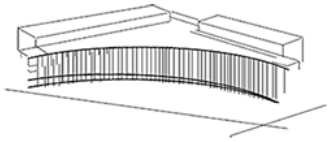
PRINCETON UNIVERSITY

Department of MOLECULAR BIOLOGY

Lewis-Sigler Institute for Integrative Genomics



Summer Undergraduate Research Program Molecular and Quantitative & Computational Biology

Poster Session

Abstracts

2017

Student Name	Advisor	Institution	Title	Page #	Poster #
Kristen Albrecht	Enquist	Princeton University	Pseudorabies Virus Induces Interleukin-6 Specific Immune Response in Mice	7	1
Henry Ando	Shaevitz	Princeton University	Methods for Structural Analysis of Branching Microtubule Nucleation Sites	7	--
Metodi Balev	Ploss	Princeton University	Probing the role of cyclophilin A in hepatitis C virus (HCV) replication and host tropism	8	2
Logan Blaine	Andolfatto	Princeton University	An epistatic basis for the correlated clustering of amino acid substitutions	8	3
Sandra Carpenter	Devenport	Princeton University	Investigating the Role of the Planar Cell Polarity Pathway in Mammalian Epidermal Wound Repair	9	4
Milena Chakraverti-Wuerthwein	Leifer	Princeton University	Neural Dynamics of Food and Hunger in *Caenorhabditis elegans*	9	5
Jessica Chambers	Witten	Princeton University	The temporal dynamics of dopamine in working memory	10	---
Daniel Che	Buschman	Princeton University	Efficacy of an Adaptive Cognitive Prosthetic on Behavior in Awake Mice	10	6
Ariel Chen	Rabinowitz	Princeton University	Gene expression analysis in CCLE database identifies potentially important genes in cancer cell lines	11	7
Sam Chiacchia	Notterman	Princeton University	ARDS is Associated with Genome Wide Methylation Changes and Alterations in DNA Methylation Age in Adults	11	8
Diana Chin	Donia	Princeton University	Antimicrobial-Encoding Biosynthetic Gene Clusters from *Lactobacillus* Strains in the Vaginal Microbiome Mediate Bacterial Community Dynamics	12	9
Mary Kate Davis	Wieschaus	Princeton University	Phenotypic and Genetic Analyses of 2D61, a Unique Embryonic Lethal Mutation in *Drosophila melanogaster* (in the Wingless pathway)	12	10
Kaitlin Demarest	Korennykh	Princeton University	Investigating Translation Inhibition by RNase L	13	11

Student Name	Advisor	Institution	Title	Page #	Poster #
Nazik M. Elmekki	Brynildsen	Princeton University	Investigating the Relationship between Metabolic Activity and Persistence in Growth-Inhibited Populations of *Escherichia coli*	13	12
Cao Fang	Kang	Zhejiang University	Dissecting the Functions of LncRNA75 in Breast Cancer	14	13
Zach Feig	Brody	Princeton University	Conserved Neuronal Firing Sequences in the Hippocampus: An Application of Graph Theory's the Longest Paths Problem	14	--
Michael Fernandez	Ayroles	East Los Angeles College	Using experimental evolution to understand the contribution of the microbiome to host metabolism in *Drosophila melanogaster*	15	14
Thomas Ferrante	Bassler	Princeton University	Identification of Genes Required for the Production of an Autoinducer-2 Mimic in *Saccharomyces cerevisiae*	15	15
Paola Figueroa-Delgado	Gavis	University of Puerto Rico	The role of canonical Wnt pathway in regulating neuron morphology	16	16
John B. Finlay	Schwarzbauer	Princeton University	Nephronectin and Tenascin-C may be key players in mesangial matrix expansion in diabetic nephropathy	16	17
Zachary Flamholz	Gavis	Princeton University	Analysis of Post-Initiation Translational Regulation during *Drosophila* Oogenesis by Ribosome Footprinting	17	18
Courtni Foster	Levine and Troyanskaya	Howard University	Role of Non-Coding Single Nucleotide Variants in Autism Spectrum Disorder	17	19
Nicholas Freeman	Brody	Princeton University	Neuronal Feature Encoding in the Hippocampus CA1 Region during an Evidence Accumulation Task in Mice	18	20
Noah Han	Burdine	Princeton University	Characterizing Human RASopathy Mutations in *Danio Rerio*	18	21
Kristian Harris	Gitai	Howard University	Identifying genes involved in the regulation of the PQS operon in *Pseudomonas aeruginosa*	19	22
Rukia Henry	Enquist	Howard University	Establishing a model to study the Entry and Transport of Pseudorabies Virus Infection in Central Nervous System Neurons	19	23

Student Name	Advisor	Institution	Title	Page #	Poster #
Benjamin Huang	Simons	Princeton University	Modeling and Predicting the Global Gravitational Tide	20	24
Wentao Huang	Kang	Zhejiang University	Searching for Stemness Negative Regulators of Stem Cell Program in Mammary Gland	20	25
Vishank Jain-Sharma	Bialek	Princeton University	Understanding the Paradox in Genetic Regulator Precision	21	26
I'ran James	Andolfatto	Elizabeth City State University	Hunting for SNPs: The genomic distribution of false positive and negative variant calls	21	27
Gabriel Joseph	Shenk	Princeton University	Functional annotation of gene set enrichment analysis generated networks gives an identity to significantly modified pathways and functions	22	28
Benjamin Juarez	Gitai	University of California, Merced	Investigating *Vibrio cholerae*: Analysis of Curvature Proteins	22	29
Katie Kavanaugh	Jonikas	Princeton University	Untitled	23	30
David Kim	Ploss	Princeton University	An Investigation of a Novel Host Determinant of Live-Attenuated Yellow Fever Virus Infection	23	31
Joshua Kim	Toettcher	Princeton University	Photoswitchable Control of Protein-Protein Interactions: Engineering Light-Sensitive DARPins	24	32
Rachel Kim	Yang	Princeton University	Toward Multiple Site-Specific Labelling on Nonribosomal Peptide Synthetases	24	33
Robert Benjamin Lainer	Verma	Princeton University	Automatic Speech Recognition with Kaldi ASR	25	34
Joshua Latham	Staggs	Princeton University	Reconstituting a Cryostat Refrigerator for Testing of CMB Instruments	25	--
Ju Young Lee	Burdine	Princeton University	Design of a transgenic ERK2-GFP zebrafish line utilizing CRISPR/Cas9 technology	26	35
Miles Lee	Wang	Princeton University	Design and Optimization of Spectrofluorometer	26	36
Daniel Liu	Kang	Princeton University	Hysteresis control of TGFβ-induced epithelial-mesenchymal transition generates a distinct program with metastatic abilities	27	37

Student Name	Advisor	Institution	Title	Page #	Poster #
Jennifer Liu	Rabinowitz	Rutgers University-New Brunswick	One-Carbon Metabolism in Small Cell Lung Cancer	27	38
Olivia Long	De Leon	Princeton University	Charge State Readout of Nitrogen Vacancy Centers in Diamond	28	39
Yuting Lu	Murphy	Zhejiang University	Redacted	28	40
Yikai Luo	Murphy	Zhejiang University	Cobalt Chloride-mediated Hypoxia affects reproduction, fat metabolism and aging in *Caenorhabditis elegans*	29	41
Luisa Mercado	Murphy	Webster University	Identifying Biomarkers of Aging in Human Blood	29	42
Jivahn Moradian	Chaney and Shevilakova	Princeton University	Visualizing Earth System Models: From Communes to Continents	30	43
Megan Ostrowski	Shaevitz	Princeton University	Quantitative Analysis of Aggressive Fruit Fly Behavior	30	44
Clayton Otter	Cristea	Princeton University	Characterization of the host-virus mitochondrial interactome during HCMV infection	31	45
Catherina Pan	Cristea	Princeton University	Interferon-inducible protein 16 (IFI16) in immune signaling during herpesvirus infection	31	46
Nana Park	Kang	Princeton University	Characterizing the role of DACT1 in Wnt signaling and breast cancer bone metastasis	32	47
Emily Pauls	Burdine	Princeton University	Exploring mechanisms of spinal curvature in zebrafish models of Idiopathic Scoliosis	32	48
Tony Phan	Muir	Princeton University	Untitled	33	49
Dennis Perez-Lopez	Wühr	University of Puerto Rico, Mayagüez Campus	Optimization of lysis conditions for quantitative proteomics in *Escherichia coli*	33	50
Katherine Pizano	Witten	Princeton University	The Role of Dopaminergic Terminal Activity in NAc and DMS during Aversive Fear Extinction	34	51
Aparna Raghu	Petry	Princeton University	Investigating the Role of Multiple Importin-Binding Sites in the Phase Transition of TPX2, a Critical Spindle Assembly Factor	34	52
Gabriel Ramirez-Arellano	Hughson	Southern Virginia University	Characterizing the Interaction between Sec39/Use1 in the Dsl1 Multisubunit Tethering Complex	35	53

Student Name	Advisor	Institution	Title	Page #	Poster #
Rachel Reed	Murphy	Princeton University	Identification of Reproductive Span Regulating Genes in *C. Elegans* Using Mos1 *Insertional Mutagensis*	35	54
Anastasia Repouliou	Brangwynne	Princeton University	How stress can phase out nucleocytoplasmic shuttling	36	55
Jorge Reyes	Rabinowitz	Princeton University	Metabolic Flux Modeling of One-Carbon Metabolism	36	56
Elizabeth Reznik	Notterman	Princeton University	Investigating the Role of DNA Methylation in Stress-Attributed Telomere Attrition	37	57
India Rogers-Shepp	Murphy	Princeton University	Creation of a single worm *Mos1* transposon screen for *Caenorhabditis elegans*	37	58
Jenelys Ruiz-Ortiz	Murphy	University of Puerto Rico, Rio Piedras Campus	Effect of lifestyle factors and aging on differential gene expression patterns in human blood.	38	59
Christopher Russo	Wingreen	Princeton University	Translational burst size and gene expression noise in E. coli: Finding an optimum	38	--
Peyton Smith	Schedl	Princeton University	Niemann-Pick Disease (NPC) family proteins in Embryonic Germ Cell Migration	39	60
Katherine Stiefel	Kocher	Princeton University	Down-Regulation of syntaxin 1A Gene in Bombus impatiens Using dsRNA	39	61
Shiye Su	Staggs	Princeton University	Feedhorn Characterisation and Sensitivity Calculation of the AdvACT Bolometer Array	40	--
Camille Sullivan	Bassler	Princeton University	Determination of PqsE active sites in *Pseudomonas aeruginosa*	40	62
Xiaoting Sun	Leifer	Princeton University	Optogenetic disruption of neural activity in *C. elegans*	41	63
Caroline Taber	Shvartsman	Princeton University	Analysis of Cell Growth and Cycle Regulation in *D. melanogaster* Nurse Cells	42	64
Tanya Tafolla	Wingreen and Donia	University of California, Merced	Creating Hidden Markov Models for Transport Protein Classification in Metagenomics Data	43	65

Student Name	Advisor	Institution	Title	Page #	Poster #
Mark-Avery Tamakloe	Levine	Princeton University	An Investigation of Distal Enhancer-Enhancer Communication for the *brinker* Locus in Drosophila	43	66
Alice Tang	Singh	Princeton University	The Enhancer Effect: Investigating the Role of Non-coding Mutations in Cancer	44	67
Kimberly Tang	Silhavy	Princeton University	Understanding outer membrane protein assembly in *E. coli* through suppression of ß-signal mutant LamB^{G439D}	44	68
Sarah Tian	Schwarzbauer	Princeton University	Investigating the Relationship of Fibronectin and LOX Family Proteins during Matrix Assembly	45	69
Ava Torjani	Notterman	Princeton University	Establishing a Mechanism for Stress---Induced Telomere Shortening in Human Foreskin Fibroblasts	45	70
Gabriel Toneatti Vercelli	Gregor	Princeton University	Modeling the fluorescence signal of transcribed loci from live imaging	46	71
Aidan Waugh	Schedl	Princeton University	*roX1* and *roX2* Chromatin Entry Sites: Dosage Compensation and Boundary Elements	46	72
Maria Wissler	Silhavy	Princeton University	Investigation of Phospholipid Trafficking to the Outer Membrane in *Escherichia coli*	47	73
Stevie Yang	Gitai	Princeton University	Effects of Environmental Conditions on Vibrio cholerae Curvature and Potential Implications	47	74

Pseudorabies Virus Induces Interleukin-6 Specific Immune Response in Mice

Kristen Albrecht
Kathlyn Laval
Advisor: Lynn Enquist

Pseudorabies virus (PRV), a neurotropic alphaherpesvirus, establishes asymptomatic infection in swine. In non-native hosts, including mice, PRV causes mad itch and death. Studies with a flank inoculation model found that mice infected with PRV develop pruritus in the flank by 40 hours post-inoculation and die approximately 72 h. Pruritus does not cause death, as has been shown by a study observing the same time of death for anesthetized and non-anesthetized mice. We hypothesize that death results from a lethal uncontrolled immune response initiated by PRV replication in the peripheral nervous system and spread to the spinal cord. This study aims to characterize the currently undefined immune response. Mice were inoculated in the hind footpad, blood was collected every 24 hours, and tissues were collected from animals sacrificed at 72 h. Among 11 pro-inflammatory cytokines, only IL-6 was significantly increased in sera of infected animals at 72 h. Immunohistochemistry staining detected PRV in the foot, bladder, kidney, heart, and sacral through thoracic segments of the spinal cord but not in the cervical spine, brain, lungs, liver, spleen, or pancreas. Cells producing IL-6 were detected only in the inoculated foot and heart. Therefore, we conclude that PRV spreads from the inoculation site to the spinal cord through peripheral neurons, and replication initiates an IL-6 specific inflammatory response. A time course study of IL-6 secretion in these tissues collected every 12 hours will be performed using ELISA and q-PCR to establish whether secretion correlates with the development of clinical signs, particularly pruritus.

This research was made possible by the generous support of the Susan W. and James C. Blair '61 P87 Endowed Senior Thesis Fund.

Methods for Structural Analysis of Branching Microtubule Nucleation Sites

Henry Ando
Sagar Setru, Matthias Koch, Guannan Liu
Advisor: Joshua Shaevitz

Microtubules (MTs) are protein _laments that are involved with structure and motion in all eukaryotic cells and some prokaryotes. In particular, they form the mitotic spindle which pulls apart the chromosome during cell division. It was recently proved that microtubules can nucleate from the sides of existing microtubules in the presence of -tubulin and augmin, a process believed to be critical for rapidly increasing MT density in the mitotic spindle during cell division. Little is currently known about the molecular players involved with, and the exact mechanism of, branching MT nucleation. We believe that a clearer understanding of the as-yet-unknown quaternary structure of the branching MT nucleation site will further our understanding of the nucleation mechanism. To this end, we began developing atomic force microscopy techniques to acquire molecular-resolution images of branching MT nucleation sites. We concluded that the AFM in our laboratory is not precise enough to resolve the shapes of the different proteins involved in the branching MT nucleation sites. However, our AFM should be able to probe the mechanical properties (e.g. stiffness) of this juncture by pushing the daughter MT in various directions, and so we have begun writing programs to perform this measurement. Additionally, we hope to gain access to a more powerful AFM in another lab on campus to acquire nanometer-resolution images of the branching MT nucleation site.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Probing the role of cyclophilin A in hepatitis C virus (HCV) replication and host tropism

Metodi Balev
Jenna M. Gaska, Brigitte Heller
Advisor: Alexander Ploss

Hepatitis C virus (HCV) is a leading cause of liver-related mortality. Although tissue culture systems allow for the study of HCV *in vitro*, *in vivo* studies have been restricted by the poorly understood host tropism of the virus and lack of an easily-accessible, externally-valid animal model. Mice genetically engineered to express critical human HCV entry factors are permissive to infection, but viral replication remains low. One potential barrier could be the incompatibility of murine orthologs of vital host factors with components of the HCV replication machinery. Recent evidence suggests that even minor deviations in the primary sequence of the peptidylprolyl isomerase cyclophilin A (CypA), a host factor crucial for HCV replication, can have noticeable effects on HCV replication efficiency in human-derived hepatocytes. In this study, we seek to explore the contribution of interspecies deviations in CypA protein sequence to both the restricted host tropism of HCV and the poor replication efficiency of HCV in murine cells. We utilized human hepatoma cells expressing an shRNA directed against endogenous CypA (Huh7.5-shRNA CypA) transduced with CypA orthologs from a variety of species and assessed these orthologs' respective ability to rescue HCV infection. Given our preliminary findings that murine CypA (mCypA) cannot facilitate efficient viral replication and conversion of certain mCypA residues to the analogous human sequence confers an increased ability to promote viral replication, we will perform AP-MS on FLAG-tagged human CypA (hCypA) and mCypA to determine the differential protein interactions during HCV infection that may affect mCypA's inability to promote viral replication.

