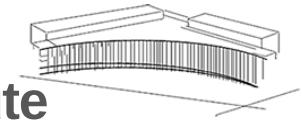
Department of MOLECULAR BIOLOGY

Lewis-Sigler Institute for Integrative Genomics



Summer Undergraduate Research Program

Molecular and Quantitative & Computational Biology

Poster Session

Abstract Booklet

2018

PRINCETON UNIVERSITY

Student Name	Advisor	Institution	Title	Page #	Poster #
Jacob Adkins	Petry	Princeton University	Investigating the role of XMAP215's C-terminus in microtubule nucleation with gamma-tubulin	8	1
Nicholas Archer	Ploss	Princeton University	Investigating the Host Factors Determining the Varied Host Tropism of Hepatitis E Virus	8	2
Ruchita Balasubramanian	Gregor	Princeton University	Characterizing the Network: Elucidating the dynamics of the Gap Gene Regulatory Network	9	3
Destiny Batton	Nelson	University of California-Los Angeles	Investigating the role of pressure-sensitive signaling in lung branching morphogenesis and epithelial cell differentiation	9	4
Julia Casazza	Murphy	Princeton University	*repx-1* regulates *Caenorhabitis elegans* reproductive span	10	5
Daniel Che	Buschman	Princeton University	Developing an Adaptive Cognitive Prosthetic: Closed-loop multisite electrical stimulation produces specific neural firing patterns	10	6
Shuqi Chen	Wang	Zhejiang University	Optimizing uDISCO-based clearing for signal preservation and increased tissue transparency	11	7
Zhenhong Chen	Korennych	Howard University	Determining the activating mechanism of GCN2 kinase	11	8
Shuyuan (Kat) Cheng	Kang	Zhejiang University	Screening of Immunoregulatory Genes in Breast Cancer Therapy	12	9
Annie (Yoolim) Choi	Levine	Princeton University	Investigating PIWI regulation in the central nervous system of Ciona intestinalis	12	10
Leo Choi	Murphy	Princeton University	Dietary Supplementation of Oleic and Linoleic Acid Rescues Mating-Induced Death in *Caenorhabditis elegans*	13	11
Farion Cooper II	Kocher	Howard University	Computational paleovirology reveals previously undescribed taxa present in the viral communities of bees	13	12
Daniela Coronado	Bassler	Princeton University	Regulation of an alternative quorum-sensing pathway in Vibrio cholerae	14	13

Student Name	Advisor	Institution	Title	Page #	Poster #
Joseph Crapse	Wuhr	Princeton University	The Arrhenius Equation and Complex Reaction Networks: Animal Development	14	14
Kerri Davidson	Burdine	Princeton University	Identifying motile ciliated cell types required for maintaining spine straightness in zebrafish	15	15
Sofia Dimitriadoy	Kang	Princeton University	Elucidating the Role of METTL14 in Breast Cancer Tumor Initiating Cells	15	16
Ashley Drengler	Schwarzbauer	Princeton University	Effect of a Novel Fibronectin Mutation on Collagen Processing and Matrix Assembly	16	17
Kerry Farlie	Notterman	Princeton University	Investigating the Role of Progesterone on Telomerase Activity and Telomere Length in Mice with Calibrated Traumatic Brain Injury	16	18
Courtni Foster	Troyanskaya/Levine	Howard University	Phenotyping Non-Coding Autism Related SNVs in Embryonic Stem Cells	17	19
Samuel Garfinkle	Hyster	Princeton University	Structure based N-methyl chloroamide cyclase design using OYE-family enzymes	17	20
Valeria Gerena	Silhavy	University of Puerto Rico, Rio Piedras	Investigating the Role of MlaA-OmpF/C Interaction in the Mla Pathway	18	21
Jessica Goehring	Donia	Princeton University	Microbiome-Derived Metabolism of Anti-Infectives: Identifying the Human Gut Microbiome Isolates Responsible and Measuring Changes in Bioactivity *Ex Vivo*	18	22
Raymond Guo	Ploss	Princeton University	The impact of spatial transcriptional differences in the liver on hepatitis delta virus infection	19	23
Kristian Harris	Gitai	Howard University	Mutating the Primary Sequence of the CrvA Protein in *Vibrio cholerae*	19	24
Zhi-Shan Hsu	Enquist	Princeton University	A Functional Comparison HSV-1 and PRV VP16 on Neuronal Gene Expression	20	25
Yechen Hu	Norman	Princeton University	Can sleep rescue memories from retrieval-induced forgetting?	20	26

Student Name	Advisor	Institution	Title	Page #	Poster #
Danielle Isakov	Toettcher	Princeton University	Investigating the Relationship between Endogenous Erk Dynamics and Downstream Cell-fates in Primary Mouse Keratinocytes	21	27
Sophia Jackman	Akey	Harvard University	Frequency and Patterns of Cancer Causing Somatic Mutations in 14 Healthy Individuals	21	28
Benjamin Jacobson	Notterman	Princeton University	Investigating a Link between Child DNA Methylation and Exposure to Fracking	22	29
Artem Khan	Rabinowitz	Princeton University	Absolute Concentration of Metabolites in *Saccharomyces cerevisiae*	22	30
Joyce Lee	Yan	Princeton University	Single-Particle Cryo-EM of the human NPC1/NPC2 complex	23	31
Sally (Ju Young) Lee	Burdine	Princeton University	Investigatng the activation and asymmetric regulation of *dand5* in zebrafish	23	32
Sarah Lee	Hughson	Princeton University	Investigating the Interaction Between the Sec1/Munc18 (SM) Protein Vps33 and its Qa-SNARE Vam3 and R-SNARE Nyv1	24	33
Lian Kirit Limperis	Notterman	Princeton University	Control of Translational Repression Through a Light-Inducible Clustering System	24	34
Gabe Lipkowitz	Ploss	Princeton University	Analysis of host responses to hepatitis B and delta viral infections in a micro-scalable hepatic co-culture system	25	35
Chaozhong Liu	Notterman	Zhejiang University	Apply European ancestry-based GWAS data to height prediction for Africa Ancestry by Gradient Boosted Regression Trees	25	36
Jelani Lyda	Levine, Mike	Howard University	Interactions between enhancers in the regulation of even-skipped transcription in *Drosophila* embryos	26	37
Christina Matl	Wang, Samuel	Princeton University	The role of Purkinje cells in Cognitive and Affective Development	26	38
John McEnany	Wingreen	Princeton University	Simulation of *Myxococcus xanthus* Motility in a Wetted Environment	27	39
Isabel Medlock	Tamir	Princeton University	The Role of Emotion Dynamics in Mental Models for Social Prediction	27	40

Student Name	Advisor	Institution	Title	Page #	Poster #
Kennedy Miller	Buschman	Princeton University	Investigating the therapeutic potential of low-dose ketamine in an *in utero* valproic acid mouse model of Autism Spectrum Disorder.	28	41
Noam Miller	Norman	Princeton University	Leabra7: A Python Package for Hippo-cortical Interactions in Learning	28	42
Janelle Nelson	Wingreen	Howard University	The Statistical Physics of Intrinsically Disordered Proteins	29	43
Phoebe Nelson	Petry	Florida International University	Structural Determination of Microtubule Nucleation Factors	29	44
Eva Parisi	Muir	Princeton University	How do cells create enhancers? Inducing H3K4me1–H3K27ac crosstalk at enhancers in live cells with protein trans-splicing	30	45
Olivia Parker	Avalos	Princeton University	Optogenetic Regulation of Biofiuel Pathways: Control of Isopentanol and 2-methyl-1-butanol Biosynthesis with Blue Light	30	46
Brooke Phillips	Devenport	Princeton University	Investigating the Novel Mutation, *rosette*, as a Potential Global Cue in the Core Planar Cell Polarity Pathway	31	47
Anagha Prasanna	Muir	Princeton University	Investigating the Role of PARP1 in the Context of Oncohistone Mutations	31	48
Justin Ramos	Gitai	Princeton University	Identification of Novel *Caulobacter crescentus* Width and Length Genetic Determinants Using a Transposon Library Screen	32	49
Ellie Randolph	Schwarzbauer	Princeton University	Requirement for Heparin Interactions with Fibronectin Module III-13 during Polymerization and Cell Association	32	50
Jose Reyes	Wuhr	University of Puerto Rico, Rio Piedras	Investigating the proteome partitioning between the nucleus and cytoplasm in using DNA coupled magnetic beads	33	51
Denay Richards	Gavis	Princeton University	Dendritic Disasters : Discovering the role of Muscleblind in Dendritic Arbor Formation	33	52

