implicated in preventing recombination. Additionally, co-immunoprecipitation showed that MORC-1 directly interacts with CSR-1, the only *C. elegans* Argonaute that is essential to fertility and embryo viability. The interesting phosphorylation behavior of MORC-1, in addition to its role in recombination and its interactions with a critical Argonaute, led us to further investigate these qualities. We verified two putative sites of MORC-1 phosphorylation through immunoblotting, and sought to identify the kinase that phosphorylate MORC-1 through a targeted screen of 58 germline kinases. Through genetic crosses, we found that knocking down *csr-1* through RNAi rescues the WWAA recombination defect. We also imaged CSR-1 localization in the absence of MORC-1 through immunofluorescence, and investigated the interaction between CSR-1 and MORC-1 in different RNAi backgrounds through co-IP. Our results have identified putative kinases responsible for MORC-1 phosphorylation and further clarified the relationship between MORC-1 and the Argonaute CSR-1.

**Interaction Between Cardiolipin and Two Mitochondrial Proteins, Drp1 and OPA1**

Zikun Zhu  
*Mentor: David C. Chan*

OPA1 and Drp1 are proteins central for mitochondrial fusion and fission, respectively. Cardiolipin, a kind of mitochondrion-specific lipid, plays a key role in mitochondrial fission and fusion. To figure out the binding of these two proteins and cardiolipin, a protein-lipid overlay assay was performed. I have been optimizing the assay by changing different conditions and got positive results using purified full length Drp1 and shortened versions of OPA1. After that, I tested the Q785R mutant of OPA1, which has been shown to be defective for lipid binding. The results showed that its cardiolipin binding ability was not influenced negatively. As a consequence, I have predicted other mutants that may lead to deficiency and purified them. Liposome floatation assay should be done in the future to quantify their binding. In order to figure out which domain of OPA1 is responsible for cardiolipin binding, we noted that OPA1 has a region similar to dynamin’s pleckstrin homology domain which is known to bind lipid. I found that this domain can bind cardiolipin, which triggered our interests to solve its structure. So far, we have got the initial crystals of pH domain and are trying to optimize it.

**Auditory-Visual Crossmodal Interactions in Perception: The Spatial Double Flash Illusion**

Ishani Ganguly  
*Mentor: Shinsuke Shimojo*

Multimodal illusions, such as the double flash experiment initially carried out in the Shimojo Laboratory, allow us to understand how humans prioritize and organize sensory stimuli when perceiving their surroundings. This classic experiment overturned predominating concepts by demonstrating that in some cases audition can be prioritized over vision, and showed that the brain can postdictively alter future perceptions based on already perceived stimuli. My project deals with a recently discovered variation of the original double flash illusion called the spatial double flash illusion. In this experiment, a central visual flash is initially depicted in conjunction with an auditory stimulus that is displaced to one side. Then, a second auditory stimulus displaced to the opposite side generates an illusory visual flash in that direction. My objective was to optimize the robustness of the illusion by noting how altering various factors influence the postdictive perception observed. To achieve this, specific experiments were coded in MATLAB and subsequently run on human subjects. Pilot results suggest that the addition of a static fiducial bar extending into central vision, in addition to the original visual flash set in the periphery, strengthens the illusion by making it easier for subjects to perceive a sense of movement.

**Neural Connectivity Within Individual During Flow Experience**

Shota Yasunaga  
*Mentors: Shinsuke Shimojo and Mohammad Shehata*

Flow is the mental state of operation in which a person performing an activity is fully immersed in a feeling of energized focus, automation, full involvement and enjoyment. Previous studies have shown that appropriate task difficulty is crucial to flow. However, few researches have applied brain imaging techniques. This project investigates the neural connectivity of flow experience with electroencephalography (EEG) within individual. We used music rhythm game as the task. For the flow condition, we used normal game and for a negative control, we used reverse and shuffled music. We used the Adaptive Mixture of Independent Component Analyzers (AMICA) to get independent component of EEG data acquired with 128 channels. We calculated partial directed coherence to those components for the measurement of the neural connectivity with a software called group-SIFT. We observed increase in connectivity around basal area and wide occipital area. Statistical tests are going to be done to obtain the difference in connectivity between conditions.
Peptidases are enzymes that play various roles in developmental processes including cell migration, where errors can have serious consequences such as mental disability, vascular disease, or tumor formation. Our goal is to provide mechanistic support for the function of four conserved peptidases (Mmp2, Tasp1, Tok1, CG9416) using the easily observable *Drosophila* caudal visceral mesoderm (CVM) cell migration as our experimental system. Tasp1 and Tok1 are primarily expressed in migrating CVM cells, while Mmp2 and CG9416 are expressed in several tissues including CVM. Although the human orthologues of these peptidases have been connected to overexpression in cancers and roles in various tissue development, little is known about their specific biological mechanisms. We analyzed tissue-specific RNAi mutants and published genetic knockout mutants of the peptidases via immunohistochemistry against CVM markers, then assayed the RNAi mutants with live imaging movies. So far, we have observed pronounced stalling, asymmetry, and other abnormal migration phenotypes in the Mmp2 mutants specifically, and are waiting on more results to confirm their statistical significance. We hope to see phenotypes in the other mutant peptidase lines, which may ultimately lay the foundation for future treatments of developmental diseases or cancer.