This research was made possible by the generous support of the Susan W. James C. Blair '61 P87 Endowed Senior Thesis Fund.

An epistatic basis for the correlated clustering of amino acid substitutions

Logan Blaine
Andrew Taverner
Advisor: Peter Andolfatto

The extent to which protein evolution is predictable remains an important open question in evolutionary biology, necessitating a better understanding of the forces of evolutionary constraint. Previous work by our lab has shown that amino acid substitutions, sites within in a protein that differ between species, tend to cluster within the DNA coding sequence (CDS) of a protein. This clustering may in part be explained by epistasis, whereby a prior substitution at a given site in a protein influences the effect that substitutions at neighboring sites have on the protein's function. An *in silico* model of purifying selection on a set of proteins was built to test whether epistasis between substitutions is sufficient to cause clustering. In this simulation, substitutions are assigned a higher fitness if they do not substantially effect the protein's free folding energy, as calculated by the software FoldX. Preliminary results from the simulation reveal that sequential substitutions occur slightly closer together under our model than those introduced alone, implicating epistasis as causative force for clustering. However, the magnitude of this clustering is far lower than the clustering observed in actual genomic data, suggesting that positive selection, rather than purifying selection, may produce most of the observed clustering. To complement the *in silico* study, I also have begun a computational analysis of *Drosophila* genomes to identify clusters of substitutions that cannot be explained by site-specific differences in evolutionary constraint. Through this, it will be possible to quantify the prevalence of epistatic clusters of substitutions across the genome.

This research was made possible by the generous support of the Susan W. and James C. Blair '61 P87 Endowed Senior Thesis Fund.

Investigating the Role of the Planar Cell Polarity Pathway in Mammalian Epidermal Wound Repair

Sandra Carpenter
Sara Stahley
Advisor: Danelle Devenport

Efficient repair of wounds to the skin is a fundamental capability of an organism's survival. Achieved by migrating keratinocytes, reepithelialization is an inherently polarized process involving coordinated and directional collective cell movements. These highly orchestrated movements remain incompletely understood. Recent work indicates a potential link between the Planar Cell Polarity (PCP) pathway and a murine cytoskeletal effector necessary for epidermal repair. PCP regulates a diverse range of embryonic morphogenetic processes that are recapitulated in wound closure. It is conceivable that the wounded state triggers upstream polarity cues that then exert an influence on the complex coordination of collective cell movements during healing. To define a more precise role for PCP in wound repair, this research seeks to learn how epithelial polarity changes at the molecular level during and after wounding as well as to describe how disrupted PCP affects cell behavior. Using *in vitro* and *in vivo* wounding assays on murine skin, this work will specifically characterize patterns of protein localization in response to wounding, predicting that PCP components will reorient at the wound margin to demarcate leader cells from more distant follower cells. The results from these experiments may provide direction for the future study of cell polarity as an informer and regulator of healing with emphasis on dissecting wound-specific associations between PCP proteins and their downstream cytoskeleton effectors. A better understanding of the molecular mechanisms underlying wound repair is valuable for medical advances in surgery, tissue repair and regeneration, and certain cancer pathways.

This research was made possible by the generous support of the Susan W. and James C. Blair '61 P87 Endowed Senior Thesis Fund.

Neural Dynamics of Food and Hunger in Caenorhabditis elegans

Milena Chakraverti-Wuerthwein
Advisor: Andrew Leifer

Caenorhabditis elegans can be used as a model organism to study brain activity because they have a rudimentary brain and fully mapped out neuron wiring. Using fluorescent markers and a specialized imaging set up, neuronal activity of the entire brain can be recorded in real-time. Whole-brain imaging techniques assist in quantitatively studying the neural activity corresponding to specific behaviors, and up until now have been used primarily with short-term behaviors, such as turning or reverse locomotion. We hope to expand on this and explore the brain chemistry and neural activity of long-term behaviors, such as feeding. *C. elegans* exhibit a local search behavior when they are taken off a feeding plate and placed onto an unseeded plate. This local behavior has been found to be affected by the conditions that the worm previously fed in. Applying experimental methods presented in previous literature, we hope to add to the understanding of feeding behaviors by observing the neural dynamics with resolution of differences in single-neuron activity for the entire brain.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

The temporal dynamics of dopamine in working memory

Jessica Chambers
Advisor: Ilana Witten

Working memory, the basis of learning and cognition, is a system for temporary storage and management of novel information that directs subsequent behavior. Several lines of research suggest that dopamine plays a crucial role in working memory function and may improve its capacity. However, the specific aspects of working memory that the dopaminergic system support have yet to be fully explored, as previous experimental approaches have been hindered by a lack of temporal resolution and cell-type specificity. Optogenetics presents a promising solution to these difficulties, as it enables direct manipulation of neuronal activity at a rapid timescale that matches dopaminergic neuronal activity. Our unpublished preliminary data shows that SNc dopaminergic neurons are causal to working memory, as optogenetic inhibition of these neurons throughout a working memory task impairs performance accuracy. These results motivate my project: focused investigation of neuronal activity during specific epochs of the task, each of which maps to a phase of working memory. The experiments proposed here will unveil fascinating and central aspects of behavior, while furthering our understanding of related disease processes. Furthermore, to our knowledge, these experiments are among the first to link temporal dynamics of the dopaminergic system to the regulation of working memory.

This research was made possible by the generous support of the MOL/QCB Summer Undergraduate Research Program

Efficacy of an Adaptive Cognitive Prosthetic on Behavior in Awake Mice

Daniel Che
Sina Tafazoli
Advisor: Timothy Buschman

Neurological injuries and conditions such as stroke or traumatic brain injury can leave patients with lasting neurological and neuropsychiatric disorders. While these abnormalities in brain function can last a lifetime if untreated, previous treatments have utilized cognitive prosthetics focusing on prerecorded behavior. We test the efficacy of an Adaptive Cognitive Prosthetic (ACP) on behavior in awake head-fixed mice. The ACP consist of an adaptive learning algorithm and multi-side electrical microstimulation and functioned by recording activity from normal brain areas, transforming the signal to mimic lost function, and then stimulating another healthy brain region. The ACP was an acute implant through a craniotomy opened in the primary visual cortex (V1). The visual pathway was chosen due to its structure being relatively well understood and its location in an accessible area of the brain for craniotomy and electrode insertion. Water restricted mice are trained in a go/no-go visual discrimination task consisting to two stimuli in three phases, consisting of a habituation phase to the training setup, an intermediate conditioning phase where a water reward was given upon licking during the correct stimulus, and a final phase in which mice received a timeout and air puff upon licking during the incorrect stimulus. Preliminary training results show that mice were able to reach a stable 100% hit rate and a relatively lower false-alarm rate allowing for discrimination between the two visual stimuli. The ACP has been shown to be able to produce complex cortical firing patterns in anesthetized mice and preliminary results show that the ACP is capable of both recording and stimulating neurons in awake mice.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Gene expression analysis in CCLE database identifies potentially important genes in cancer cell lines

Ariel Chen
Zhaoyue Zhang, Bassel Ghaddar, Lukas Tanner
Advisor: Joshua Rabinowitz

The Cancer Cell Line Encyclopedia (CCLE) provides mRNA expression levels of several thousand genes in many cancer cell lines. Analysis of this database provides us the ability to link expression of certain genes with specific cancer types, which could aid in discovering new anticancer targets. To achieve this goal, we developed a computational method to categorize the cell lines by cancer type and find correlations between cancer type and gene expression levels, identifying potentially important genes whose expression is most skewed towards one or a few specific types of cancer. One gene of interest, ALDH1L2 (aldehyde dehydrogenase 1 family, member L2) codes for a mitochondrial enzyme that is active in one-carbon metabolism and in production of the pivotal cellular reducing equivalent, NADPH. In our analysis, the ALDH1L2 gene expression is correlated with a highly malignant plasma cell cancer type, multiple myeloma (MM). Among all cancer cell lines, MM is the most highly expressed category in ALDH1L2, while this gene is in the top 1% of overexpressed genes in multiple myeloma, suggesting the significance of the pairing. We then perform in vitro characterization of ALDH1L2 in multiple myeloma cell lines by creating ALDH1L2 knockout cells with lentiviral CRISPR and confirming the batch knockout with flow cytometry. Future work takes two directions - first, obtaining a complete list of gene/cancer type combinations of interest; second, further characterization of ALDH1L2 knockout cells, including identification of suppression methods in order to develop possible therapies for multiple myeloma.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

ARDS is Associated with Genome Wide Methylation Changes and Alterations in DNA Methylation Age in Adults

Sam Chiacchia
Lisa Schneper, Iulia Kotenko, Gaby Soto
Advisor: Daniel Notterman

Acute respiratory distress syndrome is a lethal condition of acute bilateral lung disease associated with trauma, sepsis, and shock that occurs as a result of fluid build up in the alveoli. Prior work in the field suggests that ARDS and its pathophysiology may be mediated by epigenetically controlled factors, such as DNA methylation, and that the syndrome itself may incur genome wide DNA methylation changes. In this investigation, we compare DNA methylation data derived from blood collected from 39 adults with ARDS, 75 ICU controls, and 30 healthy individuals made available by Szilagyi et al. Probe wise analysis revealed genome wide methylation changes across 24 chromosomes in the promoters of 13,475 genes, while regional analysis of differential methylation revealed 67 significant regions across 21 chromosomes. GO analysis revealed significant enrichment for CpG sites associated with response to wounding and stress. Taken together, these results demonstrate that the human methylome is significantly altered in response to ARDS, suggesting a potential role in epigenetic response to the stress of acute lung injury.

This research was made possible by the generous support of the Center for Health and Well Being, MOL/QCB SURP, and Office of Undergraduate Research.

Antimicrobial-Encoding Biosynthetic Gene Clusters from *Lactobacillus* Strains in the Vaginal Microbiome Mediate Bacterial Community Dynamics

Diana Chin
Jared Balaich
Advisor: Mohamed Donia

The lactic acid-producing genus *Lactobacillus* dominates in vaginal bacterial communities in healthy women and is thought to play a key role in protecting against pathogens. Consistent with this idea, bacterial vaginosis – involving reduced proportions of vaginal *Lactobacillus* species – is correlated with adverse health outcomes such as preterm birth and HIV transmission. Specific *Lactobacillus* species within the vaginal microbiome are associated with different outcomes, indicating factors besides lactic acid that influence host-pathogen response. The mechanisms behind the community dynamics of the vaginal microbiome remain poorly characterized. We hypothesized that strains of *Lactobacilli* may secrete antimicrobial small molecules that confer a competitive advantage over both commensals and pathogens. To investigate, we queried the genomes of 38 *Lactobacillus* strains for antibiotic-encoding gene clusters found in bacteria and identified sixteen distinct biosynthetic gene clusters (BGCs). Using a twofold approach, we cloned candidate BGCs and expressed them heterologously in *Bacillus subtilis* and also expressed them endogenously in the native *Lactobacillus* strains. We tested both types of cell extracts for activity against a panel of pathogens and found evidence for activity against Gram-negative bacteria from two *L. jensenii* strains carrying different BGCs. We are currently testing additional strains using disc diffusion bioassays. Together, our preliminary results suggest that antimicrobials encoded by BGCs found in certain *Lactobacillus* strains could play an important role in mediating the dynamics of the vaginal bacterial community by inhibiting certain pathogens.

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund.

Phenotypic and Genetic Analyses of 2D61, a Unique Embryonic Lethal Mutation in *Drosophila Melanogaster* (in the Wingless pathway)

Mary Kate Davis
Advisor: Eric Wieschaus

Drosophila melanogaster develop utilizing a set of signaling pathways that are highly conserved in most species. The basic structure of these pathways is well understood, but their regulation during development is not known. The 2D61 mutant stock produces abnormal embryos with a phenotype resembling a hypomorphic mutation in the wnt/wingless pathway. Understanding the genetic basis of this stock could provide new insight into the modulation of these pathways. I compared the cuticle from 2D61 with known developmental mutants in various signaling pathways. 2D61 has similar denticle belt malformations as a hypomorphic armadillo/betacatenin mutant. My preliminary analyses using classical genetic mapping shows that there are two lethal mutations in 2D61, localized at 2-57 and 2-64. The first lethal mutation causes both the denticle belt and head defects, while the second lethal mutation shows a normal phenotype. Because there are no known components of the wingless pathway in this area of the genome, a novel gene is likely perturbed by this mutation. Another mutation is available from the nation stock center, 3263, that is reported to show a similar phenotype but in my analysis complements 2D61. Their head phenotypes are similar, but the denticle belt malformations in 3263 are similar to an arrow mutant, another component of the wingless pathway. The two lethal mutations were isolated and along with 2D61 will be tested further by utilizing scanning electron microscopy for brown head defect analysis, movies for temporal developmental divergence, and genetic localizations using deletion and deficiency complementation testing. Ultimately this project will determine the role of this novel gene in the regulation of such a highly conserved pathway.

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund.

Investigating Translation Inhibition by RNase L

Kaitlin Demarest
Sneha Rath
Advisor: Alexei Korennykh

In response to cell stress (more specifically, double-stranded RNA), the endoribonuclease RNase L is activated. When active, RNase L cleaves noncoding RNAs and inhibits translation, although the complete mechanism of this inhibition has yet to be determined. By measuring translation activity after activation of wild type and cleavage-deficient RNase L, we have demonstrated that its cleavage ability is necessary for translation block. We have also shown that RNase L is capable of blocking translation in *Escherichia coli*. Thus, we hypothesize that RNase L is targeting a part of translation conserved in humans and bacteria. We aim to deepen our understanding of RNase L's method of inhibition in a number of ways. First, we plan to study its effect on tRNA in bacteria by analyzing their cleavage and acylation status. Second, in order for the stress response to be successfully carried out, we believe that certain transcripts must be able to evade this block. We hope to identify these transcripts, as this could clarify how the block is evaded and provide insight into how RNase L establishes it. To do this, we plan to use puromycin immunoprecipitation to isolate nascent polypeptides and identify them via mass spectrometry. These findings will expand our knowledge of the cell in times of stress, and could ultimately lead to the development of cancer therapies that target translation, as there are a growing number of anticancer drugs that shut down protein synthesis.

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund.

Investigating the Relationship between Metabolic Activity and Persistence in Growth-Inhibited Populations of *Escherichia coli*

Nazik M. Elmekki
Grant M. Gelderman
Advisor: Mark P. Brynildsen

Resistance to antibiotic therapies is an increasing global health problem, and nonheritable bacterial persistence poses an independent set of challenges. Persisters are bacterial cells that can tolerate antibiotic treatment independent of resistance genes and naturally comprise a small fraction of bacterial populations. These persistent cells are thought to be responsible for relapse infections, with nongrowing populations being particularly difficult to kill. There have previously been several investigations into persister metabolism, but with a focus on metabolic activity prior to antibiotic treatment. For this study, we used the fluorescent dye Redox Sensor Green (RSG) to characterize the metabolic activity of stationary phase *Escherichia coli* populations during ofloxacin treatment, followed by fluorescence-activated cell sorting (FACS) to assess the dependence of persistence on metabolic activity during treatment. To survive ofloxacin treatment, DNA repair machinery must be active in persisters, and thus we hypothesize that persisters require higher rates of cellular metabolism. Preliminary experiments have established that RSG does not affect the ability of ofloxacin to kill stationary phase cells and that an RSG signal accumulates over the course of a 6 hour incubation following antibiotic treatment. The findings of this study will provide insight to persister cellular metabolism and their response to antibiotic stress, paving the way for the development of new drug therapies that are more effective against persistent populations. Future studies may investigate the metabolic response of persisters when exposed to additional antibiotics, explore different nutrient and environment conditions, or more closely examine the metabolic processes during treatment that enable survival.