Student Name	Advisor	Institution	Title	Page #	Poster #
Joseph Ryu	Troyanskaya	Princeton University	Genetic and Biochemical Analysis of MAB21 Proteins in Neuroblastoma Proliferation	34	53
Nicholas Schmeller	Mittal	Princeton University	Trust but Verify: Analysis of Flawed Trust Beliefs for Tor and Optimal Trust Modeling	34	54
Yoni Schoenberg	Shenk	Princeton University	Developing a Tool to Organize and Annotate Gene Expression Metadata	35	55
Sara Shatela	Gavis	California State University-Northridge	Investigating the relationship between mRNA localization and fertility in *Drosophila*	35	56
Jessica Sheng	Silhavy	Princeton University	Characterizing the mechanism of the Bam complex during outer membrane protein assembly	36	57
Anoopkumar Sonar	Majumdar	Princeton University	PAC-Bayes Control: Synthesizing Controllers that Provably Generalize to Novel Environments	36	58
Yanilka Soto	Devenport	University of Puerto Rico, Rio Piedras	Mapping the adhesion profile of hair follicles during placode polarization	37	59
Ashley Stone	Zakian	Princeton University	Characterizing the functions of motifs A, B, and C and identifying the nuclear localization signal in *Saccharomyces cerevisiae* Pif1 helicase	37	60
Samvida Sudheesh Venkatesh	Cristea	Princeton University	Mapping intracellular HCMV infection using a computational and proteomic approach	38	61
Dominique Summerville	Bassler	Princeton University	Autoagglutination of O1 biovar El Tor *Vibrio cholerae* strain	38	62
Tingting (Xiaoting) Sun	Wang	Princeton University	Automated tracking of freely-moving mouse behavior	39	63
Amy Tien	Cristea	Princeton University	Sirtuin-3 vs. pUL37 in mitochondrial changes during human cytomegalovirus infection	39	64
Annan Timon	Toettcher	Princeton University	Optogenetic Control of Apical Contraction in the *Drosophila* embryo	40	65

Student Name	Advisor	Institution	Title	Page #	Poster #
Awele Utomi	Scharzbauer	Howard University	Investigating Expression of Tenascin-C Isoforms in Diabetic Conditions	40	66
Kelly Van Baalen	Jonikas	Princeton University	Localizing the Most Highly Conserved Proteins in Photosynthetic Organisms	41	67
Nancy Wenger	Levine	Princeton University	Determining a role for *Lmx* in neural tube formation	41	68
Lillian Wilkins	Schedl	Princeton University	Functional analysis of *germ cell less* during germ cell formation and specification in *Drosophila melanogaster* embryos	42	69
Brittany Williams	Toettcher	University of Arizona	Increasing CRISPR-Mediated Homologous DNA Recombination Efficiency	42	70
Lucy Williamson	Donia	Princeton University	The Antimicrobial Role of Candidate Biosynthetic Gene Clusters from *Actinomyces* Strains within the Oral Human Microbiome	43	71
Scott Wolf	Ayroles/ Shaevitz	University of Arkansas, Little Rock	RATrak: A Novel Software Suite for the Analysis of *Drosophila melanogaster* Locomotor Data	43	72
Evelyn Wu	Ploss	Princeton University	Identification of antiviral signaling pathways restricting hepatitis delta virus persistence	44	73
Yihua Xie	Jonikas	Princeton University	Determining the localization of uncharacterized proteins required for photosynthesis	44	74
Alondra Arciga	Shaevitz	University of Illinois at Urbana-Champaign	3D Tracking Calibration for Biplane Microscopy		75
LeeRoy Borders	Astro	Rutgers University	Quasars		76
Jeremy Caparotta	Stone	Rutgers University-Newark	Promoting cell-adhesion in microfiber hydrogels		77
Eyoel Demissie	Puchalla	Bunker Hill Community College	Chaperone Mediated Protein Aggregation Using Agent-Based Modelling		78
Mouhammed Nassir Diagne	Linear and Nonlinear Photonics Group	Illinois Institute of Technology	Phase retrieval		79

Student Name	Advisor	Institution	Title	Page #	Poster #
Pedro Fernandez, JR and David Garcia	Pelczer	Harold Washington College and Rutgers University	NMR Analysis of Bodily Fluids		80
Boris Franklin	Psychology	Rutgers University	Diversity Framing in Photos		81
Ryan Harvey	Yan	Manchester Community College	Purification of the Beta Subunit for Cav 1.2		82
Paul Kazelis	Niv	Rutgers (New Brunswick)	The psychological space that represents numerosity is scaled logarithmically		83
Toluwalase Olusola	Jonikas	Camden County College	Revealing the reason for proteins that compose a microcompartment called Pyrenoid in the single cell green alga Chlamydomonas reinhardtii Using bioinformatics		84
Chinweoke Onejeme	Prud'homme	Camden County College	Evaluating the Effects of Nanoparticle Assembly on Enzyme Activity		85
Kofi Jose Ries, Sr.	Stigma and Social Perception Lab	Rutgers University	Diving Deeper into Diversity: How does the Way Postsecondary Schools Frame Diversity Affect Student Outcomes		86

Investigating the role of XMAP215's C-terminus in microtubule nucleation with gamma-tubulin

Jake Adkins, **Sabine Petry**

The proper nucleation of microtubules is essential to the formation of a variety of cellular structures, including the mitotic spindle. For several decades, gamma-tubulin has been considered to be the universal microtubule nucleator of the cell. Recently, however, work from our lab has demonstrated that XMAP215, the microtubule-associated protein responsible for catalyzing microtubule growth via its TOG domains, is also necessary for proper microtubule nucleation. During microtubule nucleation, the ill-defined C-terminus of XMAP215 interacts directly with gamma-tubulin and is necessary for wildtype levels of nucleation to occur. However, the exact position and function of this interaction with gamma-tubulin remains unknown. I hypothesize that a newly-predicted TOG domain in the XMAP215 C-terminus is responsible for this interaction. To investigate this potential interaction, I have designed and optimized the purification of a series of truncated XMAP215 C-terminal constructs. Using these constructs and purified gamma-tubulin, I will investigate the position and functional role of the XMAP215-gamma-tubulin interaction through a series of pull-down, size-exclusion chromatography, and microtubule nucleation visualization assays in *Xenopus laevis* egg extracts. Purification of these constructs will also enable the use of X-ray crystallography to establish an atomic-resolution 3D structure of the XMAP215 C-terminus, the structure of which remains to be established. Completion of this study will provide novel insights into the nucleation of microtubules, as well as serve to elucidate the role of the XMAP215 C-terminus in mitotic spindle formation.

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

Investigating the Host Factors Determining the Varied Host Tropism of Hepatitis E Virus

Nicholas F. Archer, Ila Nimgaonkar, Benjamin Y. Winer, **Alexander Ploss**

Each year, hepatitis E virus (HEV) infects ~20 million people and claims the lives of over 40,000. While mortality rates in healthy individuals are around 2%, HEV causes death in ~30% of infected pregnant women, as well as chronic hepatitis in the immunocompromised, such as organ transplant recipients and HIV-positive patients. HEV is transmissible via the fecal-oral route, and thus contaminated water is a major source of infection in developing countries. Additionally, zoonotic transmission of the virus occurs worldwide via consumption of undercooked pork meat. This mode of disease spread, along with nosocomial transmission associated with organ and blood donations, contributes to the status of HEV as a global threat to public health. A safe, highly effective, prophylactic vaccine against HEV has been developed, although it is currently licensed only in China. Moreover, the few explored medical treatments remain unsafe for pregnant women and fetuses, and have otherwise proven to be only partially effective. Despite HEV's wide host range, its study and preclinical evaluation of new treatments are hindered by a scarcity of susceptible small animal models. Consequently, we aim to elucidate the barriers to cross-species HEV infection, particularly in mice. We have designed a human-murine heterokaryon assay to discern whether positive human host factors or negative murine restriction factors may be responsible for the observed tropic limitations of HEV. Our preliminary data suggest that one or more endogenous negative restriction factors in murine cells account for the inability of HEV to complete its life cycle in a murine host.

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

Characterizing the Network: Elucidating the dynamics of the Gap Gene Regulatory Network

Ruchita Balasubramanian, **Thomas Gregor**

The gap gene network is a critical biological network that governs patterning development of *Drosophila melanogaster*. Previous studies utilized fixed-tissue methods to deduce macroscopic regulatory interactions in this network. As a result, the dynamics of this network could only be inferred by timing multiple fixed and therefore dead embryos and linking them to artificially construct a time-scale. This study aims to characterize gap gene expression endogenously in a single embryo in real-time in wild-type or mutant backgrounds, to elucidate the dynamic parameters of expression and the regulatory interactions that govern this network. We have developed a live-imaging approach using the MS2-MCP-GFP system to visualize the expression of multiple gap genes live in an endogenous context. We were able to show that more minute fluctuations in endogenous transcription of gap genes *knirps, Kruppel, hunchback,* and *giant* can be observed using this approach, validating the use of this system to capture dynamics of the network. We were able to characterize with a greater degree of temporal resolution, onset of transcription, fluctuation of activity over time, nuclei-to-nuclei variability in expression, and evolution of spatial domains of expression over time of *Kruppel, giant* and *knirps* in wildtype backgrounds. Finally, the evolution of the spatial domains of *Kruppel* and *knirps* expression in a single embryo could also be assessed. Future work will investigate the dynamics of regulatory interactions through real-time visualization of endogenous expression of all four canonical gap genes in mutant backgrounds, to obtain a more thorough understanding of the gap gene regulatory network in *D. melanogaster.*

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

Investigating the role of pressure-sensitive signaling in lung branching morphogenesis and epithelial cell differentiation

Destiny M. Batton, Katharine E. Goodwin, Sarah V. Paramore, and **Celeste M. Nelson**

Disrupted lung development in the embryo can lead to lung deficiencies which can result in lifelong respiratory ailments and diseases. A deeper understanding of lung development can provide insight as to why various lung conditions arise and inform how to improve disease outcomes. Lung development occurs through branching morphogenesis (BM), during which epithelial tube networks develop. The liquid flow between the luminal and pleural cavities of the embryonic lung creates a transmural pressure gradient. Recent work with embryonic mouse lungs demonstrates that this pressure gradient regulates BM, as increased transmural pressure results in increased growth of the epithelium. Further, RNA-sequencing data suggest that increased transmural pressure stimulates epithelial cell (EC) differentiation. Based on these findings, we hypothesize that pressure sensitive signaling within the lung stimulates EC differentiation. To test this hypothesis, we are first determining whether markers of epithelial differentiation are detectable at different stages of embryonic mouse lung development. Next, we are investigating two possible transducers of the pressure signal: YAP, a mechanosensitive transcription factor, and stretch-activated cation channels, including Piezo1. So far, we have examined the effects of YAP inhibition and Piezo1 inhibition on BM in E12.5 lungs. To inhibit YAP and Piezo1, we treated the lungs with verteporfin and spider venom toxin in ex vivo culture, respectively. In both cases, we find that the lungs undergo less branching morphogenesis than in the control, suggesting that YAP and Piezo1 may play a role in BM. Future work will examine the effects of YAP and Piezo1 on EC differentiation.

This research was made possible by the generous support of the Howard Hughes Medical Institute.

repx-1 regulates *Caenorhabitis elegans* reproductive span

Julia A. Casazza, Rebecca S. Moore, **Coleen T. Murphy**

Despite its significant social costs, reproductive aging remains poorly understood. A forward genetic screen conducted in the Murphy lab identified *repx-1* as a potential regulator of *Caenorhabditis elegans* reproductive longevity. Previous work also conducted in the Murphy lab showed that loss of *repx-1* extends self-fertilized reproductive span. To determine whether increases in self-sperm or oocyte quality account for this phenotype, we conducted a mated reproductive span, which uncouples self-sperm production from reproductive span. *repx-1* mutants reproduced for longer than a rescue strain, indicating that *repx-1*-mediated extension of reproductive span relies upon maternal factors rather than self-sperm. Imaging and oocyte quality assessment indicated that *repx-1* mutation preserved oocyte quality into mid-adulthood (day 5) relative to wild-type worms. Thus, preservation of oocyte quality accounts for *repx-1*'s extended reproductive span. Imaging using a *Prepx-1*::GFP reporter strain indicated that *repx-1*'s endogenous promoter acts in the intestines. Our results indicate that, in wild-type animals, *repx-1* expression in the intestine contributes to age-related reproductive decline. These results suggest that genes active in somatic tissues can affect oocyte quality and contribute to reproductive aging. Future work will aim to uncover the mechanism by which *repx-1* extends reproductive span, which could eventually inform treatments for mammalian infertility.

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

Developing an Adaptive Cognitive Prosthetic: Closed-loop multisite electrical stimulation produces specific neural firing patterns

Daniel Che, Sina Tafazoli, Camden MacDowell, Kate Letai, **Timothy Buschman**

Electrical micro-stimulation is commonly used to study neural circuits and manipulate their dynamics through neural prosthetics. We have shown that multi-site micro-stimulation in mouse neocortex can produce a multi-dimensional neural response that closely resembles natural sensory stimulation. However, it isn't clear if responses produced by electrical micro-stimulation would allow for control of regional or whole-brain dynamics. Controlling neural dynamics is challenging; precise stimulation patterns elicit a desired neurological response are hard to identify (due to high-dimensionality of stimulation patterns) and change over time (plasticity). We developed and tested a closed loop stimulation system with an integrated adaptive algorithm that could learn to generate specific spatiotemporal patterns of activity in specific neural populations. Using a 64-channel (32 recording/32 stimulation) silicon microelectrode in the primary visual cortex of awake head-fixed mice, we first used random stimulation patterns to map the neural response space. Next, we set individual response patterns as the desired output for the learning algorithm, which minimizes the difference between measured evoked neural response and the desired output, achieving sophisticated control over neural firing patterns. This novel system reliably learned to produce target neural activity patterns. In subsequent work, we will try to show that this closed-loop stimulation system can bias perception in a go/no-go visual discrimination task. This system serves as an important building block for the next generation of neural prosthetics.

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

Optimizing uDISCO-based clearing for signal preservation and increased tissue transparency

Shuqi Chen, Thomas J. Pisano, **Samuel S.H. Wang**

Tissue clearing is a set of techniques with the goal of making tissues optically transparent for volumetric imaging. This approach removes the need for tissue sectioning consequent damage and misalignment of sections. In these techniques, the basic goal is to extract lipids and immerse the tissue into a refractive index matching medium, rendering the tissue optically transparent for volumetric imaging. Among the clearing techniques, we selected uDISCO (ultimate 3DISCO) for optimization, as it already enables intact organs and whole-rodent body clearing, but only preserves strong endogenous fluorescent signals for several days. The uDISCO technique uses tert-butanol as a stable dehydrating agent and dichloromethane (DCM) as the delipidation solvent. The sample is finally kept in di-phenyl ether (DPE)/BABB with vitamin E for anti-oxidation. The goal of this project is to optimize the uDISCO protocol by changing the delipidation steps, thereby enhancing the resolution of volumetric imaging, preserving signals for a greater amount of time and rendering a better exploration into the microscopic organizations of tissues. As a benchmark for the improvement of our technique, we are comparing the dimensional shrinkage and transparency of mouse brains, using calipers and light-sheet microscopy, between original or optimized uDISCO clearing protocols.

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

Determining the activating mechanism of GCN2 kinase

Zhenhong Chen, Kristina Solorio, **Alexei Korennykh**

Protein translation is an energy demanding process, but this energetic cost is mitigated when eukaryotic cells undergo Integrated Stress Response (ISR). In response to diverse cellular stressors, eukaryotes activate ISR, an adaptive pathway to restore cellular homeostasis. The main event of this pathway is the phosphorylation of eIF2-α which is mediated by one of the four members of eIF2-α protein kinase family. GCN2 is a metabolic-stress sensing protein kinase that plays the role of ISR activator, which is necessary for the adaptation of amino acid starvation. In response to amino acid starvation, GCN2 phosphorylates eIF2-α which inhibits eIF-2B, a competitive inhibitor of general protein translation. Thus, leading to overall protein synthesis repression and reduced utilization of amino acids. GCN2 also upregulates transcriptional activators such as GCN4 which allows for alleviation of nutrient depletion through reprogramming of amino acid biosynthetic pathways. GCN2 is the only eukaryotic kinase that senses the quality of tRNAs and controls protein synthesis. However, the precise mechanism of signal transduction from ligand-binding domains to kinase remains unclear. My aim is to utilize X-ray crystallography to elucidate and yield a detailed understanding of this intricate mechanism which, in turn, will contribute to providing a complete and accurate mechanistic description of this essential cellular process.

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

Screening of Immunoregulatory Genes in Breast Cancer Therapy

Shuyuan Cheng, Minhong Shen, **Yibin Kang**

Breast cancer is the most commonly diagnosed cancer in women. It has poor prognosis with marked resistance to chemotherapies and other standard treatments. Recently, immunotherapy has achieved great success in many aggressive cancer types; however, limited response was observed in breast cancer. Although clinical trials indicated that partial breast cancer patients could benefit from immunotherapy, suggesting immunotherapy is promising for breast cancer treatment, breast cancer cells may suppress immune response, and therefore, escape from immunosurveillance. To enhance the response rate and efficacy of immunotherapy, and explore the mechanism underlying breast cancer cell-mediated immune suppression, we aimed to systematically identify immunoregulatory genes in breast cancer cells in this study. To this end, we first established *in vitro* tumor/immune cell co-culture system. Ovalbumin (OVA) antigen stably expressed breast cancer cell line Py8119 was co-cultured with immune cells, which are from OT-1 mice and can specifically recognize OVA antigen. The result indicated that Py8119 cells were effectively killed by immune cells upon the co-culture. Next, we put CRISPR/Cas9 machinery and genome-wide guide RNA libraries into Py8119 cells to either knockout or overexpress genes. Together with the co-culture approach, tumor cells with individual gene knockout or overexpression that resistant to the immune killing were identified. The tumor cells were then collected, and genomic DNA was extracted for PCR to amplify the guide sequences. The PCR product was submitted for next generation sequencing to identify the candidate genes that mediate immune suppression. The sequencing is still on going. By analyzing the sequencing result, we expect to uncover the mechanism through which tumor cells mediate immune suppression. Moreover, novel targets could be identified to synergize current immunotherapy strategies to achieve curative treatments for breast cancer.