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund.

Dissecting the Functions of LncRNA75 in Breast Cancer

Cao Fang
Fenfang Chen
Advisor: Yibin Kang

TGF-β (transforming growth factor beta) signaling is first reported to inhibit cell growth, whereas further discoveries identified its important role in promoting cancer progression. Therefore, details about the dual role of TGF- β signaling in cancer progression needs to be examined. Long non-coding RNAs (LncRNAs), defined as non-coding RNAs longer than 200 nucleotides, have been reported to regulate various biological processes such as gene transcription, translation and protein interactions. Here we seek to find out LncRNAs that are involved in TGF- β signaling regulation by conducting a LncRNA microarray screening in different epithelial cells which are sensitive to TGF- stimulation. Together with qPCR confirmation, we found out that LncRNA75 was upregulated upon TGF-β stimulation. LncRNA75 expression was downregulated by high cell density which is similar to Hippo/YAP signaling. In addition, knock down of either LncRNA75 or YAP can inhibit cell proliferation, indicating LncRNA75 may participate in Hippo/YAP pathway. We further investigated the relationships between LncRNA75 and Hippo/YAP signaling and found that knockdown of LncRNA75 reduced the expression level of YAP-targeted genes. Future research should involve detailed mechanisms about their reciprocal regulation and influences on cell growth. Our research revealing the interactions between LncRNA75 and YAP in breast cancer may offer a possible explanation for the dual role of TGF- in cancer progression.

This research was made possible by the generous support of the Department of Molecular Biology.

Conserved Neuronal Firing Sequences in the Hippocampus:

An Application of Graph Theory's the Longest Paths Problem

Zach Feig
Edward Nieh
Advisor: Carlos Brody

The mouse hippocampus is an important brain region implicated in navigation, decision-making and reward seeking behavior. Activity in this region can be measured via two-photon calcium imaging. Recent advances in the imaging techniques, surgical procedures, mouse restraint and virtual reality allow both wider field and higher resolution hippocampal imaging while the mouse completes behaviors. These advances allow us to resolve 10 times as many Regions of Interest (ROIs) as previous comparable studies, however, the ability to analyze this new data lags behind our capability to collect it. Stereotyped patterns of neuronal activation are notoriously hard to find in hippocampal brain data. Hippocampal data cannot be generalized between subjects, or between imaging sessions because this region of the brain has a tendency to remap. Even in a single recording session, complete consensus sequences are unlikely to be found in single trials because ROIs are activated probabilistically.
This problem can be reduced to the graph theory problem longest paths. In order to identify long sequences we composed a graph consisting of all ordered pairs (doublets) of ROIs occurring in all trials. Edge weights were equal to the number of trials in which the doublet appeared and the directionality was the activation. This graph should contain long, weighty, and acyclic paths – in other words, viable consensus sequences. Unfortunately the longest paths problem is NP-hard, meaning it cannot be solved in polynomial time with any known algorithm. We suggest several case specific heuristics that make this problem more manageable and allow discovery of possible paths.

This research was made possible by the generous support of the Lane '73 Senior Thesis Fund.

Using experimental evolution to understand the contribution of the microbiome to host metabolism in *Drosophila melanogaster*

Michael Fernandez
Lucas P. Henry
Advisor: Julien F. Ayroles

The declining state of health in America is the result of many factors, including diets high in fats and sugar. Recent research has discovered the important role of the microbiome, the community of microbes found within an organism, in strongly affecting host nutrition signaling and processing of food. Here, our goal was to understand how the microbiome could enhance host metabolism, independent of host genes underlying adaptation. Due to the difficulty that a controlled human microbiome study would encounter, we utilized a long-term evolution experiment in fruit flies (*Drosophila melanogaster*) as a simple model of host-microbiome interactions where flies were evolved on high sugar diets for>100 generations. To investigate the role of the microbiome in host metabolism, we manipulated the microbiome of conventional flies that never experienced the high sugar diet with the microbiome from flies evolved on a high sugar diet. We performed microbiome transplants to inoculate conventional flies with a conventional microbiome, high sugar adapted microbiome, and no microbiome. Developmental time and metabolic traits (protein, glucose, glycogen, and triglyceride percent in flies) were measured to demonstrate how the microbiome affected host metabolism for each of the microbiome treatments. Preliminary results suggest that the evolved microbiome influences host metabolism, independent of host evolution. Our results suggest that the microbiome enhances host metabolism and provides insight into mechanisms underlying the metabolic response of organisms to high sugar diets.

This research was made possible by the generous support of The Graduate School at Princeton University.

Identification of Genes Required for the Production of an Autoinducer-2 Mimic in *Saccharomyces cerevisiae*

Thomas Ferrante
Julie Valastyan
Advisor: Bonnie Bassler

Quorum sensing is a bacterial communication process used to regulate group activities such as biofilm formation and virulence. Through the secretion and detection of molecules known as autoinducers, bacteria are able to assess the size of their population and respond accordingly. Autoinducer-2 (AI-2) is a shared quorum-sensing molecule used by many different bacterial species, and thus thought of as an inter-species bacterial signal. Recently it has been found that bacteria can induce the production of an AI-2 mimic in human epithelial cells, suggesting a potential role of the AI-2 signaling pathway in inter-kingdom communication between bacteria and eukaryotes. Here we report the production of an AI-2 mimic in *Saccharomyces cerevisiae*, more familiarly known as brewer's yeast. We screened a library of yeast gene deletion mutants, testing for mutants that are unable to produce this AI-2 mimic, and identified several genes required for the production of an AI-2 mimic molecule. Deletion of these genes, one of which is a protein of unknown function, resulted in a complete loss of AI-2 mimic production. This further supports a potential role of the AI-2 pathway in fostering inter-kingdom signaling between bacteria and eukaryotes.

This research was made possible by the generous support of the Department of Molecular Biology.

The role of canonical Wnt pathway in regulating neuron morphology

Paola V. Figueroa-Delgado
Hendia Edmund
Advisor: Elizabeth Gavis

Neurons are specialized cells that receive and transmit electrical and chemical signals. Dendrites are essential for signal transduction and adopt diverse morphologies that aid in their function. Therefore, dendrite branch patterning is strictly controlled and maintained during development. Defects in signaling pathways important during development or cell regulation can be attributed to defects in neuron morphology (NM), more specifically under- or over-branching of dendrites. Post-transcriptional gene regulation, specifically the localization of certain mRNAs, has been implicated in the morphogenesis of *D.melanogaster* sensory neurons. Through the process of mRNA localization, proteins involved in neuron morphogenesis are targeted to dendrites, where their expression governs branch patterning and sensory field formation. A serine/threonine kinase, Frayed (fray), was identified in a genome-wide screen for dendritically localized mRNAs in class IV dendritic arborization (da) neurons. RNA interference (RNAi) knockdown of *fray* results in a significant increase in the number of terminal branches (NTB), while overexpression yields dramatically under-branched neurons. Fray is part of the Wnt pathway and is phosphorylated by Wnt target Wnk, to regulate ion channel activity and expression of several transcription factors. Wnt signaling has been extensively studied; however, the role of canonical Wnt pathway in NM and function is unclear. For this reason, RNAi was used to knockdown expression of canonical Wnt pathway components and assess morphological defects in third instar larval class IV da neurons. The β-catenin homolog *armadillo* (*arm*) acts as a master regulator of canonical Wnt. Arm protein translocates into the nucleus where it activates *dishevelled* (dsh) transcription. Statistical analysis of the NTB and loss-of-self-avoidance in class IV da neurons expressing *arm* and *dsh* RNAi may reveal a role for canonical Wnt in regulating NM. Furthermore, if a loss of *dsh* and *arm* rescues *fray* overexpression phenotype it would suggest *fray* acts downstream of both *arm* and *dsh*. Identifying the proteins involved in signaling cascades could lead to better understand neurological diseases associated with morphological defects and potentially lead to novel therapeutics and treatments.

This research was made possible by the generous support of The Genentech Foundation.

Nephronectin and Tenascin-C may be key players in mesangial matrix expansion in diabetic nephropathy

John B. Finlay
Maria E. Vega
Advisor: Jean E. Schwarzbauer

Diabetic nephropathy (DN) is a leading cause of end-stage renal disease and results from an overproduction of extracellular matrix (ECM) in the glomeruli of the kidneys where filtration occurs. Our lab has shown that mesangial cells, which normally provide support to surrounding tissue in the glomeruli, increase fibronectin matrix assembly when grown in 30mM versus 5mM glucose. However, the precise molecular mechanisms of gene and pathway regulation that contribute to ECM over-assembly remain unclear. We performed RNA-Seq on mesangial cell cultures grown in 5mM or 30mM glucose for 72 hours to determine high glucose induced changes in gene expression. DESeq2 and KEGG pathway analysis revealed induction of the bone morphogenetic protein 2 (BMP2) signaling axis and significant upregulation of the potentially downstream kidney developmental protein nephronectin (NPNT). Interestingly, NPNT has recently been suggested as a reliable marker for DN, but its role in the adult kidney is unknown. Immunoblotting revealed that NPNT localizes to the medium as a secreted protein, and time course quantitative-PCR, consistent with our hypothesis of gene regulation, showed that BMP2 expression increases before NPNT. Furthermore, the ECM glyco-protein Tenascin-C (TNC) was significantly upregulated in high glucose, which we show is sufficient to cause alternative splicing of TNC. These results indicate that ECM over-accumulation in high glucose conditions may be mediated specifically by NPNT and TNC.

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund.

Analysis of Post-Initiation Translational Regulation during *Drosophila* Oogenesis by Ribosome Footprinting

Zachary Flamholz
Yingshi Peng
Advisor: Elizabeth R. Gavis

Post-transcriptional gene regulation is increasingly being recognized for its role in spatial and temporal control of protein expression. Research of one such mechanism, translational regulation, has traditionally focused on regulation of the pre-initiation and initiation stages of translation. However, there are situations that necessitate the ability to quickly switch between translation and repression. Such ability could be conferred by a post-initiation regulation mechanism, where ribosomes are paused in the process of elongation and can resume translating upon removal of the pausing mechanism. Early *Drosophila* oogenesis presents a good model for studying such a mechanism, as many maternal mRNAs dumped from the nurse cells into the oocyte require immediate translational repression. Additionally, previous biochemical work in our lab suggests that *nanos* is regulated at the post-initiation step. Ribosome footprinting allows for mapping the position of translating ribosomes with codon level resolution. Using previously collected footprint data as well as data published by the Bartel lab at MIT, a pipeline was developed for identifying possible ribosome stalling sites, manifesting as peaks in fragment coverage within a gene coding region. Five peak containing genes were discovered across three independent wild-type *Drosophila* late-stage oocyte footprint samples. Three of the peaks are maintained in footprint data from three early embryo datasets. Additional footprint samples will be used to confirm peaks and identify others. The results suggest the prevalence of an elongation based translational regulation mechanism. Further research will focus on characterizing the distribution of peaks as well as elucidating the mechanism of peak formation.

This research was made possible by the generous support of the W. Reid Pitts, Jr., M.D. '63 Undergraduate Research Fund.

Role of Non-Coding Single Nucleotide Variants in Autism Spectrum Disorder

Courtni Foster
Vanessa Zhang, Jian Zhou, Laurence Lemaire, Chandra Theesfeld
Advisors: Michael Levine and Olga Troyanskaya

Individuals diagnosed with Autism Spectrum Disorder (ASD) express a range of conditions characterized by social and communication challenges. Autism is a highly heritable complex genetic trait. Sequence and genetics based approaches estimate 400-1000 genes are involved in susceptibility, however, only a few dozen are known: mostly genes with loss-of function coding mutations. Many individuals with autism do not carry mutations in genes, but carry several mutations in non coding regions of the genome. Using deep learning and conventional computational methods, along with experimental approaches, this research aims to characterize the contribution of non coding *de novo* single nucleotide variants (SNVs) to ASD, and uncover the mechanisms of action for prioritized mutations. Using a deep learning method called DeepSEA, a list of SNVs from affected individuals were ranked in order of their predicted impact on chromatin biology and contribution to disease. We show that genomic regions containing some mutations (such as one near NEUROG1), drive differential expression in cells, and show differential regulation at the chromatin level. A model organism, Ciona intestinalis, was used to ask if the SNV genomic regions can drive tissue-specific expression *in vivo*, and whether the expression is in autism-relevant tissue such as the CNS. Our results for the NEUROG1 proband allele (and others) support the predictions made by DeepSEA, and suggest a model whereby the mutation affects gene regulation. Future research may include editing the genome of human embryonic stem cells to carry autism alleles, such as NEUROG1, and assaying phenotypic consequences in neuronally-differentiated cells.

This research was made possible by the generous support of The Genentech Foundation.

Neuronal Feature Encoding in the Hippocampus CA1 Region during an Evidence Accumulation Task in Mice

Nicolas Freeman
Edward Nieh
Advisor: Carlos Brody

The hippocampus has been previously described to encode spatial information as a 'cognitive map' of the current environment. Although individual cells, including place cells, have been shown to encode certain features of the environment, further research is required to more thoroughly classify and understand how these cell types are linked to behavior. In this study, we have utilized two-photon microscopy to measure large-scale neuronal activity in the hippocampus CA1 of *Mus musculus* during an evidence accumulation task. For the evidence accumulation task, a head-restrained mouse runs along a virtual reality T-maze during which visual cues appear on the left or right side; at the end of the maze, the mouse decides to turn left or right, the side with the greater number of cues indicating the correct choice. By recording several features of the task, including cue position, head angle, and task difficulty, we can analyze activation events of single neurons along multiple feature dimensions. First, one-dimensional feature analyses of a neuron's probability density function suggest which features a neuron may encode. Multidimensional k-means clustering analyses may be used to determine how the features a neuron encodes interact with each other during the behavioral task. Future steps include neuron classification across the population based on feature encoding. Validity of the classification analyses will be analyzed with resampling shuffle tests.

This research was made possible by the generous support of the Lane '73 Senior Thesis Fund.

Characterizing Human RASopathy Mutations in *Danio Rerio*

Noah Han
Victoria Patterson
Advisor: Rebecca Burdine

RASopathies are diseases that arise from germline mutations in the RAS/MAPK signaling pathway and affect ~ 1 in 1000 live births. Patients with RASopathies display symptoms including congenital heart disease, short stature, craniofacial defects, and neurocognitive delay. One of the most common RASopathies is Noonan syndrome (NS), affecting 1 in 2000 births, displays typical RASopathy symptoms as well as an increased risk of developing cancer. Half of all Noonan syndrome cases arise from mutations in Src homology phosphatase 2 (SHP2). Mutations that cause an increase in SHP2 activity lead to the upregulation of RAS/MAPK signaling leading NS, while mutations inactivating SHP2 cause Noonan Syndrome with multiple lentigines (NSML). In this study, I will rank mutant SHP2 variants for their impact on development events of *danio rerio* (also known as zebrafish) by microinjecting site-directed mutagenized mRNA into the 1-cell stage embryo. Since RASopathies in humans appear as heterozygous polymorphisms I will generate RASopathies patient mutations in *ptpn11a* (the gene encoding Shp2) through CRISPR/Cas9 effects of these mutations in *ptpn11a* on development. Lastly, I will use a heat-sensitive transgenic line of RASopathy mutant embryos to measure impact on cardiac development when mutations are activated at different times during early development. Hopefully through continued research, we can determine whether this ranking of mutations can be used to predict mutation severity in patients and allow us to quickly assess new variants identified in SHP2 and determine if they are likely to be deleterious to SHP2 function

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund.