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

Investigating PIWI regulation in the central nervous system of Ciona intestinalis

Annie Choi, Kai Chen, **Michael Levine**

The *Ciona intestinalis* (Ciona) is an ascidian species that has been a popular subject of research due to its proximity to vertebrates in evolution. The motile, swimming larvae of Ciona turn into sessile adults via metamorphosis. Metamorphosis of Ciona thus requires the regeneration and restructure of cells after reaching maturity. Recent research suggested that ependymal cells in the Ciona nerve cord may serve as neuronal stem cells by showing that part of the adult neurons differentiate from these cells. Our lab's RNA single cell transcriptomics data has shown P-element–induced wimpy testis (PIWI) expression in the central nervous system, including some overlapping expression with the ependymal marker CRALBP gene. Based on previous reports of PIWI function in maintaining genome integrity in the germline and neural progenitor cells, we investigated the regulation of PIWI expression in the Ciona central nervous system. In order to find the minimum enhancer of the piwi locus, we conducted a series of reporter assays using deletion constructs of the approximate promoter region fused with the photo-convertible fluorescence protein Kaede. Interestingly, deletion constructs including a 1kb region around 3kb upstream of the promoter showed enhanced expression of PIWI. Quantification of the fluorescence of each embryo using corrected total cell fluorescence (CTCF) showed an approximate two-fold increase in fluorescence for the constructs including this region. We hypothesize that this region is a repressor region. Further analysis of PIWI regulation, including overexpression of potential piwi transcription factors including MyT1, will give insight on PIWI's role in the central nervous system in light of the stem cell-like ependymal cells.

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

Dietary Supplementation of Oleic and Linoleic Acid Rescues Mating-Induced Death in *Caenorhabditis elegans*

Leo Choi, Cheng Shi, **Coleen T. Murphy**

Mating-induced death is a phenomenon in which an organism's lifespan is significantly reduced after mating; this effect is conserved between sexes and across multiple species. Mating-induced death is preceded by evident fat loss; understanding the role of fat metabolism in reproduction and lifespan will help identify methods to mitigate the post-mating lifespan reduction in many organisms, including humans. We investigated the role of fat metabolism by dietarily supplementing *Caenorhabditis elegans* mothers with fatty acids after 24 hours of mating. We hypothesized that such supplementation will restore the fat loss and mitigate the post-mating lifespan reduction. Fat staining confirmed that supplementation of oleic acid significantly restores the overall fat storage in mated worms. Remarkably, we found that controlled dietary supplementation of oleic and linoleic acid completely rescues the post-mating lifespan reduction. This lifespan rescue seems to be independent of the reproductive status and the total brood size of mated worms. Our results suggest that the level of fat storage plays a significant role in regulating the lifespan of mated *C. elegans* after completion of progeny production. Future studies will examine the specificity of oleic and linoleic acid, the conservation of this rescue in different sexes, and transcriptional changes mediated by supplementation. Such investigations will elucidate the scope and the mechanism of the lifespan rescue and further verify dietary supplementation of oleic and linoleic acid as a potent method to combat mating-induced death.

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

Computational paleovirology reveals previously undescribed taxa present in the viral communities of bees

Farion Cooper II, **Benjamin Rubin**, **Sarah Kocher**

Understanding bee virus diversity is key to preserving their ecological function as pollinators. The honey bee, *Apis mellifera*, provides essential pollination services to wild and agricultural plants and, thus, has been the focus of extensive epidemiological efforts. However, other bee taxa provide similar levels of pollination, yet little is known about the pathogens that infect these non-model species. Paleovirology, the discovery of fragments of viral genomes, otherwise known as endogenous viral elements (EVEs), in their hosts' genomes has offered new insights into the complexities of viral evolution and host-virus relationships. To begin to understand the community of viruses that infect the vast diversity of non-model bee taxa, we use computational approaches to characterize the EVEs found in the genomes of 11 bee species, confirming their presence using complementary molecular biological techniques. We find evidence of the presence of many different families of viruses previously undescribed in association with *Apis mellifera*, or any other bee. More specifically, we find signatures of the virus family Totiviridae across bees, suggesting that this taxon integrated into the genome of the ancestral bee lineage before it diverged into the many different species of today. Additionally, evidence for recent integration of Partitiviridae viruses into the bee *Dufourea novaeangliae* demonstrates the extant pathogen diversity that is missed by focusing almost exclusively on honey bee-associated pathogens. No members of either the Totiviridae or Partitiviridae or even the group to which they belong, the double-stranded RNA viruses, are known to infect honey bees. The discovered EVEs indicate that many different families of viruses, including several with no relatives found to associate with honey bees, infect these non-model bee species. Studies of the viral communities associated with non-model taxa are essential to completely capture the diversity of pathogens associated with these critical agricultural pollinators.

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

Regulation of an alternative quorum-sensing pathway in Vibrio cholerae

Daniela Coronado, Justin Silpe B.S., **Bonnie Bassler Ph.D**.

Quorum sensing (QS) is a cell-cell communication process bacteria use to orchestrate collective behaviors including virulence and biofilm formation. Biofilms are groups of bacteria adhered to surfaces. QS relies on the production, detection, and group-wide response to extracellular signal molecules called autoinducers (AI). *Vibrio cholerae*, the causative agent of the disease cholera, transitions between motile and biofilm lifestyles during disease transmission. Thus, understanding QS control of biofilm formation could be crucial for development of new antibiotics. The small RNA VqmR is a QS-controlled repressor of *V. cholerae* biofilm formation. *vqmR* expression, in turn, is activated by the QS AI called DPO and its partner receptor, VqmA. However, there exist additional VqmA-dependent, DPO-independent mechanisms of *vqmR* regulation. Here, we found through a genetic screen that a mutation in the gene coding for the phosphotransacetylase (Pta) enzyme, which controls acetate metabolism, activates *vqmR* expression above wild-type levels. While DPO was dispensable for Pta regulation of *vqmR*, VqmA was required implying that Pta may alter VqmA activity. Furthermore, *vqmR* expression was restored to wild-type levels by supplying medium conditioned by growth of the wild-type strain, suggesting that a signal produced by the wild-type, but not the *pta* mutant, is required for *vqmR* expression. Future experiments will clarify the identity of the mediating signal, its connection to VqmA, and its subsequent regulation of *vqmR* expression and biofilm formation. An improved understanding of these mechanisms will help in the creation of new alternatives to the antibiotics currently used.

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

The Arrhenius Equation and Complex Reaction Networks: Animal Development

Joseph Crapse, **Martin Wühr, Eric Wieschaus**

It has long been known that the Arrhenius equation describes the relationship between temperature and simple chemical reaction rates. Recent evidence from Eric Wiechaus' work with *Drosophila melanogaster* and Mustafa Khokha's lab's work with *Xenopus laevis* suggests Arrhenius' equation may also describe complicated reaction networks, such as embryonic animal development. This is counter-intuitive based on the simplicity behind Arrhenius' law, which starkly contrast with the complexity involved in embryonic animal development. To reconcile this question we investigated the temperature dependence of animal development in *X. laevis* model embryos. To this end we set up a microscope under tight temperature control and have started to carefully record frog development from the fertilized egg to the onset of embryonic movement at various temperatures. To determine if development matches Arrhenius we calculate mock reaction rates by inverting developmental times between phenotypically different developmental events. Once calculated we can plot these mock reaction rates as an Arrhenius plot to determine which developmental events follow the Arrhenius equation, and over what temperature range. Should development follow Arrhenius a new question arises. What is the meaning of activation energy? To address this puzzle, in future experiments we will run simulations of linearly coupled reaction networks, made up of individual Arrhenius reactions with random activation energies. Different distribution types will determine activation energies. Comparing the overall reaction network rate against the activation energies distribution we hope to gain insight into what determines the network's reaction rate. If rate is limited by the slowest or most plentiful activation energies then the network's activation energy is simply the limiting reactions' activation energy. More interestingly, if the network's activation energy is not shared by any of those reactions in the simulation then further research must be done to characterize a network's activation energy.

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

Identifying motile ciliated cell types required for maintaining spine straightness in zebrafish

Kerri A. Davidson, Daniel T. Grimes, **Rebecca D. Burdine**

Idiopathic Scoliosis (IS), characterized by a 3D lateral curvature of the spine, impacts around 4% of children worldwide. We know little about the causes of the disease, let alone a possible cure. My research will help us gain a deeper understanding of the cell biological mechanisms behind IS. In zebrafish, perturbed cerebrospinal fluid (CSF) flow is associated with IS-like spinal curves in juveniles. Flow is lost in mutants defective in motile cilia - organelles that protrude into the extracellular space and move fluid by coordinated beating – and spinal curves develop as a result. However, there are several cell types harboring motile cilia that could be relevant to the IS-like phenotype observed. To identify cell types that contain the transcriptional signatures of motile ciliogenesis at the onset of curvature, I am using single cell RNA-sequencing (sc-RNA seq). To date, I have developed a protocol that obtains a single cell suspension from juvenile fish just prior to when cilia motility-deficient mutants develop spinal curves. Fluorescence activated cell sorting (FACS) is used to isolate all cells expressing *foxj1a*, a master regulator of motile ciliogenesis, providing a mixed population of motile ciliated cell types. The next step is to perform sc-RNA-seq to determine the identity of the motile ciliated cell types. These data will be used to perform cell type-specific rescue experiments to determine the requirement of these cell types in spine development. Lastly, I will describe my CRISPR/Cas9 mutagenesis experiments to investigate the role of *poc5*, a causative IS gene, in cilia formation and spine development in zebrafish.

This research was made possible by the generous support of the Susan W. and James C. Blair ’61 P87 Endowed Senior Thesis Fund. Summer Undergraduate Research Program in Molecular and Quantitative & Computational Biology.

Elucidating the Role of METTL14 in Breast Cancer Tumor Initiating Cells

Sofia Dimitriadoy, Nicole M. Aiello, **Yibin Kang**

The N^6-methyladenosine (m^6A) modification is the most prevalent internal modification in eukaryotic messenger RNAs (mRNAs) and has been shown to promote mRNA decay, suggesting its role as a negative regulator of gene expression. Several studies have shown METTL14, a component of the m^6A methyltransferase complex, may promote or inhibit stemness in tumor initiating cells (TICs) depending on the cancer subtype. In breast cancer, patient clinical analysis shows that low METLL14 expression correlates with poor overall survival and recurrence free survival in breast cancer. Here, we report that METTL14 is downregulated in normal mammary stem cells (MSCs), suggesting METTL14 promotes differentiation in TICs and may serve as a tumor suppressor in breast cancer. To investigate the role of METTL14 in breast cancer tumorigenesis, we overexpressed METTL14 in breast cancer cell lines and tested proliferation and self-renewal in vitro through tumorsphere assays. Because METTL14 has never before been studied in the context of breast cancer, recognizing its role serves as an important step in further understanding breast cancer TICs and the clinical data.

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

Effect of a Novel Fibronectin Mutation on Collagen Processing and Matrix Assembly

Ashley Drengler, Eli Cadoff, **Jean Schwarzbauer**

Proper assembly of extracellular matrix (ECM) is essential for normal cell and tissue development; improper ECM regulation can lead to serious bone and cartilage defects including skeletal dysplasia. Fibronectin (FN) and collagen are crucial ECM proteins that assemble into matrices and regulate various cellular activities such as cell migration and differentiation. The discovery of a novel FN mutation (FN^{C97W}) linked to an individual with spondylometaphyseal dysplasia (SMD), a specific skeletal developmental defect, offers the opportunity to investigate the mechanisms through which FN^{C97W} affects cell function and leads to skeletal dysplasia. Previous studies show that FN^{C97W} cells accumulate mutated FN within the endoplasmic reticulum (ER) and assemble reduced FN and collagen matrices. In this study, I used primary dermal fibroblasts from an individual heterozygous for the FN^{C97W} mutation to show the effects of a mutation in FN on collagen assembly and matrix formation. Through immunoblot and immunofluorescence techniques, I have found that while there is no difference in intracellular collagen levels between mutant cells and age- and gender- matched wild type cells, mutant cells present collagen in large inclusions within the ER. Additionally, collagen levels are elevated in the culture media of the mutant cells, while immunofluorescence shows no collagen matrix formation in the FN^{C97W} cells. Future studies aim to investigate the possible steps of collagen processing, secretion, or matrix formation collagen that are affected. This work will suggest the mechanisms by which mutated FN can impact collagen assembly and matrix formation, providing important insights into the role of these essential ECM components pathogenesis of skeletal dysplasia.

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 Progesterone on Telomerase Activity and Telomere Length in Mice with Calibrated Traumatic Brain Injury

Kerry Farlie, Iulia Kotenko, Lisa Schneper, **Daniel Notterman**

Traumatic Brain Injury (TBI), a "silent epidemic" that kills approximately 50,000 people annually, has no known effective monotherapy. TBI can result in profound neurological damage, including necrosis, oxidative damage, edema, inflammation, and numerous other sequalae. Recent literature highlights progesterone, a female ovarian hormone, as a candidate treatment due to its neurogenerative and neuroprotective properties. Despite progesterone's remarkable successes in animal models and phase II clinical trials, multi-site phase III trials failed to support its use to attenuate TBI. In cell culture studies, higher doses of progesterone correlate with suppressed telomerase activity while lower doses increase telomerase activity. Furthermore, in prior work our lab has shown that TBI is associated with changes in telomerase activation in the thalamus, cortex, and hippocampus of mouse brains, suggesting a role for telomere homeostasis in the damage response. We hypothesize that the dose of progesterone administered in the phase III trials may have been too high; thus, telomerase activity was suppressed and progesterone did not confer maximum neuroprotection. This study will determine the optimal concentration of progesterone to induce maximal telomerase activity in cell culture. Using pharmacokinetic calculations, this concentration will be extrapolated to a dosage for intraperitoneal administration in mice following calibrated blunt-force head trauma. To explore whether the negative phase III trial could be attributed to an excessive dose of progesterone, I will compare telomere length and telomerase activity at varying progesterone doses in both mice with and without (i.e. control) TBI. Finally, I will assess histopathology under the microscope and changes in soluble TBI biomarker levels in brain tissue from the different groups of mice, correlating these findings with progesterone dosage and telomerase activity.

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

Phenotyping Non-Coding Autism Related SNVs in Embryonic Stem Cells

Courtni Foster, Chandra Theesfeld, **Olga Troyanskaya**

Autism is a highly complex genetic trait caused by both *de novo* and heritable mutations that may manifest consequences in early-mid fetal development, when early neural stem cells are populating the rapidly forming brain. From whole genome sequencing more than 70% of autistic individuals in the Simons Simplex Collection do not have mutations in genes, but rather non-coding sequences. As there is no “code” for the noncoding genome, we used a deep learning computational method, DeepSEA, to distinguish high impact non-coding mutations from benign mutations. We experimentally demonstrated the transcriptional effects of dozens of predicted high impact mutations compared to corresponding benign alleles using luciferase assays. Two of the highest impact noncoding mutations are predicted to affect the expression of genes sending and receiving input to the Wnt developmental signaling pathway: HES1 and NEUROG1. To study the impact of these mutations in stem cells we will use the newest generation of CRISPR/Cas9 single base editors (Koblan et al 2018). With these new Cas9s, the target base must lie in the edit window between nucleotides 5 and 10 in the gRNA (labelling distally from the PAM sequence). We tested stem cell transfection methods and found Lipofectamine was more successful than nucleofection; and less DNA gave a higher percentage of transfected cells with less cell death. We are screening the edited cells and planning phenotype assays to determine impact of noncoding mutations on cell growth, Wnt signaling, and the ability of the stem cells to differentiate into neural precursor cells and neurons.

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

Structure based N-methyl chloroamide cyclase design using OYE-family enzymes

Samuel E. Garfinkle, Megan Emmanuel, Simon Cooper, Xin Gao, Kyle Biegasiewicz, **Todd Hyster**

The demand for effective, affordable, and stereospecific catalysts, especially in fields such as drug discovery and medicinal chemistry, has given new life to the field of protein engineering. Historically a challenging pursuit, the engineering of enzymes in particular is currently accomplished through a combination of rational design and combinatorial directed evolution. Both of these approaches share an important root: the engineering of function based on structural changes to the peptide sequence and fold. This project harnesses the power of structure-based protein design to engineer a set of N-methyl chloroamide cyclase enzymes using OYE-family ene-reductases as a starting peptide template. X-ray crystallography is used to determine the first full structure of wild-type GluER as well as catalytically favorable mutants, while computational sequence alignment and ligand docking are used to predict and rationalize catalytically relevant mutated amino acid residues. Directed evolution of likely active site residues in OYE1 using degenerate-codon-generated plasmids is screened for changes in yield by LC-MS and for selectivity by chiral chromatography, allowing us to optimize OYE1 to perform one variation of an enantioseletive cyclization with high yield and stereoselectivity. Apart from optimizing novel light-driven reactivity, this approach showcases the effectiveness of hybrid rational/combinatorial approaches for structure-based enzyme design.