Identifying genes involved in the regulation of the PQS operon in *Pseudomonas aeruginosa*

Kristian Harris
Geoff Vrla
Advisor: Zemer Gitai

Pseudomonas aeruginosa (PA14) is a pathogen found to cause many infections in human tissues. It is the major cause of hospital acquired infections and also lung infections in cystic fibrosis patients. Virulence in PA14 was discovered to be surface associated and has been shown to respond to mechanical cues via the use of type IV pili. The PQS operon was discovered to be important in this surface induced virulence. Two molecules produced in the PQS operon, HHQ and PQS, were originally thought to only perform a signaling function but was eventually discovered to be toxic, and thus required for virulence. In this experiment, a red fluorescent protein is used to tag the PQS operon to monitor its regulation and identify what genes are responsible for its regulation, as well as if these genes are surface induced. Transposon mutagenesis was used to generate random genetic mutations in PA14 (wild type *Pseudomonas aeruginosa*) on a surface (agar plates). It is shown, via the reporter, if the expression of the operon is affected. The identified genes involved in this process are unknown and therefore genetic sequencing must be done to identify which genes have been affected by the transposon. So far, 14 potential hits have been sequenced and are being analyzed. Their roles in regulating this operon will be determined and ultimately, if a surface is required for the regulation of these genes.

This research was made possible by the generous support of The Lewis-Sigler Institute for Genomics.

Establishing a model to study the Entry and Transport of Pseudorabies Virus Infection in Central Nervous System Neurons

Rukia Henry
Hao Huang
Advisor: Lynn Enquist

Pseudorabies virus (PRV) is a pathogen of the alphaherpesvirinae family and is neurotropic in nature. PRV is of swine origin and infection is initiated in the epithelium, primarily infecting the peripheral nervous system of the host. This propensity makes PRV a useful and safe model for studying the pathogenesis of alphaherpesvirus in the nervous system. Upon infection with PRV, the virus travels retrograde from the epithelium inside the axon of neurons, and can establish a quiescent infection in the host. Upon reactivation, the virus travels anterograde to either the epithelium, and on rare occasions, the virus can spread to the central nervous system (CNS), causing lethal encephalitis. The infection of the CNS by alphaherpesvirus is rarely studied, hence, developing a useful model for studying alphaherpesvirus infection of the CNS is highly recommended. This study seeks to investigate the entry and transport of PRV in a CNS cell line. We designed experiments using Cath.a-differentiated cells, (CAD). CAD cells were established from a brain tumor in transgenic mice. Neuronal differentiation of these cells can be induced by serum deprivation and can be maintained without addition of protein. We analyzed the CAD cells using immunofluorescence neuronal makers. The CAD cells were then infected with PRV and the entry and transport of the virus was studied using live-cell fluorescence imaging techniques. We observed that CAD cells were permissive to PRV infection by detecting fluorescently tagged viruses in the nucleus and axons. Our findings will help to establish a useful model to study the infection of alphaherpesvirus in the CNS.

This research was made possible by the generous support of The Graduate School at Princeton University.

Modeling and Predicting the Global Gravitational Tide

Benjamin Huang
Jessica Irving
Advisor: Frederik J. Simons

Tidal models based on potential fields and fluid dynamics have existed since Pierre-Simon Laplace formulated his tidal equations. However, despite iterative development on these models into the modern day, theoretical models fail to fully capture the complexity of the global tidal system. In this project, empirical data from tidal gauges were cleaned and categorized using spectral filtering. These data were compared with predictions from theoretical models as well as simple harmonic models that many governmental agencies release to the public (NOAA, EU, etc.). Data were decomposed with statistical signal processing, such as Fourier and spherical harmonic analysis. This decomposition allowed direct comparison to gravitational models of the Sun-Moon-Earth system in the spherical harmonic domain. This work allows further understanding of the importance of gravitational potential's influence on the tide, as opposed to the effect of local variables, such as bathymetry.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Searching for Stemness Negative Regulators of Stem Cell Program in Mammary Gland

Wentao Huang
Nicole M. Aiello
Advisor: Yibin Kang

Mammary stem cells (MaSCs) are a group of cells possessing the ability of self-renewing and differentiation. They are responsible for maintaining the hierarchy of the normal mammary gland throughout life. The discovery of cancer stem cells (CSCs) in breast cancer revolutionized our perception towards the initiation of breast cancer, indicating the stem cell program plays a key role in tumor growth and reoccurrence. A better understanding of the stem cell program could lead to precise therapies targeting cancer stem cells. Numerous positive regulators of stem cell program have been reported in the previous studies. However, little is known about the negative regulators of the stemness, which serve as “brake” for cell growth. Through microarray and qPCR analysis of normal mammary stem cells versus differentiated mammary epithelial cells, we showed *RBPMS* is significantly highly expressed in mammary stem cells. However, clinical RNA-seq data of breast cancer cells showed lower expression level of *RBPMS* than normal cells, and patients with low level expression of *RBPMS* have a shorter survival time. Thus, we hypothesized *RBPMS* is a negative regulator of stemness. To further validate our hypothesis, we generated CRISPR-Cas9 mediated *RBPMS* knockout cell line and *RBPMS* overexpression cell lines to test their ability to form mammospheres, which is an *in vitro* test for self-renewal. We also plan to perform single cell RNA-seq in normal and malignant mammary gland cells for further searching of more stemness regulators. Our research will lead to a new approach towards treating breast cancer.

This research was made possible by the generous support of the Department of Molecular Biology.

Understanding the Paradox in Genetic Regulator Precision

Vishank Jain-Sharma
Advisor: Bill Bialek

The understanding of biological behavior using physical principles, the goal of biophysics, hinges on transcription factor interactions with cellular DNA. Such factors, typically proteins or nucleic acids, determine through their genome-wide binding affinities the level of genetic expression, in the form of proteins, that an organism may utilize to develop and operate. Typically, the concentrations of the original cellular regulators donated maternally to the embryo are precisely conserved across individual organisms. For example, in the fruit-fly embryo, the maternal Bicoid protein forms a precise concentration gradient across the cell which regulates expression of the gene Hunchback and other important developmental genes. However, many such regulators, including Bicoid, are present at comparatively lower concentrations compared to other common cellular products such as calcium ions. This poses an apparent paradox: given that raising the transcription factor concentration would reduce the regulatory input noise, thus advantageously increasing the output precision, why does an organism not simply operate with higher transcription factor concentrations as it does for other molecules? This broad conceptual question applies to all organisms, and investigating potential hypotheses for a resolution has been the focus of this project. Two physical metrics may provide insights: the information transmitted during regulation, and the energetics of the transcription factor binding. Attempts using the former to mathematically analyze a first order model of specific and nonspecific sites have not resolved the question, but work with the latter on binding site specificity, using high-throughput data, shows more promise, and work is currently ongoing.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Hunting for SNPs: The genomic distribution of false positive and negative variant calls

I'ran James
Patrick Reilly
Advisor: Peter Andolfatto

The focal point of this research is to determine the genomic distribution of false positives and false negative variant calls in non-model organisms, as exemplified by Drosophila yakuba. Variant calling is the process of searching for single nucleotide polymorphisms (SNPs) and other genetic differences that can be found between individuals of a species. These variants in DNA may lead to changes in gene expression or protein sequence and function. Simulating the sequencing of genomes of distantly related individuals helps determine where errors in variant calling occur. Using a computational approach, we will leverage our knowledge of the ground truth from simulations to analyze various spatial biases in the distribution of false positive and false negative calls. Our main hypotheses are that false positives and false negatives occur more frequently in non-coding regions than coding regions, and are also enriched in repetitive regions relative to non-repetitive regions of the genome. This experiment is beneficial because it informs scientists as to which particular regions in a genome are reliably called, and thus should be used in further population genetic analyses. In conclusion, we have found that noncoding and repetitive regions are significantly enriched for false positive and false negative variant calls relative to coding and non-repetitive regions, respectively.

This research was made possible by the generous support of The Genentech Foundation.

Functional annotation of gene set enrichment analysis generated networks gives an identity to significantly modified pathways and functions

Gabriel Joseph
Adam Oberstein
Advisor: Tom Shenk

Gene set enrichment analysis reduces the redundancy in differential expression assays, which measure changes in gene expression levels. By functionally annotating communities of gene sets that are concordantly regulated, we can name and identify processes and functions that are significantly modified under experimental conditions. This analysis, while computational in nature, allows us to sift through large amounts of data and understand mechanistically how viruses (or any family of perturbation) effect endogenous cell function. We set out to determine whether it would be possible to look at a community of related gene sets, and automate a system that uses their identities and naming conventions to characterize the family with one term. Our method involves climbing up a clustering tree where members of a community that are closest together, are most closely connected on a tree – the x-y coordinates of a graphical representation of the gene set network, define proximities. Our system allows us to input gene set expression data, pass it into a Javascript that generates a force directed graph, and process it once again in R to name communities. This tool has incredible potential, as you can very quickly identify which cellular functions are significantly modified across different viruses, different time points, all through a pipeline that allows for immediate and side-by-side comparisons.

This research was made possible by the generous support of the Molecular Biology Summer Senior Thesis Fund.

Investigating *Vibrio cholerae*: Analysis of Curvature Proteins

Benjamin Juarez
Edith Blackman
Advisor: Zemer Gitai

Bacteria come in a variety of shapes, and each shape is significant to cell fitness. Bacteria rely on their cell shape to interact with the environment. Some cell shapes allow bacteria to navigate through hosts, obtain nutrients, or develop extracellular structures with other bacteria. Of particular interest to us is *Vibrio cholerae*, a gram-negative curved bacterium that causes the *diarrheal intestine infection, cholera.* Recently, the Gitai Lab discovered *crvA* and *crvB,* two genes responsible for the cell's curvature. When running straight *V. cholerae* cells, ΔCrvA, in an animal model a colonization defect occurred. Not only showing cvrA's linkage to curvature but that curvature effects host colonization. Previous work done on crvA has helped us begin to understand how the cell determines and maintains its shape. However, very little has been done toward understanding its homologue, crvB. Which is also a necessary factor for curvature in *V. cholerae.* Thus, a complete understanding of the mechanism by which cholerae is curved has yet to be achieved. By using biochemical approaches such as right-angle spectroscopy and thermal shift assays, I have begun to explore the structural properties of crvB and how these two proteins might interact and work together in curving cholera.

This research was made possible by the generous support of The Graduate School at Princeton University.

Untitled

Katie Kavanaugh
Advisor: Martin Jonikas

By converting sunlight into organic compounds, photosynthesis produces essential biomatter necessary for the maintenance of life on Earth. However, photosynthetic processes such as carbon concentrating mechanisms (CCM) remain relatively undefined. Research into the genetic underpinnings of photosynthetic unknowns can be conducted in model organisms like *Chlamydomonas reinhardtii* which is ideal for the genetic study because of its 1) well characterized genome 2) known reproductive patterns 3) significant comparability to higher plants. Insertional mutagenesis of known DNA cassettes into the Chlamydomonas genome has been used to generate a photosynthetic mutant Chlamydomonas library. This study aims to examine 264 previously uncovered low confidence candidate genes with potential correlation to Chlamydomonas photosynthesis. The progeny arising from test crosses between mutants of interest and WT strains will be phenotypically screened for photosynthetic abnormalities. PCR assays will be used to verify the presence and location of the insertion within candidate genes predicted to correlate to photosynthetic processes. This work will eliminate mutants containing background secondary site mutations from the pool photosynthetic candidate genes. Additionally, relevant literature will be examined to determine if any of the 264 unannotated candidate genes have been previously implicated in photosynthetic pathways to elucidate their function and importance in Chlamydomonas photosynthesis.

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund.

An Investigation of a Novel Host Determinant of Live-Attenuated Yellow Fever Virus Infection

David Kim
Florian Douam, Gabriela Hrebikova, Qiang Ding
Advisor: Alexander Ploss

Yellow fever virus (YFV) is a prototypic member of the *Flavivirus* genus that continues to cause a significant global burden. A live-attenuated vaccine strain, termed YFV-17D, is available and provides long-term immunity with a single dose. Despite the vaccine's high efficacy and a mere 32 amino acid variations from its virulent parental strain, YFV-Asibi, the mechanism of viral attenuation and the specific determinants governing the virulence of the YFV strain remain poorly understood. Data from our lab support the notion that the effective viral replication process of YFV-17D is specific to primate hosts, and therefore specific host components in humans may contribute to the efficacy of the YFV-17D. However, no such host components, to our knowledge, have currently been identified. Recently, the Ploss lab performed a genome-scale CRISPR-Cas9 functional screen on hepatoma cell lines to identify such host factors essential for YFV-17D replication. We isolated a cell clone exhibiting a strong resistance to YFV-17D infection. This resistance phenotype was associated with a disruption in a gene locus coding for a protein – tentatively referred to as S23 – involved in cell signaling. Preliminary experiments indicated that this protein might play a role on viral RNA replication and viral particle release. In depth characterization of this protein's involvement in the replication cycle of YFV-17D, along with the identification of human-specific determinants, could provide novel insights into the immunogenicity mechanisms of YFV-17D. This research could then potentially help develop novel vaccine strategies for other members of *Flavivirus* family such as Zika, Dengue, and West Nile viruses.

This research was supported in part by the Berry '21 Molecular Biology Senior Thesis Fund, the Grand Health Challenge program from Princeton University (to A.P.) and by an Investigator in Pathogenesis Award by the Burroughs Welcome Fund (to A.P.).

Photoswitchable Control of Protein-Protein Interactions: Engineering Light-Sensitive DARPins

Joshua Kim
Agnieszka Gil
Advisor: Jared Toettcher

The recently developed technology of optogenetics utilizes photoactive proteins to enable control over intracellular processes. One of the well-studied families of photoactive proteins is the Light, Oxygen, Voltage (LOV) domain family, including the Avena Sativa LOV2 (AsLOV2) protein that has been used to create a toolbox of photoactivatable signaling systems *in vitro* and *in vivo*. Recently it was reported that by inserting AsLOV2 into a solvent-exposed loop of a protein of interest, it is possible to destabilize a protein upon light stimulation, reversibly. However, this technique requires a solved crystal structure of the protein of interest along with extensive computational work to correctly insert the AsLOV2 domain. Here, we propose a novel approach based insertion of the LOV domain to construct a library with insertions at all possible positions. We will then screen for photoswitchable binding to a native target of the protein of interest. We will test the viability of this approach in the context of two ERK2-specific DARPins, E40 and pE59. DARPins represent a large class of engineered proteins that all utilize a highly conserved epitopes, therefore we expect the photoswitchable binders we identify to be highly generalizable and easily translated to other classes of DARPins. If successful, our approach will demonstrate the viability of a transposase insertion approach to insert LOV domains into any protein of interest to implement light-controlled binding interactions with the benefit of engineering new optogenetic tools without prior knowledge of the protein's structure. We have already shown that AsLOV2-DARPin fusion proteins can be expressed and purified, with functional and photoswitchable AsLOV2. The AsLOV2-DARPins will be tested for photoswitchability by using LED plates, and tested for ERK2 binding in mammalian cells using Kinase Translocation Reporters (KTR) assays.

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund.