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

Investigating the Role of MlaA-OmpF/C Interaction in the Mla Pathway

Valeria A. Gerena-González, Randi L. Guest, **Thomas J. Silhavy**

Bacteria often develop resistance to toxins, such as antibiotics and detergents, which may translate to severe and life-threatening infections. Gram-negative bacteria are able to resist antibiotics due to the sophisticated organization and composition of their outer membrane (OM), which has an asymmetric lipid distribution with lipopolysaccharides (LPS) in the outer leaflet and phospholipids in the inner leaflet. Disruption of LPS can lead to mislocalization of phospholipids in the outer leaflet of the OM. This is detrimental for the bacterium, as it reduces OM barrier function, and increases the cells sensitivity to toxic molecules. To restore OM lipid asymmetry, bacteria have a pathway that mediates the retrograde transport of misplaced phospholipids in the outer leaflet of the OM, called the maintenance of lipid asymmetry pathway (Mla). This pathway consists of several proteins that are located in the OM, the periplasm, and the inner membrane. The MlaA is an OM transport lipoprotein that removes phospholipids from the outer leaflet to maintain the asymmetry of its membrane. Previous studies have shown thatMlaA forms a stable complex with the outer membrane proteins F and C (OmpF/C) by Van der Waals interactions, yet the OmpF/C do not play and active role in the transport of phospholipids. The aim of this study is to understand the reason that MlaA interacts with the OmpF/C. We will measure MlaA-OmpF/C interaction by UV photocrosslinking. We will then disrupt this interaction and determine the effect this has on the activity of the Mla pathway. Understanding the molecular details of Mla can lead to more efficient targeting of drug resistance bacteria and as well develop more efficient antibiotics.

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

Microbiome-Derived Metabolism of Anti-Infectives: Identifying the Human Gut Microbiome Isolates Responsible and Measuring Changes in Bioactivity *Ex Vivo*

Jess Goehring, Bahar Javdan, **Mohamed S. Abou Donia**

Variation in the efficacy of anti-infective medication among individuals is often attributed to either differences in host genome expression or the variability of anti-infective resistance among diseases; however, microbiome-mediated metabolism is a component of drug metabolism that has been overlooked despite the role it plays in drug modification. During enterohepatic circulation and prior to first-pass metabolism, xenobiotics are exposed to a subset of up to 100 trillion microbes within both the small and large intestines, providing ample opportunity for bacterial enzymes to metabolize them prior to host metabolism. In this study, we analyzed anti-infective drugs for microbiome-mediated metabolism by exposing them to both the entire microbiome of a healthy donor (HD1) as well as isolates representative of four of the six main phyla present in the gut microbiome. The combined separation power (High Performance Liquid Chromatography) and identification power (Mass Spectrometry) of Liquid Chromatography Mass Spectrometry was used to determine when the incubation of a specific isolate-drug pair resulted in a change in structure of the parent drug molecule. Where microbiome-derived metabolism is observed, there is also potential for a change in bioactivity-- the ability of a molecule to trigger a response or cause a reaction in living tissue-- of the modified molecules in comparison to their parental counterparts, which can have dramatic effects on drug response *in vivo*. Changes in bioactivity can also be easily measured *ex vivo* through the use of anti-infective disc diffusion bioassays. Because an anti-infective-soaked disc can induce a zone of inhibition within the field of growth of a susceptible microbial strain, changes in the diameter of inhibition reveal changes in bioactivity. By understanding how the gut microbiome can potentially modify therapeutics, clinicians can move towards more highly personalized medicine, where the composition of an individual's microbiome can determine appropriate treatment regimens for that patient.

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

The impact of spatial transcriptional differences in the liver on hepatitis delta virus infection

Raymond Guo, Benjamin Winer, **Alexander Ploss**

Hepatitis delta virus (HDV) is a satellite virus requiring hepatitis B virus (HBV) co-infection; HBV/HDV co-infection of human liver hepatocytes causes accelerated liver damage and elevated risk of hepatocellular carcinoma. Environmental cues affect hepatocytes' transcriptional profile, and transcriptional differences divide the liver into zones. However, it is unclear if all hepatocytes are equally susceptible to HDV infection; we hypothesize that liver zonation patterns distinctively affect permissiveness to viral infection. To test this hypothesis, we have generated a novel reporter system fusing HDV to Cre recombinase which can label hepatocytes that take up this "HDVLP-Cre." To characterize and optimize this construct, we infected a liver cell line expressing the hNTCP receptor permitting HBV/HDV viral entry. Following successful *in vitro* infections, we will infect a novel genetically humanized mouse model supporting HDV infection with HDVLP-Cre. *In vitro* HDV-Cre DNA transduction experiments and hydrodynamic delivery to humanized mice confirm recombinase activity. To reduce cytotoxicity effects, the timeline for HDVLP-Cre production and infection has been reduced, and additional cloning experiments are in progress to generate a smaller viral like particle containing only the nineteen amino acids of HDV necessary for viral packaging fused to Cre recombinase, which will improve cellular uptake of the construct. Moving forward, fluorescent microscopy will quantify the frequency of HDVLP-Cre permissive cells in humanized mice and record their spatial location, which will be compared to the frequency and spatial location of hepatocytes bearing HDV RNA following wild-type HDV infection. This comparison will demonstrate whether or not all HDV entry-permissive hepatocytes also support subsequent viral replication, potentially revealing differences in the susceptibility of hepatic subregions to HDV and suggesting future targeted therapeutic treatment of HDV.

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

Mutating the Primary Sequence of the CrvA Protein in Vibrio cholerae

Kristian Harris, Nicholas Martin, **Zemer Gitai**

Vibrio cholerae is a gram negative, human bacterial pathogen that is responsible for one of the most serious diarrheal diseases. The characteristic curved-rod shape of V. cholerae improves its swimming capabilities through viscous media and increases pathogenesis in disease models. The curvature determinant CrvA produces this curved shape by forming a polymer that localizes to one side of the cell and reducing the amount of cell wall growth at that surface. CrvA does not enzymatically alter the cell wall, suggesting an interaction with cell wall synthesis machinery is required. However, these possible interaction partners and the parts of the CrvA protein which facilitate them are currently unknown. The purpose of this research is to identify amino acids in the CrvA protein required for its localization and activity. We used error-prone PCR to introduce point mutations into a fluorescently labelled CrvA protein to observe its phenotypic changes. We discovered three classes of mutants that alter protein localization or curvature: a straight cell with diffuse protein signal, a straight cell with patchy polymers, and cells with the wild-type curve shape but with a diffuse protein signal. We also discovered that within the patchy polymer class, there are mutants that have localization patterns with distinct temporal dynamics.

This research was made possible by the generous support of the Genentech Foundation and the Lewis-Sigler Institute for Genomics

A Functional Comparison HSV-1 and PRV VP16 on Neuronal Gene Expression

Zhi-Shan Hsu, Orkide O. Koyuncu, **Lynn W. Enquist**

Herpes simplex virus (HSV-1) is a virus that establishes a reactivatable latent infection in the nervous system, but as a human pathogen, the animal models available to study its infection cycle are limited. On the other hand, pseudorabies virus (PRV), a related alpha herpesvirus, has a broad host range and can infect most mammals. The HSV-1 VP16 protein, a viral tegument protein, is delivered to cells immediately after infection. VP16 has viral gene transactivating capabilities that turn on viral mRNA synthesis required for escape from latency and may have functionally conserved domains in PRV VP16. My studies seek to determine if HSV-1 VP16 and PRV VP16 have similar gene activating properties in neurons. We constructed adeno-associated virus (AAV) vectors expressing the fluorescent reporter protein mTurqoise2 alone or as a fusion protein with HSV-1 or PRV VP16 and tested them in the rat2 epithelial cell line and primary sympathetic neurons. For protein analyses, we made a polyclonal antibody specific for PRV VP16 and used a commercially available HSV-1 VP16 antibody. These reagents enabled us to compare protein expression and subcellular localization of PRV and HSV-1 VP16 during infection and AAV transduction. Following these characterizations, superior cervical ganglia (SCG) neurons from embryonic rats were grown and transduced with these AAV vectors. Their RNA was extracted and sequenced (RNA-seq) to identify and quantify any changes to gene expression caused by HSV-1 or PRV VP16. This comparative study will reveal the similarities and differences of these alpha herpesvirus tegument proteins to activate neuronal genes.

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

Can sleep rescue memories from retrieval-induced forgetting?