Toward Multiple Site-Specific Labelling on Nonribosomal Peptide Synthetases

Rachel Kim
Hao Li
Advisor: Haw Yang

Nonribosomal peptide synthetases (NRPSs) are a subject of interest because of their unique role as "molecular assembly lines" for different antibiotic compounds. Specifically, we want to understand the physical movement of these structures as they interact in different situations. To do so, we have implemented site-specific labelling techniques to examine the Gramicidin S biosynthetic pathway. We successfully labelled the peptidyl carrier protein (PCP) domain of Gramicidin S synthetase I (GrsA) in two ways using different cofactors. First, we attached a dye to the coenzyme A (CoA) used in the holo-conversion of the protein. We also successfully labelled the substrate azido-phenylalaline (AzF), an unnatural amino acid that reacts covalently at the same location. We are currently working on another labelling technique which uses the cofactor pyridoxal-5-phosphate (PLP) to form a ketone group on the N-terminal of the protein. This allows the dye label to attach specifically to this site. We aim to use these labelling techniques simultaneously in FRET-based analyses to gain a much deeper understanding of how these complex structures change in conformation as they perform their designated tasks. This research could be applied to a wide range of bio and physical chemical studies on protein dynamics and their functional roles, allowing the field to pursue more detailed insight into various microscopic systems.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Automatic Speech Recognition with Kaldi ASR

Robert Benjamin Laitner
Advisor: Naveen Verma

Automatic speech recognition (ASR), sometimes referred to as speech to text, represents the problem of teaching machines to identify and transcribe speech data from an audio stream to recognizable words. Research in such technology has been ongoing for decades; however, recent developments in machine learning, namely the implementation of deep neural networks, has led to major breakthroughs in the past few years leading to a rise in consumer applications. Unfortunately, modern consumer ASR systems only work online through cloud computing due to the computationally intensive nature of speech recognition. Using the open-source software Kaldi ASR, we built a system that records speech data and processes it on the consumer end rather than on the service provider end. The system was implemented with a long short-term memory recurrent neural net that was trained on 460 hours of English speech data to develop an audio model, a dictionary, and language model which the neural net utilizes to decode speech. Audio is passed to the model which develops a multi-dimensional vector known as an ivector to represent qualities of the audio such as background noise and amplitude. The neural net then processes the adapted audio data with the ivector using the dictionary to match phonetic sounds, known as phonemes, to words. The words are then matched in the language model to determine their likelihood of use given the words around them. Accuracy results of the system are pending. If the system performs well, the next goal will be speech recognition in real time.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Reconstituting a Cryostat Refrigerator for Testing of CMB Instruments

Joshua Latham
Michelle Baird, Tristan Shoemaker, Kevin Crowley, Lyman Page,
Advisor: Suzanne Staggs

Studying the CMB tells us about the history of the universe, providing clues about ancient gravity waves, early epochs of inflation, and the original formation of stars. In order to have very sensitive measurements of the Cosmic Microwave Background (CMB), the advACT experiment uses Transition Edge Sensors, electronic components which must be brought down to temperatures of 300 mK to work.

A refrigerator assembly was made in our lab ten years ago to test superconducting components for measuring the CMB. We are testing it to make sure it is in working order, as well as using thermometers and electrical resistor-heaters to better understand the thermal dynamics inside the refrigerator for computer simulations, done through *SolidWorks*.

The refrigerator has two stages, 40 K and 4 K, which are wrapped in super-insulation to prevent radiation from causing thermal exchange between the stages. The refrigerator must operate under vacuum to ensure thermal isolation. With each modification of the refrigerator, we check for leaks and find, at room temperature, suitable leak rates such as 0.25 mTorr/min.

We adjusted the insulation to minimize the amount of heat that reaches the inner stage, finding the optimal insulation set-up that can be used for future projects. Using the thermometers, we saw that the 40 K stage in our configuration reaches 45 K, and the 4 K stage reaches 3.7 K. These are suitable temperatures for testing the He-3/He-4 sorption refrigerators, which will be suspended inside the 4 K stage and reach 300 mK temperatures, temperatures with which we can test the electronic components to ultimately measure the CMB.

Funding for this project came from the MOL/QCB Summer Undergraduate Research Program and the Princeton Physics Department.

Design of a transgenic ERK2-GFP zebrafish line utilizing CRISPR/Cas9 technology

Ju Young Lee
Granton A. Jindal
Advisor: Rebecca D. Burdine

The Ras/ERK (extracellular signal-regulated kinase) signaling pathway is a highly-conserved set of protein interactions and phosphorylations that transduces extracellular signals to affect intracellular processes, such as changes in gene expression and cell proliferation. Due to its importance to cellular function, mutations in this pathway can disrupt normal processes, from complete loss of function to developmental disorders to cancer. The final step of the Ras/ERK pathway is the ERK2 protein, which upon phosphorylation, translocates to the nucleus and affects transcription of downstream genes. The most common methods to read out the state of the pathway is to determine the presence of p-ERK by antibody staining, but such methods require fixed cells, preventing visualization of signaling dynamics. By creating a transgenic line of zebrafish with the GFP fluorescent protein attached to the ERK protein, however, we would be able to visualize a live readout of the pathway by quantifying the levels of fluorescence in the cytoplasm, where ERK is present in the inactive state, and the nucleus, where it is active. Through CRISPR/Cas9 genome editing technology, we create double stranded breaks at the ERK gene and repair the breaks with a homology donor template containing an optimized version of GFP gene – mNeonGreen. With both the Cas9/gRNA complex and the donor repair template synthesized, we conduct injections into the embryo at the one-cell stage – initially the Cas9 complex without the template to verify guide RNA efficacy, and then, upon validation through a high lethality of the embryos, inject the complex with the template to check for fluorescence. This transgenic line will provide insight into the dynamics of the MAPK pathway and contribute to further research in analyzing cell signaling dynamics in living embryos.

I would like to thank everyone in the Burdine lab for their helpful discussions, Rebecca Burdine for funding, and the Department of Molecular Biology at Princeton University.

Design and Optimization of Spectrofluorometer

Miles Lee
Advisor: Quan Wang

Fluorescent labelling and detection has played an increasingly important role in modern biological research by providing researchers with a way to specifically label entities of interest. However, commercially available spectrofluorometers, used to measure the emission spectrum of fluorescent samples, are expensive, bulky, have complicated user interfaces and little option for customization. Therefore, we designed a simple, portable, cost-effective and customizable spectrofluorometer, with an interchangeable light source. Data taken could be viewed live from a PC and stored as a txt file for further analysis in different programs. We optimized the light collection efficiency of the system using both empirical data and theoretical modelling through geometrical optics. We then benchmarked our spectrofluorometer against a commercial design using the intrinsic Raman scattering of water, and achieved a more than 5-fold higher signal to noise ratio. Due to the high sensitivity and ease of customization, this spectrofluorometer system is suitable for daily usage for sample characterizations in any biophysics lab. In future, we hope to implement an LED light source spanning the visible spectrum and interchangeable excitation filters to allow recordings to be taken at any visible excitation wavelength, further improving the customization and usage of our system.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Hysteresis control of TGFβ-induced epithelial-mesenchymal transition generates a distinct program with metastatic abilities

Daniel D. Liu
Toni Celià-Terrassa, Caleb Bastian, Brian Ell
Advisor: Yibin Kang

The epithelial-mesenchymal transition (EMT) is a key step during cancer metastasis, endowing epithelial tumor cells with migratory and invasive capabilities, thus allowing tumor cells to disseminate through the body and establish distant site metastases. The signals and transcription factors (TFs) inducing EMT are widely known, which converge to suppress the expression of epithelial genes, most prominently E-cadherin. However, less is known about the intrinsic spatiotemporal dynamics of the transition. Mathematical modeling and experimental validation of TGFβ-induced EMT shows a non-linear response resulting in a hysteretic bistable switch in E-cadherin expression, tightly controlled by the Zeb/miR-200s double-negative feedback loop. In other words, cells are predicted to undergo all-or-nothing transitions between epithelial- and mesenchymal-like states. Using CRISPR/Cas9-mediated genome editing, we disrupted a critical interaction breaking the Zeb/miR-200s feedback loop, which successfully eliminated hysteresis *in vitro*. We show that hysteresis ensures quick responses and molecular memory, which sustains the EMT-like state long after withdrawal of stimuli. Remarkably, cells gained invasive properties equally regardless of the type of transition; however, only cells following hysteresis increased lung metastatic colonization. Accordingly, only cells undergoing hysteretic EMT differentially expressed subsets of stem cell-like and extracellular matrix genes with clinical prognosis value. Overall, these findings illustrate the existence of distinct EMT programs which have functional effects on metastatic properties.

This research was made possible by the generous support of the Hickok Molecular Biology Senior Thesis Fund.

One-Carbon Metabolism in Small Cell Lung Cancer

Jennifer Liu
Gregory S. Ducker
Advisor: Joshua D. Rabinowitz

In all organisms, one-carbon (1C) metabolism is essential for normal cell growth and functioning as it fuels multiple cellular activities including nucleotide, DNA, RNA, and protein synthesis. 1C units can enter the pathway through the activity of the enzyme serine hydroxymethyltransferase (SHMT), which catalyzes the transfer of a carbon atom from serine to glycine and tetrahydrofolate (THF). For most proliferating cells, including cancer cells, the amino acid serine is the dominant source of 1C units; thus, inhibiting this process may have antiproliferative effects. Using an inhibitor of SHMT and isotope tracing, we investigated serine and 1C metabolism in small cell lung cancer (SCLC) cell lines, a tumor type with limited existing therapeutic options. Isotope tracing results comparing control and SHMT inhibitor treated cells in media containing 2,3,3-^{2}H serine show decreased amounts of M+1 labeled serine in treated cells, especially H1930 cells, suggesting SHMT inhibition is effective in suppressing the 1C metabolism pathway. This result is further supported by low amounts of M+2 labeled glutathione, which indicates decreased activity in mitochondrial SHMT to convert 2,3,3-^{2}H serine to M+2 labeled glycine. Growth assays conducted to measure the effect of SHMT inhibition on cell proliferation have yet to yield results due to the difficult culture conditions of these cells. Further investigation of the requirement for SHMT activity in SCLC cell lines may help highlight novel avenues for designing treatments targeted to this disease.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Charge State Readout of Nitrogen Vacancy Centers in Diamond

Olivia Long
Advisor: Nathalie De Leon

Nitrogen vacancy (NV) centers are naturally-occurring defects in diamond consisting of a substitutional nitrogen atom and an adjacent vacancy in the diamond's carbon lattice. The electronic spin state of NV centers can be optically detected and initialized under laser excitation. Moreover, NV centers have long spin coherence times on the millisecond timescale at room temperature. This makes them attractive candidates for quantum nanoscale sensing. However, NV centers can exist in two different charge states, leading to errors in initialization and readout. Since ionization dynamics can be drastically different near the diamond surface, we explore the ionization process between NV^- and NV^0 for shallow NV centers within 10 nm of the surface. Using two lasers to provide different excitation energies, we selectively excite both charge states (532 nm green laser) or just NV^- (589 nm orange laser) to distinguish between the two states. The ratio of the charge states at equilibrium will differ depending on the individual NV center. To determine the ratio of a given NV, we use a laser pulse sequence to first excite both states. Then, we use the orange laser to selectively excite NV^-. From the fluorescence data, we can determine the ratio of the two different states, and we achieve single shot readout of the charge state. Further work includes analyzing different NV centers and possibly correlating the ratios with other NV characteristics. Ultimately, a deeper understanding of the conversion between NV^- and NV^0 may allow us to engineer an environment for optimal nanoscale sensing capabilities.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Redacted

Yuting Lu
Thomas Heimbucher, Jasmine
Ashraf Advisor: Coleen T. Murphy

Redacted

This research was made possible by the generous support of the Department of Molecular Biology.

Cobalt Chloride-mediated Hypoxia affects reproduction, fat metabolism and aging in *Caenorhabditis elegans*

Yikai Luo
Advisor: Coleen T. Murphy

Abstract not disclosed.

This research was made possible by the generous support of the Department of Molecular Biology.

Identifying Biomarkers of Aging in Human Blood

Luisa Mercado
Jenelys Ruiz, Sean King, Rachel Kaletsky
Advisor: Coleen T. Murphy

Human aging is the result of genetic and environmental factors. Biomarkers of aging are valuable because they can report on individual's biological age compared to chronological age. Therefore, identifying biomarkers from easily accessible samples, such as blood, is important. Biomarkers can be identified through the expression patterns of genes under specific conditions. High-throughput sequencing technologies and bioinformatic tools have enabled a more comprehensive study of the expression patterns within the entire genome simultaneously, thus easing the identification of genes that can serve as biomarkers. We aimed to identify biomarkers of age on blood and analyze the effects that unhealthy life-style factors, such as smoking, have on them, by performing gene expression analysis on a database of RNA-sequenced blood samples of women who belong to different age groups. Advanced bioinformatic tools such as TopHat, HTSeq, and DESeq2, along with diverse visualization tools, resulted in the detection of a differentially expressed Antisense RNA (asRNA) identified as FO704657, exhibiting a p-value of 0.049. The effects of smoking in the expression level of FO704657 were not significant, however, there was reduced expression in the group of 30-year-old individuals who smoke, in contrast to an increase in the group of 40-year-old individuals who smoke, suggesting an age-dependent factor involved. Expanding this project on a much bigger sample size, could better support these results. Additional work is needed to determine whether this RNA could be used as a biomarker of aging. The effects of different life-style factors, such as smoking would add clarity to those studies.

This research was made possible by the generous support of The Genentech Foundation and The Graduate School at Princeton University.

Visualizing Earth System Models: From Communes to Continents

Jivahn Moradian
Advisors: Nathaniel W. Chaney and Elena Shevliakova

Earth System Models (ESMs) simulate the global coupled water, energy, and carbon cycles at spatial resolutions ranging between 25 to 100 km; computational constraints limit running these models at higher spatial resolutions. To account for the role of fine-scale heterogeneity observed over land, within ESMs, each grid cell is divided into different tiles, which are derived by clustering available environmental data (e.g., elevation and soil types). Although these tiles provide higher resolution model output than the overlying ESM, this data is rarely used and instead summarized through a grid cell weighted average -- this is a missed opportunity. Our objective is to develop a program that enables the user to adapt the model's resolution to his or her own specific needs, allowing them to view the region of interest with a finer resolution than they could with the default model output. Extensively using the Geospatial Data Abstraction Library (gdal), we created code that adjusts the data obtained from ESMs to enable individuals to input the temporo-geographic frame of interest as well as their desired resolution. It accomplishes this by using parallel processing to create a template grid determined by the selected region of interest, then upscales the useful data to the specified resolution. We next fine-tuned the code by removing most predefined directories and allowing the user to modify a greater number of parameters, thus broadening their control of the program. We intend to publish the code in the form of an open-source Python library, and potentially adapting it to a more user-friendly Google Maps style presentation. By doing so, we hope to make the data obtained from these ESMs accessible to a wider audience.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Quantitative Analysis of Aggressive Fruit Fly Behavior

Megan Ostrowski
Ugne Klibaite
Advisor: Joshua W. Shaevitz

Aggressive behavior is crucial for survival in nature, and therefore it is important to understand the particular interactions that underlie aggression. However, behavior is complex and notoriously difficult to quantify and we know little about the specifics of aggressive behavior. We employ a novel imaging and analysis technique for quantifying behavioral interactions involving the tracking and registration of individuals in order to describe a wide and highly resolved range of fly behaviors. Previous research of aggressive behaviors in fruit flies has shown differences between the behaviors of males and females in same-sex pairings. We extend this analysis to compare behavior across several strains as well. We use the model organism Drosophila *melanogaster* to explore the behavior of interacting individuals with an experimental focus on aggression. To do so, we generated movies of same-sex pairings of wild-caught and lab strains of Drosophila melanogaster during a competition for access to food. We analyzed these movies for positional data, the distance between individuals, and proximity to the shared resource. We find differences using the positional and distance data in the way aggression is expressed between wild-caught and lab strains, as well as across the sex of the pairings. We are further analyzing the movies to map specific behaviors into a 2-dimensional behavioral density space using unsupervised machine learning techniques and hope to further enumerate the full set of behaviors that constitute the repertoire of aggression, and determine if there are quantitative differences between wild-caught and lab strains and between males and females.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Characterization of the host-virus mitochondrial interactome during HCMV infection

Clayton Otter
Joel Federspiel
Advisor: Ileana Cristea

In order to successfully infect its host, human cytomegalovirus (HCMV) must both evade the host immune response and hijack cellular machinery to manufacture and export virus progeny. Recent work by our lab and others has shown that HCMV has a profound effect on the mitochondria, as increased mitochondrial fragmentation and localization to the periphery of the virion assembly complex (AC) during infection has been observed. Our project aims to identify virus-host interactions that may be responsible for these changes in the mitochondria. In order to characterize the mitochondrial virus-host interactome, we used immunoprecipitation-mass spectrometry (IP-MS) techniques. We first optimized mitochondrial fractionation, viral protein extraction from the mitochondria, and isolation of tagged viral proteins. Then, we applied these optimized methods to examine the interactome of viral protein pUL37x1. Though the mitochondrial-localized pUL37x1 has been studied extensively for its role in preventing host cell apoptosis, no interaction network studies for this protein have been conducted. Following infection by a virus that expresses pUL37x1 tagged with GFP, mitochondria were isolated, IPs were conducted to pull-down pUL37x1 along with any interacting proteins, and MS was used to identify these interactors. These preliminary experiments have identified specific host and viral proteins as pUL37-interacting partners that play a role in various cellular pathways, including mitochondrial translation and protein processing. A better understanding of this interactome will provide further insight into the progression of infection and can help identify targets for the development of antiviral treatment.

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund.

Interferon-inducible protein 16 (IFI16) in immune signaling during herpesvirus infection

Catherina Pan
Krystal K. Lum
Advisor: Ileana M. Cristea

Human interferon-inducible protein IFI16 is a sensor of exogenous DNA critical for intrinsic and innate immunity. During herpes simplex virus-1 (HSV-1) infection, IFI16 represses viral gene transcription and induces antiviral cytokine expression to suppress pathogenesis. However, how IFI16 initiates these downstream immune responses remains unclear. Our lab and others have previously demonstrated that IFI16 co-localizes with incoming viral genomes and promyelocytic leukemia (PML) proteins, and that its pyrin domain (PYD), which functions in homotypic oligomerization, is necessary for these localizations. Since protein oligomerization occurs among several immune signaling proteins, and PML proteins are well-known for their antiviral functions, we hypothesize that the ability to oligomerize is necessary for IFI16 antiviral functions. Moreover, this study aims to determine the downstream mechanism by which IFI16 oligomerization exerts these effects. To address these questions, we created oligomerization-disrupted PYD mutants and assessed the functional consequences of non-oligomerizing IFI16 in the context of HSV-1 infection. Using immunofluorescence, western blotting, and quantitative real-time PCR, we found that these PYD mutants exhibited increased viral gene expression and protein levels when compared to wildtype IFI16, indicating that PYD oligomerization is important for IFI16 antiviral functions. With the rise of gene therapy and DNA-based vaccines, this study's findings will lend key insight into the interaction of the immune system with herpesviruses, and may enable the development of safe and effective vaccines against viral, bacterial, and parasitic diseases.

This research was made possible by the generous support of the Berry '21 Molecular Biology Senior Thesis Fund.

Characterizing the role of DACT1 in Wnt signaling and breast cancer bone metastasis

Nana Park
Mark Esposito
Advisor: Yibin Kang

Metastasis, the dissemination of tumor cells to distant parts of the body, accounts for 90 percent of cancer-related fatalities, and among breast cancer patients, 70 percent of metastatic burden can be attributed to metastasis to the bone. Yet, no reliable markers exist to prognose breast cancer patients at high risk of developing bone metastasis. To identify bone metastasis promoting factors, we screened for genes over-expressed at least two-fold in highly bone metastatic derivatives of breast cancer cell lines. Among genes identified, Dishevelled Binding Antagonist of Beta-Catenin 1 (DACT1) was the only independent prognostic marker of reduced distant bone metastasis-free survival. Preliminary evidence suggests that DACT1, though poorly characterized in the literature, is a Tgf-β-induced upregulator of Wnt signaling, demonstrated by our lab to be active in metastatic cells adjacent to the bone vasculature. Given the implicated role of DACT1 and Wnt signaling in bone metastasis, we then generated a bone metastasis signature using Wnt target genes highly correlated with DACT1 ($R^2 > .25$) both in clinical data sets and bone metastatic cell lines. Interestingly, high expression of the DACT1-derived gene signature had no predictive value for patients with lung, breast or liver metastasis, but was associated with a nearly two-fold higher risk of poor distant bone metastasis-free survival. Together, this evidence suggests that Wnt signaling and DACT1 mediate a bone metastasis-specific program. Ongoing investigations attempt to validate the gene signature and characterize DACT1-dependent activation of Wnt signaling in bone metastasis through in vitro assays and in vivo xenograft experiments.

This research was made possible by the generous support of the Hickok Molecular Biology Senior Thesis Fund.

Exploring mechanisms of spinal curvature in zebrafish models of Idiopathic Scoliosis

Emily Pauls
Dan Grimes
Advisor: Rebecca Burdine

Idiopathic Scoliosis (IS), a disease that causes chronic pain, is characterized by 3D spinal curvatures with severe curves requiring surgery. IS onsets in adolescence and curves are more prevalent in girls than boys. Zebrafish (Danio rerio) have recently emerged as excellent models of IS. Our lab found that defects in motile cilia – an organelle projecting from the cell surface into extracellular space that beats back and forth to generate fluid flow – and cerebrospinal fluid (CSF)-flow cause spinal curves. However, we do not know which ciliated cells are involved nor do we know how CSF flow is sensed in the spine.

My work tests the hypothesis that CSF-contacting neurons sense flow using ciliary-localized Polycystin proteins (Pkd2l1 and Pkd1l2a). To test this, I am assessing spinal curve development in mutants in these genes. I have found mild spinal curves in pkd2l1 mutants. By contrast, lethality in pkd1l2a homozygous mutants has thus far prevented robust analysis of IS-like phenotypes. My second goal is to use cell-specific ablation techniques to determine which cell types in the zebrafish brain and spine are required for spinal linearity. To date, I have broadly ablated radial glia and neurons and am observing expected embryonic defects including twitching, early spinal curvature, paralysis, and pericardial edema. Ablations at later time points will assess the requirement of these cell types in spinal linearity.

By conclusion, I hope to improve our knowledge of the role of cilia in scoliosis, knowledge which can be applied to refine treatments for those afflicted with IS.

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund, the Summer Undergraduate Research Program in Molecular and Quantitative & Computational Biology, and Phil Johnson for fish care.

Untitled

Tony Phan
Stephen Xie
Adviser: Tom Muir

Methylation of lysine-4 of histone H3 (H3K4) plays a crucial role in the regulation of chromatin structure and gene transcription. KDM5B is an enzyme responsible for demethylating this residue, and its dysregulation has been implicated in the proliferation and invasive migration of cancer cells. This enzyme contains a PHD1 finger domain, which has previously been shown to associate with the unmodified N-terminal tail of histone H3. An unpublished finding from our lab suggests that a peptide sequence in KDM5B may mimic this histone tail. Thus, PHD1 may interact with this histone-tail mimic peptide to induce a conformation in KDM5B that intrasterically inhibits the enzyme. However, it is not known whether this interaction take place and what effect it has on H3K4 demethylation by KDM5B. This current study aims to investigate whether this interaction takes place using ^{1}H-^{15}N HSQC NMR titration. We also examine whether this interaction modulates H3K4 demethylation by KDM5B using immunofluorescent staining of cells transfected with KDM5B mutants aimed to abrogate the interaction. Successful completion of this study will provide insight into the regulation of KDM5B, potentially contributing to therapeutic strategies for the treatment of cancers and other gene regulatory disorders.

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund.

Optimization of lysis conditions for quantitative proteomics in *Escherichia coli*

Dennis Pérez-López
Meera Gupta
Advisor: Martin Wühr

E. coli is a classical model organism that has been widely used to study, among many other subjects, mRNA and protein level control, metabolomic regulation, and evolutionary fitness. For many of these studies it will be beneficial to quantify the absolute levels of protein abundances and their change upon perturbations. Recent advances in mass-spectrometry based proteomics make protein abundance measurements possible on a genome-wide scale. However, an essential requirement for these measurements is that we are able to reproducibly and comprehensively extract all proteins from the sample. This is particularly difficult in hard to lyse gram-negative bacteria like *E. coli*. To this end we have optimized the lysis method for *E. coli*. We compared various lysis conditions with different chaotropes, detergents, and physical perturbations. We obtained the best lysis conditions using 6M GuaCl, with 2% Hexadecyltrimethylammonium bromide (CTAB) and repeated sonication. Importantly, this lysis condition was able to extract a significantly larger fraction of proteins than the condition used in a previous *E. coli* proteomics study but also more proteins than any of the other conditions tested. Next, we will use the newly developed lysis protocol to measure *E. coli*s' protein abundances under well-defined growth conditions.

This research was made possible by the generous support of The Graduate School at Princeton University.

The Role of Dopaminergic Terminal Activity in NAc and DMS during Aversive Fear Extinction

Katherine Pizano
Lili X. Cai
Advisor: Ilana Witten

Dopamine (DA) neurons have long been implicated in encoding prediction errors in reward-based learning. However, the role of striatal DA neurons in aversive learning remains largely unknown. We hypothesize that DA terminal activity within the striatum, specifically the NAc and DMS, helps encode a teaching signal which enables extinction learning. To test this hypothesis, we use an aversive fear conditioning behavioral paradigm, along with calcium imaging, to record from striatal DA terminals during extinction. This allows us to analyze how these subpopulations of neurons correlate with conditioned response of freezing to aversive associations. Our preliminary findings suggest DA activity in the NAc and DMS during CS (tone) offset is, in fact, correlated with the aversive action (freezing) within the current trial. Moreover, a decrease in DMS terminal activity at the CS onset also correlates with freezing during extinction trials. Together, both these signals may help drive the conditioned aversive response by preventing extinction learning from occurring. Our next steps will include using optogenetics to manipulate and record from these DA terminals. This will provide causational evidence as to whether dopamine is involved in reinforcing aversive actions or updating the CS-US association. Understanding dopamine's role during extinction of aversive stimuli can help better our understanding of psychiatric disorders like PTSD, which show strong evidence for the involvement of the DA pathway.

This research was made possible by the generous support of the Crecca '46 Molecular Biology Senior Thesis Fund.

Investigating the Role of Multiple Importin-Binding Sites in the Phase Transition of TPX2, a Critical Spindle Assembly Factor

Aparna Raghu
Raymundo Alfaro-Aco, Matthew King
Advisor: Sabine Petry

The rapid and accurate assembly of microtubules to form the mitotic spindle is necessary to ensure proper chromosome segregation during cell division. This process relies on the coordination of many spindle assembly factors (SAFs). One important SAF in *Xenopus laevis* is TPX2, a protein that promotes microtubule nucleation throughout the spindle. TPX2, in association with the importin α/β complex, can potentially undergo phase transition. The resulting phase droplets may prevent diffusion and delay TPX2 activation until molecules are released through the interaction of importins with RanGTP. To understand the role of importin-binding sequences in TPX2 phase transition, this study will first identify the importin-binding sites of TPX2. Using GST pull-down assays, I will map the nuclear localization sequences (NLSs) of TPX2 by characterizing the interactions between sequential truncations of TPX2 and importins. In addition to the two previously-identified NLSs of TPX2, I predict that other sequences, particularly those identified as putative NLSs *in silico*, will interact with importins. I will then determine whether the presence of multiple NLSs enhances TPX2's ability to undergo phase transition by mutating basic residues within identified NLSs to alanines, rendering them nonfunctional. I hypothesize that the presence of multiple importin-binding regions helps promote TPX2 phase transition by providing multivalency. This study will provide insight into how the presence of multiple importin-binding domains contributes to phase transition of TPX2. This research will also help inform future studies exploring how phase transition regulates mitotic spindle assembly by concentrating and coordinating activation of TPX2 and other SAFs.

This research was made possible by the generous support of the W. Reid Pitts, Jr., M.D. '63 Undergraduate Research Fund.

Characterizing the Interaction between Sec39/Use1 in the Dsl1 Multisubunit Tethering Complex

Gabriel Ramirez-Arellano
Sophie Travis
Advisor: Frederick Hughson

Vesicular transport of luminal cargo from one organelle to another is crucial in the eukaryotic cell. The Dsl1 complex is required for intracellular vesicular trafficking from the Golgi apparatus to the endoplasmic reticulum (ER). The complex is localized to the ER by tight interactions between two of its three subunits, Tip20 and Sec39, and the Q-SNAREs Sec20 and Use1 respectively, and appears to remain associated with the SNAREs throughout its catalytic cycle. While an atomic model of the Dsl1 complex bound to the SNARE Sec20 has recently been solved, the position of the remaining SNARE, Use1, remains unknown. Here we report progress toward characterizing Sec39 in complex with the N-terminal domain of Use1. Through ortholog screening, we have found that the subcomplexes Candida glabrata Sec39/Use1^{N} and Kluyveromyces lactis Dsl1^{C}/Sec39/Use1^{N} are promising targets for crystallography: highly soluble, monodisperse via gel filtration, and comprised of a stoichiometric amount of both proteins. We have begun sparse-matrix screening of crystallization conditions of the full-length complex, as well as identifying stable fragments via limited proteolysis. Obtaining the structure of the Sec39-Use1^{N} complex will provide the final piece required for a complete atomic model of the SNARE-bound form of the Dsl1 complex, the functional form of the complex in the cell. As the only complete atomic model of a tethering factor bound to SNAREs, it will also yield insight into conserved mechanisms of tethering factor function.

This research was made possible by the generous support of The Genentech Foundation.

Identification of Reproductive Span Regulating Genes in *C. Elegans* Using Mos1 *Insertional Mutagensis*

Rachel Reed
India Rogers-Shepp, Rachel Kaletsky, William Keys
Advisor: Coleen T. Murphy

Reproductive decline in females is one of the earliest signs of aging in the human population. With more women having children later in life, understanding the genetic and molecular causes of reproductive decline is becoming increasingly important. Using the model organism *C. elegans,* a screen can be run using *Mos1* insertional mutagenesis to find mutations in the genome that result in extended reproductive span. *Mos-1* mediated insertional mutagenesis takes advantage of the capabilities of the *Mos-1* transposon of *Drosophila mauritiana.* Using *Mos1*-mediated mutagenesis, we can mutate the animals, screen them for extended reproduction, then identify and characterize the lesions created by the *Mos1* sequence. This protocol allows easy identification of the insertioned mutations in worms that were screened. Although the Mos-1 transposon mutagenesis method has a lower efficiency than classical chemical mutagens, it is advantageous because it provides an opportunity for high-throughput sequencing and mapping, since all mutations would be tagged with the *Mos1* gene. The work completed this summer has involved with testing out the protocol and measuring the transposition efficiency of the *Mos1* sequence. With the final results of this screen, several key genes will be characterized further to determine their function in extending reproductive span.

This research was made possible by the generous support of the Crecca '46 Molecular Biology Senior Thesis Fund.

How stress can phase out nucleocytoplasmic shuttling

Anastasia Repouliou
David W. Sanders
Advisor: Cliff P. Brangwynne

To face the growing burden of neurodegenerative diseases, we must develop pathological models. Increasing evidence suggests that the hallmark protein aggregates found in neurodegenerative disease patients result from aberrant liquid-to-solid phase transition of proteins that physiologically phase-separate into liquid-like droplets. During stress, proteins and mRNA are recruited to stress granules (SGs), a class of phase-separating organelles. Nucleocytoplasmic shuttling through the nuclear pore complex (NPC), which is often disturbed in stress, is mediated by nuclear transport receptors (NTRs) that avidly bind to the intrinsically disordered domains of FG-Nups, the proteins guarding the NPC channel. FG-Nups have also been shown to phase-separate into gel-like assemblies, which NTRs traverse like liquids, suggesting NTRs hold a disaggregase in addition to chaperone potential. This study aims to probe the poorly understood interplay between stress response and nuclear transport. We propose that, during stress, SGs sequester NTRs, which act to enable the entry of larger molecules into SGs, while simultaneously disrupting nucleocytoplasmic shuttling. We will therefore test NTR localization to canonical and artificial SGs and to FG-Nups nucleosolic inclusions, and mark changes in the complex body dynamics in NTR overexpression and knockout. Moreover, we will quantify the nuclear to cytosolic ratio of protein and mRNA reporters in the artificial SG-model. The expected results will yield valuable insights into the synergistic transport and disaggregase potential of NTRs, and SG and NPC mechanisms and physiology.

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund and the Class of '43 Molecular Biology Senior Thesis Fund.