Yechen Hu, Elizabeth A. McDevitt, **Kenneth A. Norman**

Retrieving a memory can improve access to that memory, while also impairing memory for related competitors, a phenomenon known as retrieval-induced forgetting (RIF). However, restudying the damaged competitor may help differentiate the neural representations of these memories, resolving the competition and reducing RIF. We hypothesize that sleep may be an opportune time to revisit the damaged competitors, an idea lent support by prior work showing that sleep can eliminate RIF (Baran et al., 2010). We aim to expand upon this work, using a different task, by 1) replicating the traditional RIF effect with no time delay and 2) examining how RIF is modulated by time spent awake or asleep.

Participants study animal-name pairs from six animal categories (e.g., Abner the Ape and Olaf the Otter). Then, participants are shown an animal and practice retrieving its name. Critically, for 3 of 6 categories, half of the names are practiced (Rp+) and half are not (Rp-); all names in the remaining 3 categories are unpracticed (Nrp). In the final memory test, participants recall the names of all animals. In Study 1, we expect to replicate typical RIF results: Memory for Rp+ names is the best, whereas memory for Rp- names is worse than Nrp names, even though both are studied the same amount. In Study 2, a period of sleep or wake will be inserted between the retrieval practice and final test phase. We predict that sleep will boost memory for Rp- names, thereby reducing the overall RIF effect.

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

Investigating the Relationship between Endogenous Erk Dynamics and Downstream Cell-fates in Primary Mouse Keratinocytes

Danielle Isakov, Siddhartha Jena, **Jared Toettcher**

The Ras/Erk mitogen-activated protein kinase (MAPK) pathway is a central signaling pathway that is involved in diverse cell fates such as proliferation, differentiation, and apoptosis. Although the molecular players of the pathway have been extensively characterized, the method by which the cell uses a single pathway to interpret varying inputs into different downstream outputs is still unknown. Uncovering this missing process of regulation is crucial for understanding not only natural phenomena, but also pathogenic cellular behavior, as mutations in the Ras/Erk pathway account for nearly a third of all cancers and 80% of melanomas. Our lab is interested in one dimension of control—dynamics of Erk activity, involving duration, intensity, and frequency of pulses. Changes in Erk oscillations in cells have shown to result in different downstream cellular outcomes. Our model system of choice, primary mouse keratinocyte epithelial cells, display endogenous Erk dynamics and also exhibit strong proliferation and differentiation behaviors, making them an ideal candidate for studying the significance of Erk dynamics in a physiologically relevant paradigm. In this project, we explore the central question of whether endogenous Erk dynamics in keratinocytes lead to transcriptional outputs of downstream targets of Erk. Using PCR cloning, we created constructs in which endogenous promoters of AP1, SRE, and Involucrin drive a live transcriptional reporter (MS2/MCP) and a translational reporter (YFP). In order to create stable keratinocyte cell lines with these constructs, we optimized the TARGATT site-specific knock-in system for transfection in primary cells. We will also investigate how Erk dynamics affect differentiation in distinct layers of the air-liquid interface tissue model. Through these inquiries, we will characterize the role of endogenous Erk dynamics in a primary cell model, and further elucidate on how the Ras/Erk signaling pathway translates inputs into diverse cell fates through dynamic control.

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

Frequency and Patterns of Cancer Causing Somatic Mutations in 14 Healthy Individuals

Sophia Jackman, Selina Vattathil, **Joshua Akey**

Tumor formation in healthy tissues is partially a result of accumulated somatic mutations, especially cancer driver mutations, which promote tumor development. Patterns of somatic mutations in tumors have been studied in depth and documented. However, we do not know the patterns of somatic mutations in non-tumor tissues. We expect a small subset of the mutations observed in healthy tissues to be known cancer causing mutations. The goal of this project was to find baseline levels and patterns of somatic mutations, especially cancer drivers. We analyzed 267 tissue samples from 14 healthy individuals using exome sequencing, a process that sequences the protein coding regions of the genome. Somatic single-nucleotide variants were identified using the software MosaicHunter and the data was analyzed using Python and R scripts. Mutations in these healthy individuals were then compared to pathogenic mutations found in tumors, compiled in the Catalogue of Somatic Mutations in Cancer database, and cancer driving genes found in the Cancer Gene Census. In total, 2524 mutation calls in 281 unique positions were found. Four pathogenic somatic mutations were found, one of which was found in a confirmed cancer driving gene. It is possible that some of these calls are false positives due to sequencing artifacts. Although we have determined that low read depth did not account for these findings, further analysis should be done to identify criteria to be used to remove mutation calls due to non-random sequencing errors. Our findings suggest a non-negligible level of cancer driving somatic mutations exist in healthy tissues. Future directions include studying somatic cancer driving copy number changes and mutations associated with other diseases in healthy individuals.

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

Investigating a Link between Child DNA Methylation and Exposure to Fracking

Benjamin Jacobson, Lisa Schneper, Louis Donnelly, **Daniel Notterman**

The development of hydraulic fracturing (fracking) for natural gas recovery over the past decade has radically transformed the nature of U.S. energy production, playing a large role in achieving energy independence. However, the associated increase in fracking site development has proceeded with a limited understanding of its consequences on the health of nearby residents. While recent studies suggest significant health risks associated with fracking exposure, the fact that fracking has only emerged in the past few years has limited the number of extensive, longitudinal studies investigating these impacts on human health. Further, previous studies have focused primarily on phenotypic outcomes, such as birthweight and pneumonia. We hypothesize that alterations in DNA methylation will be a more sensitive indicator of environmental toxic exposures due to fracking. The Fragile Families and Child Well Being Study (FFCWS) dataset, while not collected with fracking in mind, aligns well with the advent of fracking, with epigenetic data collected from over 3000 children just before the fracking surge (children 9 years of age) and again after six years of sustained fracking development. Geographic, temporal, and genetic factors will be used to model the relative exposure of FFCWS subjects to fracking. Correlation of this exposure with DNA methylation age changes will seek to uncover the geographic range of fracking effects. Subsequent differential methylation analysis will investigate the specific ways in which fracking exposure impacts the epigenome, further elucidating the impacts of fracking exposure on human health.

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

Absolute Concentration of Metabolites in *Saccharomyces cerevisiae*

Artem Khan, Tianxia Xiao, **Joshua D Rabinowitz**

Absolute concentration of metabolites is an essential aspect for an understanding of cellular metabolism. It enables an estimation of such important functions as the change in Gibbs free energy (ΔG) and provides useful information about kinetics of reactions. Hence, knowledge of absolute concentration of metabolites can reinforce our understanding of biochemical pathways inside the cell. Although there is an already existing extensive list of absolute concentration of metabolites for such model organisms as Baker's strain of *Saccharomyces cerevisiae*, many important compounds are still missing. In this project, we tried to quantify and augment the list of metabolites in *Saccharomyces cerevisiae*. We applied liquid chromatography - mass spectroscopy (LC/MS) to find absolute concentration of metabolites in yeast by finding the ratio between the peaks of a ^{13}C-labeled compound inside the cell and a prepared unlabeled ^{12}C standard solution with a known concentration. By using this method, first, we verified previously made measurements of several compounds: alanine, ATP, citric acid, fructose-1,6-bisphosphate, glucose-6-phosphate, leucine, malic acid, serine, and succinic acid. Then, we tried to find concentrations of metabolites that have never been measured such as histidine, 3-hydroxybutyrate, mannitol, pantothenic acid, pyroglutamic acid, shikimic acid, and sorbitol. The next step of the project could be a further extension of the list of metabolites as well as an application of the absolute concentration data in the context of thermodynamics (for example, calculating Gibbs free energy) and flux analysis. Moreover, the obtained results can be useful in quantification of metabolites in other yeast strains or species. The data is important for a bioengineering research: for instance, some of the recently measured metabolites potentially may be significant in understanding of pathways relevant to the biofuel production.

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

Single-Particle Cryo-EM of the human NPC1/NPC2 complex

Joyce Lee, Xuelan Wu, Hongwu Qian, **Nieng Yan**

Niemann-Pick disease type C (NPC) is a rare genetic neurodegenerative disorder that results from mutations in the *NPC1* and *NPC2* genes, causing an accumulation of cholesterol within the lysosome and late endosomes of cells. The NPC1 and NPC2 proteins are responsible for low-density lipoprotein (LDL)-derived cholesterol trafficking and egress within the lysosomes. Therefore, these proteins play a central role in cellular cholesterol metabolism. Existing biochemical evidence strongly suggests the hypothesis that NPC2, a soluble protein, delivers cholesterol to NPC1, a polytopic lysosomal membrane protein, through a hydrophobic handoff mechanism, but there is no direct evidence of this occurring. Therefore, the goal of the investigation is to capture images of multiple conformational states of the NPC1/NPC2 complex through protein expression, protein purification, and high-resolution imaging with single-particle electron cryomicroscopy. Investigating the physical complex formed by these proteins, along with their interactions with cholesterol, may elucidate broader mechanisms behind cholesterol trafficking, metabolism, and related diseases. Furthermore, NPC1 in particular has a sterol sensing domain (SSD) that is highly conserved in many proteins involved in cholesterol metabolism and signaling. Understanding how this domain interacts with cholesterol may provide a strong framework for understanding how other biochemical pathways are possible. Using a 3-D classification algorithm and reconstitution from particle picking, analysis reveals a preliminary <10Å structure of the complex, which confirms NPC1 and NPC2's complex formation and shows promising early evidence supporting the hydrophobic handoff hypothesis. Further optimization of the cryo-sample will be necessary to improve the resolution to beyond 4Å and to potentially allow imaging of additional states.