Metabolic Flux Modeling of One-Carbon Metabolism

Jorge Reyes
Lifeng Yang, Sheng Hui
Advisor: Joshua D. Rabinowitz

Mass spectrometry presents a relatively simple and accurate method of tracing molecular labeling for metabolite identification in a metabolic network. However, the information provided by mass spectrometry is not limited to such static measurements. With knowledge of the atomic transitions in the chemical reactions that constitute a metabolic network, mass isotopomer (MI) measurements can be employed to determine flux values. Here we present a coarse-grained whole-body network concerning folate-mediated one-carbon (1C) metabolism using MI measurements from *in vivo* infusion experiments. 1C metabolism supports numerous biological processes including (1) purine and thymidine biosynthesis, (2) amino acid homeostasis, and (3) redox defense. Using cumulated isotopomer (cumomer) fractions we reduce the nonlinear system of equations that describe the bimolecular reactions of this network to readily solvable linear systems of equations. In doing so, stoichiometrically constrained flux values can be calculated by optimization with MI measurements. Additionally, we present a simple method for determining the contribution of specific fluxes to formate, a key metabolite in 1C metabolism, in numerous tissues by MI measurements. These fluxes present a source of information that enhances the systematic understanding of 1C metabolism. Allocation of biochemical resources to any metabolic network depends on the state of the organism. Ultimately, the fluxes of a network inform us on the metabolic economics of such a state.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Investigating the Role of DNA Methylation in Stress-Attributed Telomere Attrition

Elizabeth Reznik
Lisa Schneper
Advisor: Daniel Notterman

Epigenetics is the study of covalent modifications to DNA that lead to structural changes within chromatin structure and alter transcription. DNA methylation is the addition of a methyl group to a cytosine that precedes a guanosine, and hypermethylation of CpG islands at promoter regions is associated with inhibition of gene expression. Such genes can include those that regulate the lengths of telomeres, regions of repetitive non-coding sequences at the ends of chromosomes that are implicated in stress-attributed health responses. Few studies have examined the relationship between DNA methylation and telomere length and the molecular mechanism involved. To understand this relationship, data from the Fragile Families and Child Wellbeing Study, a longitudinal analysis of a cohort of approximately 5000 children from 20 cities, will be used to identify candidate promoter regions that are differentially methylated under stressful situations and correlated with telomere attrition. The identified promoter regions will be cloned into pCpGL, a novel luciferase reporter vector that completely lacks CpG dinucleotides on its backbone, thus limiting the effects of DNA methylation on the promoter of interest. By establishing a Dual-Luciferase Reporter Gene Assay, we hope to test the effects of DNA methylation of a promoter region implicated in telomere length regulation on reporter gene expression. This work can provide evidence on a molecular level that chronic stress, particularly at a young age, may predispose an individual to disease via epigenetic modifications.

This research was made possible by the generous support of the W. Reid Pitts, Jr., M.D. 63 Undergraduate Research Fund.

Creation of a single worm *Mos1* transposon screen for *Caenorhabditis elegans*

India Rogers-Shepp
Rachel Kaletsky
Advisor: Coleen T. Murphy

Caenorhabditis elegans are a useful model organism for studying both reproductive and somatic aging. Similar to humans, this hermaphroditic worm ovulates for one-third to one-half of their lifespan. Additionally, oocyte development is highly conserved between the two species. Likewise with respect to somatic aging, the Insulin/IGF-1 pathway, TOR signaling, AMP kinases, and sirtuins, have all been shown to be conserved in both species. Discovering the genes responsible for regulating aging is often random. Our lab has proposed to expedite the search for these genes by creating a screen for genes either responsible for longevity. I propose to use a *Mos1* transposon mediated insertional mutagenesis to create a random pool of mutants, followed up with a screen for longevity mutants. I then plan to test individual worms with a *Mos1* screen for single worms to detect the location of the mutation: by testing individual worms I will be able to isolate the strains and study their progeny as well. After many experiments we successfully isolated the *Mos1* transposon element from single worms with a known *Mos1* insert, showing that a single worm contains enough DNA for a genetic screen. Next, I plan to see if aged worms contain enough genetic material for the *Mos1* test. Once I have determined the appropriate parameters for this screen, and I move forward to use it with unknown mutants, this screen will both reduce the amount of time needed to find longevity mutants and significantly increase the known number of longevity genes.

This research was made possible by the generous support of the Crecca '46 Molecular Biology Senior Thesis Fund.

Effect of lifestyle factors and aging on differential gene expression patterns in human blood

Jenelys Ruiz-Ortiz
Luisa Mercado, Sean King, Rachel Kaletsky
Advisor: Coleen T. Murphy

Previous research has shown that there are differences in gene expression across various aging cell types in healthy individuals. Moreover, certain diseases and factors that affect lifestyle, such as obesity, have genetic underpinnings and are known to affect lifespan. However, the genetics of these factors have rarely been assayed for their expression patterns across ages in different cell types. It is especially convenient to do this type of analysis in a readily obtainable tissue such as blood, as this cell type could be a potential target for the design of accessible diagnostic tests. Therefore, our goal was to study the effect of lifestyle factors in combination with aging on gene expression patterns in blood cells to identify biomarkers of aging. To study this, 24 women were surveyed and whole blood sequencing was performed using RNA-sequencing (RNA-seq), and the data was analyzed using bioinformatics tools. The first factor chosen to investigate was Body Mass Index, which was categorized as normal weight (18.5 – 25.0 BMI) and overweight (>25 BMI). An identified differentially expressed transcript corresponding to an antisense RNA (asRNA) in the healthy aging group was isolated for all groups and correlation analysis was performed to find patterns of high and low expression. This analysis is currently ongoing, but the outcome of this project will help establish a protocol for the analysis of a higher quantity of samples, to compare patterns of expression across adjacent ages, and to study how different factors affect longevity at the transcriptome level.

This research was made possible by the generous support of The Graduate School at Princeton University.

Translational burst size and gene expression noise in *E. coli*: Finding an optimum

Christopher Russo
Zhiyuan Li, Sophia Li
Advisor: Ned Wingreen

Gene expression is an inherently noisy process due to the stochastic nature of biochemical events in the cell. High throughput experiments using RNA-seq and ribosome profiling have shown translational efficiency and mRNA degradation rate to be widely-varying and evolutionarily tunable parameters in *E. coli*.These parameters define the translational burst size, which contributes to gene expression noise. Cost benefit theory was used to determine whether translational burst size is optimized to balance the fitness costs of gene expression noise and mRNA turnover rate.This theory was applied to the lac operon, and experimentally determined values for various biophysical parameters were used to approximate an ideal translational burst size for the repressor
LacI. Our results preliminarily suggest that burst size is in fact optimized. Cost benefit theory was also used on the lac operon to understand whether noise suppression or response time of the regulatory system is more important for determining optimal mRNA degradation rate. Estimating this ratio was made difficult by the effect of the level of variation in environmental conditions on the fitness benefit of a fast response time. Analysis of genome wide translational efficiency datasets show it is significant correlated with gene expression noise, but that this is primarily due to the strong correlation between translational efficiency and mean protein number, which also affects gene expression noise. Our work shows the application of cost benefit theory to translational burst size can give insight into the evolutionary factors that determine parameters like translational efficiency and mRNA degradation rate.

This research was made possible by the generous support of the Lewis Sigler Institute and the Princeton Department of Molecular Biology.

Niemann-Pick Disease (NPC) family proteins in Embryonic Germ Cell Migration

Peyton Smith
Chris Ng, Diane Manry, Girish Deshpande
Advisor: Paul D. Schedl

Proper migration of Primordial Germ Cells (PGCs) in *Drosophila melanogaster* embryos requires attractive and repulsive cues that guide the germ cells to the somatic gonadal precursor cells (SGPs). During embryogenesis, the PGCs migrate through different cell layers in order to associate with the SGPs and form a primitive gonad. Recent studies of Drosophila germ cell migration have shown that the signaling ligand Hedgehog (Hh), the biosynthetic enzyme HMGCoA reductase (Hmgcr), and the ABC transporter Multi-Drug-Resistant-49 (Mdr49) all have roles in the germ cell migration pathway. Because the *mdr49* homologs in mammals (Mdr1/P-cg) provide a cell autonomous function of transporting cholesterol, proteins that are involved in the transport of sequestration of cholesterol, such as niemann-pick disease type C 1 (Npc1) family proteins npc1a and npc1b, are of interest due to the implication of cholesterol's role in PGC migration. Preliminary findings have shown that *npc1a* mutants display germ cell migration defects and that feeding high levels of cholesterol can mitigate migration defects. The proposed experiments aim to confirm and extend the role of Npc1 family proteins in both PGC migration and Hh signal transduction pathway. We will test the model that npc1a and npc1b are involved during transport of 'cholesterol modified' Hh ligand in different developmental contexts.

This research was made possible by the generous support of the Crecca '46 Molecular Biology Senior Thesis Fund.

Down-Regulation of syntaxin 1A Gene in Bombus impatiens Using dsRNA

Katherine J. E. Stiefel
Advisor: Sarah Kocher

Bees are an excellent model system for studying the genetic mechanisms underlying behavior because they harbor extensive variation in social behavior both within and between species. Previous work from our research group suggests that variation in a regulatory region for the gene, syntaxin 1a (syx1a), is associated with the loss of eusociality within a single bee species. To functionally test if changes in expression levels of syx1a can alter social interactions in bees, we designed an RNAi experiment in the eusocial bumble bee, Bombus impatiens. We generated dsRNA sequences of syx1a and non-native yellow uorescent protein (yfp) as a control and injected it into the mushroom bodies of nurse bees from a single colony. We expected this treatment to down-regulate syx1a expression in the brains of injected bees. To evaluate the efficacy, we used quantitative, reverse transcription PCR (qRT-PCR). In the future, we will combine this approach with individual behavioral tracking assays to determine if the modulation of syx1a expression in adult bees alters social interactions within the hive.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Feedhorn Characterisation and Sensitivity Calculation of the AdvACT Bolometer Array

Shiye Su
Supervisor: Suzanne Staggs

At 10^{10} – 10^{11}Hz, Cosmic Background Radiation (CMB) remnant from the last scattering pervades the universe with remarkable large-scale homogeneity and isotropy. CMB polarisation and anisotropy measurements provide constraints on universal parameters such as the tensor to scalar ra- tio, the sum of neutrino masses, and dark matter. AdvACT is to be outfitted with four polarisation- sensitive TES bolometer arrays, whose radiation response profiles are largely determined by the geometry and material of the feedhorn wafers. Simulating a single feedhorn in ANSYS HFSS, we find as expected that signal propagation strongly resembles the fundamental TE_{11} mode of a cir- cular waveguide, and that response is maximised around the line of symmetry at azimuthal $\theta = 0$. A frequency sweep finds the current design to be most sensitive at 110 GHz (reflection coefficient minimised at -40 dB). Beam response patterns examined for 80 − 160 GHz, corresponding to the MF array's target range, generally reveal wide central maxima and shoulder features. Further, we characterise the net bolometer sensitivity from inputs including properties of the optic compo- nents, bath and critical temperatures, and test results, along with the Pardo atmospheric model, to quantify detector noise by per-detector and total array noise equivalent temperatures (NETs). The characterisation of radiation response and sensitivity allows for detector calibration and design optimisation.

This research was made possible by the generous support of the Princeton University Physics Department and the Lewis-Sigler Institute for Integrative Genomics.

Determination of PqsE active sites in *Pseudomonas aeruginosa*

Camille Sullivan
Sampriti Mukherjee, Justin Silpe
Advisor: Bonnie Bassler

Quorum sensing (QS) is a cell density based chemical communication system used by bacteria to control group behaviors. Accumulation of chemicals (autoinducers) secreted by the bacteria coordinate expression of virulence and biofilm genes. This synchronizes bacterial responses to population changes and environmental cues. Bassler laboratory specializes in studying QS pathways and previous work has been done using *Pseudomonas aeruginosa (Pae)* as a model organism. *Pae* utilizes three interconnected QS system pathways to regulate its virulence; one pathway is the PQS pathway, where the autoinducer PQS is synthesized by the enzymes encoded by the *pqsABCDE* operon. The protein PqsE has been discovered to be dispensable for the biosynthesis of the PQS autoinducer. PqsE however is essential for *Pae* QS as a *pqsE* null mutant fails to produce QS regulated public goods such as rhamnolipids and phenazines. How PqsE regulates QS is unknown. Through this project I will determine the residues of PqsE protein that are required for QS. I am performing random mutagenesis of the *pqsE* sequence and observing ensuing levels of the downstream pyocyanin and rhamnolipid production outputs. Mutant strains have been obtained in residues near the N-terminus of the protein and Western blot analyses will be done to determine if the mutants obtained produce stable proteins. A better understanding of the functions and roles of PqsE in QS will introduce new knowledge about the mechanisms of QS which could be used in the future when creating new antibiotics.

This research was made possible by the generous support of the Crecca '46 Molecular Biology Senior Thesis Fund.

Optogenetic disruption of neural activity in *C. elegans*

Xiaoting Sun
Advisor: Andrew Leifer

Animal behavior can be encoded in several ways in the brain, representing a wide range of time-scales — for example, neural activity, neural excitability, and synaptic strength. Neural activity changes almost instantaneously, while changes to synapses occur on the order of minutes. Behavior also extends beyond short term movements to long term states such as mating and hunger, but it remains to be seen how behavior on different time-scales maps onto brain states on different time-scales. Recent advances in whole-brain imaging systems for the model organism *C. elegans* have shown correlations between neural activity and locomotion, in particular for reversals. Our goal was to build an imaging set-up capable of optogenetic stimulation in order to test the dependence of *C. elegans* short term locomotion behavior, specifically omega turns and reversals, on neural activity. If the worm's locomotion is encoded in neural activity, disrupting this activity should cause it to forget its behavioral state. Preliminary observations were made on extra-chromosomal transgenic worms expressing pan-neuronal Arch, an inhibitory light-gated ion channel, or Chrimson, an excitatory light-gated ion channel, as candidates for the 'brain reboot' experiment. Further observations on the Arch worm suggested that disrupting neural activity decreases the rate of behavior continuation. These promising results encourage continued research using this approach — the next steps are integrating the Arch strain and creating a pipeline to automate the stimulation and behavioral analysis.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics as well as guidance from Andrew Leifer, Anuj Sharma, Mochi Liu, and Milena Chakraverti-Wuerthwein.

Analysis of Cell Growth and Cycle Regulation in *D. melanogaster* Nurse Cells

Caroline Taber
Caroline Doherty
Advisor: Stas Shvartsman

Overall, little is known about the mechanisms that control a cell entering different phases of its life. The roles of cyclin dependent kinases (cdks) have been studied, but there is much more to be understood. To better study these controls, there are some cells do not follow the typical cell division pattern pattern, but rather cycle between periods of DNA replication/growth and rest without any divisions. These cells undergo the endocycle, and become polyploid. One example of such cells are the nurse cells within the eggs chambers of *D. melanogaster*. These cells are of particular interest due to the way in which they come to be - each egg chamber begins with one cytoblast which undergoes four incomplete divisions, creating fifteen nurse cells (and an oocyte) connected by ring canals. Due to the presence of these ring canals and the nature of the divisions, the order of creation of these cells can be determined, and therefore they can be numbered one through sixteen, allowing for comparisons of specific nurse cells between egg chambers. These egg chambers are ideal for analyzing the question of what it is that is controlling the periods of growth and rest within these cells. To study these cells, the ovaries of *D. melanogaster* are stained with DAPI and Phalloidin, which show the DNA content and F actin, respectively. Antibodies such as Phosphotyrosine (for ring canals) and Nuclear Pore Complex (for the nuclear envelope) are also utilized and tagged with fluorescent secondary antibodies. The ovaries are then imaged under a confocal microscope to create Z-stacks of images which are then used in conjunction with the program IMARIS, which allows for the 3D reconstruction and labeling (one through sixteen) of the nurse cells, as well as using the DAPI signal to locate nuclei, and the Phalloidin and Phosphotyrosine for the ring canals. IMARIS is also used to find the volume of the nurse cells, and is used to quantify the intensity of the DAPI signal - which can be normalized by being compared to the signal from diploid follicle cells of neighboring egg chambers. This information is then used to find the nurse cell size versus nurse cell number at various egg chamber stages, as well as normalized DNA content versus nurse cell size. Results so far show that there are general trends in size and normalized DNA content when nurse cells are compared based on how many ring canals they are away from the oocyte. Nurse cells 2, 3, 5, and 9 (those most proximal to the oocyte - only one ring canal away) all tend to be larger as well as contain more DNA than the other nurse cells, and can up to an entire endocycle ahead in terms of DNA content. Nurse cell 16, the most removed from the oocyte (4 ring canals away), tends to be the smallest and have the fewest copies of DNA. Future directions to understand what is controlling the rate of endoreplication in the different nurse cells include inducing mutations in Myc (whose levels have been shown to correlate with cell proliferation), as well as in the Hippo pathway and the effector Yki, which is known to phosphorylate cyclin E, the only cyclin necessary for the endocycle. Along with this, concentrations of transcription factors in various cells may be analyzed to see what roles, if any, they may play.

Research reported in this abstract was supported by a grant from the National Institutes of Health under award number R01HD085870.

Creating Hidden Markov Models for Transport Protein Classification in Metagenomics Data

Tanya V. Tafolla
Anna Posfai
Advisors: Ned S. Wingreen and Mohamed S. Abou Donia

The microbiome has gained a great deal of attention recently for its role in human health. Hundreds of bacterial species are present in our bodies, participating in activities such as breaking down nutrients and conditioning our immune systems. To better understand interactions between bacteria in the microbiome and their interactions with the host, we are studying transport proteins. There are thousands of known bacterial transporters and multiple superfamily categories. Our goal is to accurately and efficiently classify new protein sequences from metagenomics data into transporter classes and to further classify members of superfamilies such as the ATP-binding Cassette (ABC) into substrate specific clusters. In order to classify new sequences of transport proteins, we create probabilistic profiles of known transport-protein clusters using Hidden Markov Models (HMMs). Other profile software algorithms including BLAST do not take into account position specific information such as insertions or deletions in sequences. For our study, being able to consider insertions and deletions is crucial because homology between proteins can be difficult to spot. Thus, we created a database of HMMs for the known transport-protein clusters using HMMER [HMMER 3.1b2 (February 2015); http://hmmer.org/]. Proteins are then scored and placed into appropriate clusters. HMMs have proven to be accurate for clusters that have closely related sequences, whereas further training must be done for HMMs of clusters with more sequence variation. Next, we aim to classify test sequences from metagenomics data against our HMM database.

This research was made possible by the generous support of The Graduate School at Princeton University.

An Investigation of Distal Enhancer-Enhancer Communication for the *brinker* Locus in Drosophila

Mark-Avery Tamakloe
João Raimundo
Advisor: Michael Levine

In metazoans, the transcription of many developmental patterning genes is highly regulated, both spatially and temporally, to ensure proper gene expression in the developing embryo. In *Drosophila melanogaster*, these genes often have multiple enhancers that are believed to work in pairs with overlapping activity. Preliminary evidence in the Levine Lab suggests that these paired enhancers do not necessarily work in the widely accepted primary and "shadow" enhancer model, but interact with each other additively. However, most of the studies on these genes have been done *in vitro* and not endogenously, and thus the exact mechanism for how these enhancers communicate is not well understood. To further investigate the potential additive property of enhancers, we performed knockouts of both the early acting and late acting enhancers for the *brk* locus. We found that knocking out either enhancer seems to lead to lethality for half the male population as evidenced by the 2:1 ratio of females to males in both knockout lines. We will temporally monitor the interaction between these two distally located enhancers by tracking the early expression of the *brk* gene and measuring the corresponding levels of nascent *brk* transcripts utilizing the novel live imaging technique of the MS2 reporter system. Understanding the fundamental mechanisms underlying enhancer communication in the regulation of gene expression can provide essential insight in revealing the origin of human diseases that are triggered by abnormalities in the complex multiple enhancer regions.

This research was made possible by the generous support of the Crecca'46 Molecular Biology Senior Thesis Fund.

The Enhancer Effect: Investigating the Role of Non-coding Mutations in Cancer

Alice Tang
Pawel Przytycki
Advisor: Mona Singh

With the declining cost of genome sequencing comes a surge in sequencing data for not only populations of various species, but also various tissues within a given species. Within the scope of these efforts, limiting analysis to data generated from Whole Exome Sequencing (WES/WXS) has become the general practice due to cost and long-held beliefs that, aside from protein-coding genes, the genome otherwise consists of "junk" DNA. However, less than 2% of the human genome encodes protein. Non-coding regions have been shown to serve critical roles, ranging from gene regulation to RNA processing, yet are not well studied in cancer research.

Of non-coding regulators, enhancers are of particular interest due to their role in complex regulatory mechanisms that influence genes controlling cell growth—a key component of cancer. We investigated non-coding mutations in known enhancer regions for 130 Skin Cutaneous Melanoma (SKCM) patients, a common cancer with known recurrent non-coding mutations (e.g. in the TERT promoter), to nominate cancer-driving regions and mutations. Through complementing mutational analysis with patient-specific mutational expression data, we derived preliminary candidate lists that appear enriched for genes with Gene Ontology (GO) labels relevant to melanosomes and for genes from the Melanoma Gene Database and the Cancer Gene Census. Ultimately, since non-coding driver mutations affect the expression of protein-coding genes, identification of non-coding driver mutations will facilitate therapeutic efforts by recommending target regions for treatment.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics.

Understanding outer membrane protein assembly in *E. coli* through suppression of ß-signal mutant LamB^{G439D}

Kimberly Tang
Betsy Hart
Advisor: Thomas Silhavy

The outer membrane (OM) of Gram-negative bacteria acts as a first line of defense against toxic molecules in the external environment, including antibiotics. Outer membrane proteins (OMPs) are a major component of the OM that regulates this defense. However, the assembly of the OM and the pathways that regulate its maintenance and formation are poorly understood. In *E. coli*, the Bam complex folds and inserts OMPs into the OM. Previous work has shown that the Bam complex recognizes unfolded OMP substrates by a C-terminal ß-signal sequence. Preliminary work in the Silhavy lab has identified an assembly defective ß-signal mutant in the maltoporin LamB, *lamB$_{G439D}$,* that is suppressed by the deletion of the periplasmic protease *degP* at 30ºC. Intriguingly, *degP* deletion is lethal in this background at 37ºC, suggesting that unfolded substrate accumulates in the periplasm and creates a toxic backup of unfolded OMPs. This temperature sensitivity has allowed us to select for suppressors at the non-permissive temperature of 37C that are able to overcome the toxic accumulation of unfolded OMPs. Furthermore, we have utilized MacConkey indicator media to detect LamB assembly in these suppressors using dextrin uptake as a readout. Using this approach, we have isolated twelve suppressors of *lamB$_{G439D}$* that will provide the opportunity to understand the interactions that take place on the OMP assembly pathway.

This research was made possible by the generous support of the Crecca '46 Molecular Biology Senior Thesis Fund.

Investigating the Relationship of Fibronectin and LOX Family Proteins during Matrix Assembly

Sarah J. Tian
Katherine E. Hill
Advisor: Jean E. Schwarzbauer

Fibrosis is characterized by the excess deposition of extracellular matrix (ECM) components in an organ or tissue and increased cross-linking. Fibronectin (FN) is a key glycoprotein in ECM and regulates cell function in conjunction with other ECM proteins such as collagen. Furthermore, collagen matrix assembly is dependent on the presence of fibronectin. The lysyl oxidase (LOX) enzyme is responsible for collagen cross-linking during matrix formation. Two isoforms, LOXL1 and LOXL2, are also known to have cross-linking functions. Presently, the effect of the fibronectin matrix on collagen matrix assembly is not well understood. This study aims to understand the relationship between FN and LOX family proteins in an effort to clearly elucidate a role for the enzyme in matrix assembly. HT1080 fibrosarcoma cells were treated with dexamethasone to stimulate FN matrix assembly. HT1080s were shown to express LOX and LOXL2, but not LOXL1. mRNA expression of FN and LOX family proteins is measured by qPCR. Stimulation with dexamethasone caused a change in expression levels. A time course was developed to show how expression levels differed after 24, 48 and 72 hours of treatment wit Protein expression will be quantified by Western blot at a later time. This study will provide insight into the interaction between FN and LOX family proteins which prove useful as a target in the therapeutic treatment of fibrotic diseases and cancer progression.

This research was made possible by the generous support of the Evnin '62 Molecular Biology Senior Thesis Fund.

Establishing a Mechanism for Stress-Induced Telomere Shortening in Human Foreskin Fibroblasts

Ava Torjani
Lisa Schneper, Iulia Kotenko
Advisor: Daniel Notterman

Stress is associated with harmful outcomes including depression, cancer and cardiovascular disease. Previous studies have shown that the stress hormone, hydrocortisone, accelerates telomere attrition in human foreskin fibroblasts (HFF) and reduces telomerase activity in human T-lymphocytes. However, the link between telomerase activity and telomere length is unclear, and the mechanism behind the effects of stress on telomerase activity is unknown. This study aims to determine the effect of hydrocortisone treatment on telomere length and the mechanism of how hydrocortisone alters telomerase activity in HFF cells. Previously, the laboratory showed that exposure of HFF cells to hydrocortisone decreases telomerase activity and alters expression of telomere maintenance genes. The current study will confirm these findings by examining the effect of hydrocortisone on HFF telomerase activity using Telomere Repeat Amplification Protocol (TRAP), and subsequently on telomere length using qPCR. Additionally, altered expression of potential hydrocortisone targets, including non-coding RNA (HSAT-II), and the telomerase components, the RNA template (TERC) and reverse transcriptase (TERT), will be examined by qPCR. We confirmed a decrease in telomerase activity in response to hydrocortisone, warranting further research to confirm significance. We also expect that hydrocortisone treatment will be associated with shorter telomeres and differential expression of selected targets.
Future experiments will examine the role these targets play in altering telomerase activity in response to hydrocortisone using RNAi. These results will provide a foundation to elucidating the mechanism behind stress-induced telomere shortening.

This research was made possible by the generous support of the Evnin '62 Molecular Biology Senior Thesis Fund.

Modeling the fluorescence signal of transcribed loci from live imaging

Gabriel Toneatti Vercelli
Benjamin Zoller, Hugh Wilson
Advisor: Thomas Gregor

Live imaging enables characterization of transcriptional dynamics at the single-cell level by fluorescent labeling nascent mRNA in vivo. Recent advances in live imaging techniques such as the MS2/PP7 system have been critical for understanding transcription regulation in developing organisms. However, it remains unclear to what extent the measured fluorescence intensity correlates with nascent mRNA abundance. Here, we address this question by performing series of measurements in early Drosophila embryos. In our experiments, we used the PP7 stem loop system, creating constructs with 8, 16, 24, 32 and 48 loops inserted in a reporter driven by the hunchback P2 promoter. Using 2-photon microscopy, we measured the variability in fluorescence intensity among different embryos of the same construct and assessed how the mean fluorescence intensity scales with the number of stem loops. Assuming that fluorescent proteins saturate the stem loops, the signal should scale quadratically with the number of inserted loops, since both the length of transcripts and the number of binding sites increases significantly. However, our results support a sublinear dependence of the intensity on the number of stem loops, suggesting inefficient binding or insufficient concentration of fluorescent proteins. Future work varying the embryos' concentration of proteins fused with fluorophores and a careful sampling of embryos with less than 8 loops will shed some light on the behavior of this system at low and high fluorescence intensities.

This work was supported by MOL/QCB Summer Undergraduate Research Program.

***roX1* and *roX2* Chromatin Entry Sites: Dosage Compensation and Boundary Elements**

Aidan Waugh
Amina Kurbidaeva
Advisor: Paul Schedl

Different organisms use different mechanisms to equalize expression of X chromosome genes. In *Drosophila melanogaster*, transcription of X-linked genes in male flies (X/Y) is upregulated to match the level of expression in females (2X). The Male specific lethal complex (MSL) is responsible for dosage compensation in males. The MSL complex solely associates with active genes on the X-chromosome. Key to targeting the MSL complex to the X-chromosome are specific loci called Chromatin Entry Sites (CES). It's thought that there are more than 100 CES along the X chromosome and that MSL complexes are initially recruited to these loci and spread to nearby genes. The LBC protein complex binds to several different CES, including the canonical *roX1* and *roX2* CES. Since LBC is also a component of a Bithorax complex boundary (insulator) element called Fab-7, this discovery raised the possibility that *roX1* and *roX2* could also act as boundary elements. Boundary elements define the 3D chromosomal topology by pairing to each other and this can be demonstrated using "transvection" assays. In collaboration with the Jaynes lab (Thomas Jefferson University) we generated a set of attP transgenes containing *roX1* or *roX2* CES flanked either by enhancers or lacZ and GFP reporters. I've generated stage 13 embryos carrying the enhancer/CES transgene on one homolog and the reporter/CES transgene and probed for lacZ expression using a smFISH procedure followed by confocal microscopy. I'm refining my protocols in order to improve the image quality. We hope to show that *roX1* and *roX2* loci display pairing activity similar to known Drosophila insulators. This would help us learn more about the topology of the X chromosome and MSL spreading mechanism.

This research was made possible by the generous support of the Elkins Family Molecular Biology Senior Thesis Fund.

Investigation of Phospholipid Trafficking to the Outer Membrane in *Escherichia coli*

Maria Wissler
Jacqueline Grimm
Advisor: Thomas Silhavy

Gram-negative bacteria have cell envelopes made up of an inner membrane, a peptidoglycan cell wall, and an outer membrane. The outer membrane is made of lipopolysaccharides, β-barrel proteins, lipoproteins, and phospholipids. The mechanisms of trafficking lipopolysaccharides, β-barrel proteins, and lipoproteins are well studied and understood. However, there is a gap in knowledge regarding the trafficking of phospholipids. *E. coli* inner membrane protein YejM is homologous to the cardiolipin transporter PbgA in *Salmonella*, and is essential even though cells can survive without cardiolipin. Previous studies have associated mutant *yejM* with defects in the outer membrane. The current study aims to investigate the possible role of YejM in phospholipid trafficking. A truncated version of *yejM* is temperature sensitive at 42°C. Spontaneous suppressors of the mutant *yejM* temperature sensitivity phenotype are currently being isolated. These mutations in associated genes will be identified and characterized to better understand the function of *yejM*. I am also recombineering a new strain with a *yejM* truncation carrying a chloramphenicol marker at the truncation site. Once constructed, this strain will be used with the EZ-Tn5 transposome to create a library of transposon insertions. Transposon insertion sequencing will then be used to identify genes that are synthetic lethal with mutant *yejM*. This study is undertaken to better understand phospholipid trafficking to the outer membrane of Gram-negative bacteria. Phospholipid trafficking is the last largely unknown component in assembly of the outer membrane, and is important for basic biological knowledge of bacterial development and may be useful in targeting infectious bacteria.

This research was made possible by the generous support of the Pitts '63 Molecular Biology Senior Thesis Fund.

Effects of Environmental Conditions on *Vibrio cholerae* Curvature and Potential Implications

Stevie Yang
Nicholas Martin
Advisor: Zemer Gitai

Cholera is the leading diarrheal disease worldwide and is caused by the pathogenic bacterium *Vibrio cholerae. V. cholerae* has a characteristic curved rod shape that provides an advantage for swimming motility and pathogenesis. The degree of curvature is dynamic, and it can be perturbed by mutation of the curvature determinant *crvA* or growth in high osmolarity media; however, the effects of other environmental conditions on *V. cholerae* have not been thoroughly investigated. The objective of this research is to discover new environmental conditions that change the dynamics of *V. cholerae* curvature and to understand their underlying mechanisms, with an ultimate goal of uncovering how these conditions change *V. cholerae* life cycle. By screening *V. cholerae* curvature in a variety of growth conditions, I have identified new carbon sources that alter curvature. Preliminary data show that growth rate may contribute to curvature changes. These findings also strengthen the negative correlation between osmolarity and curvature. Since motility is important for *V. cholerae* virulence, future motility assays will be performed on *V. cholerae* grown under the curvature-altering conditions. This research will better inform the scientific community on the curvature dynamics of *V. cholerae* and its relationship to pathogenic behavior. These results can potentially offer insight on how to target virulence in bacteria based on cell shape without selecting for resistant bacteria with antibiotics.

This research was made possible by the generous support of the Rupert and Loretta Jones Molecular Biology Senior Thesis Fund.