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

Investigating the activation and asymmetric regulation of *dand5* in zebrafish

Sally Lee, Jose Pelliccia, **Rebecca Burdine**

Establishing the left-right (L/R) axis in the vertebrate body is a crucial step in proper development, as it governs the asymmetric positioning and formation of the organs and vasculature. This is established in the early embryonic stage through the breaking of symmetry at the left right organizer. In zebrafish, cilia generated fluid flow at the organizer results in the right-side preferential expression of *dand5*, a key repressor of the Nodal signaling cascade that leads to formation of left-side morphology. However, the factors that are responsible for *dand5* expression and asymmetric regulation remain elusive. Previous studies have demonstrated that the 1.2kb upstream region (referred to as the URR) and the 3'UTR are essential to drive *dand5* transcription and asymmetric expression, respectively. Building on these findings, this work first aims to identify the regulatory sequences necessary and sufficient for *dand5* transcription through transient URR deletion assays via the Tol2 transposon system in the zebrafish embryo. Injections of plasmids containing URR-driven reporter gene *egfp* have proven we have a working model for visualizing *dand5* activation. Second, this study will investigate the 3'UTR cis-regulatory elements responsible for the post-transcriptional asymmetric regulation of *dand5* expression, using CRISPR/Cas9 to generate mutants with deletions along the 3'UTR. Phenotypic assays will be conducted through immunofluorescence and *in situ* hybridization for *dand5*. The findings from this project will provide new directions for further research into asymmetric determination and organogenesis, and provide new insights into the molecular mechanism of major laterality defects, such as congenital heart diseases.

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

Investigating the Interaction Between the Sec1/Munc18 (SM) Protein Vps33 and its Qa-SNARE Vam3 and R-SNARE Nyv1

Sarah Lee, Travis Eisemann, **Frederick Hughson**

Soluble N-ethylmaleimide-sensitive factor attachment protein receptors (SNAREs) form four-helix membrane-bridging complexes that promote membrane fusion. Such SNARE complex formation is regulated by Sec1/Munc18 (SM) proteins. SM proteins interact with two of the four SNAREs that form a SNARE complex, Qa- and R-SNAREs, but a unified mechanism for SM protein function remains elusive. Separate crystal structures of the SM protein Vps33 bound to its R-SNARE Nyv1 and its Qa-SNARE Vam3 suggested that Vps33 and potentially other SM proteins could serve as templates for the half-zippered SNARE complex. However, no experimental structure of an SM protein bound to its Qa- and R-SNAREs simultaneously has been reported for any SM protein. Therefore, this study aims to obtain the complex of *Chaetomium thermophilum* Vps33 bound to both Vam3 and Nyv1, providing structural data of an intermediate of Vps33-mediated SNARE complex assembly. A structure of the ternary complex will reveal whether and if so, how Nyv1 and Vam3 influence each other's binding to Vps33 as they initiate SNARE complex assembly. To enhance the stability of the Nyv1-Vam3 bound Vps33, heterodimers of Nyv1 and Vam3 have been formed through cross-linking the two SNARE motifs with a disulfide bond. To obtain and characterize the structure and its binding affinities, isothermal titration calorimetry and crystallization trials will be performed using the purified Nyv1-Vam3 heterodimer and Vps33. The results from this study will provide important insight into the mechanisms of SM proteins in SNARE complex assembly and vesicle fusion.

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

Control of Translational Repression Through a Light-Inducible Clustering System

Lian Kirit Limperis, Ping Wu, **Jared Toettcher**

It has been extensively reported that RNA and protein aggregation is a naturally occurring phenomenon within the cell. It has been proposed to deposit RNA transcripts asymmetrically (e.g., germ cell granules) and assemble temporary storage compartments of RNAs and translational complexes for stalling or prevention of translation (e.g., stress granules, P-bodies). RNA granules contain various proteins such as translational repressors, helicases, decapping enzymes and scaffold proteins. It is still not clear whether these functional proteins are required for translational repression or if mere sequestering of RNA by clustering is sufficient. With light-inducible control of the clustering of MCP protein, whose sole activity is to bind with MS2 RNA adapters, we can dissect these two scenarios. Our novel synthetic system fuses an optogenetically controlled clustering protein (PixELL system) with an MCP protein. The MS2 RNA is attached to a fluorescent translational reporter. MS2-MCP affinity causes recruitment of the reporter transcript to the light-controlled protein cluster. The reporter's translation, or lack thereof, will subsequently be measured. Currently, mouse fibroblasts (NIH 3T3s) have been transduced with the PixELL system and successive incorporation of the RNA reporters is in process. Suppression of reporter translation after clustering will be quantified and normalized to a control reporter that does not recruit to the cluster. We expect a significant decrease in translation after RNA clustering. If not successful, the MS2 aptamer and MCP pair may be replaced with an endogenous RNA sequence and its known binding partner, such as bicoid 3' UTR and Me31B, and translation will be quantified. This system could be used as a tool to further elucidate the role of endogenous clustering in the stalling of translation. Furthermore, it may provide the field with a novel technology for spatiotemporal control of such translational suppression via optogenetics.

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

Analysis of host responses to hepatitis B and delta viral infections in a micro-scalable hepatic co-culture system

Gabriel Lipkowitz, Benjamin Y. Winer, Jenna M. Gaska, , Amit Parekh, Lance Parsons, Robert Leach, Cheul Cho, Anil Shrirao, Eric Novik, **Alexander Ploss**

Hepatitis B virus (HBV) is a persistent global health threat with 257 million chronically infected individuals worldwide, of whom approximately 20 million are co-infected with hepatitis delta virus (HDV). Progress towards understanding the complex interplay between these two viruses and developing novel therapies has been hampered by the scarcity of suitable cell culture models that mimic the natural environment of the liver. Here, we establish HBV and HBV/HDV co- and super-infections in self-assembling co-cultured primary human hepatocytes (SACC-PHHs) for up to 28 days – longer than any have reported before – in a 384 well format. In proof-of-concept experiments, we highlight the suitability of this platform for high-throughput drug testing. Importantly, we demonstrate that successful infection does not require dimethyl sulfoxide, a supplement traditionally used to facilitate infection but, as we show, is highly disruptive of numerous host cell processes. Transcriptomic analysis demonstrates that hepatocytes in this culture format maintain a mature hepatic phenotype over time, regardless of infection status. Our unbiased transcriptomic analysis of these samples also provides a broad picture of the changes occurring during infection over time, with a more diverse transcriptional profile observed at earlier time points (8 days following infection). We demonstrate that HBV is a “stealth” virus that does not induce a strong response by the innate immune system, which we verify can be robustly activated by other stimuli in SACC-PHHs. Neither does HBV/HDV co-infection activate innate immunity. Instead, infection perturbs pathways involving metabolism, steroid hormone synthesis, and extracellular matrix, which we speculate may create an environment conducive to viral persistence. Altogether, these data provide crucial insight into short and long term host responses to HBV mono- and HBV/HDV co-infections, and establish the SACC-PHH system as a versatile platform that can be utilized in the future to deepen our understanding of such complex infections.

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

Apply European ancestry-based GWAS data to height prediction for Africa Ancestry by Gradient Boosted Regression Trees

Chaozhong Liu**,** Lisa Schneper, **Daniel A. Notterman**

Genome-Wide-Association-Studies (GWAS) have been widely used for several decades in the fields of epidemiology and genetics to identify candidate genes and single nucleotide polymorphisms (SNPs) that are statistically associated with a certain disease or phenotype. In addition to the p-value obtained, the effect size of each SNP, or the contribution of the SNP to the phenotype, is also calculated from GWAS. These effect sizes, together with genotype information, are used to construct polygenic scores (PGS), which in turn can be used to predict phenotype. While there are many different advanced algorithms to generate PGS from effect sizes, most GWAS were done using data from European cohorts. Only recently has there been an effort to include diversity of dataset. Due to the differences in genomic structure among different populations, it is impossible to make a robust prediction for other populations using effect sizes calculated from European populations. Taking the huge amount of money and time GWAS costs into consideration, this study aims to develop a machine learning model to apply European population GWAS summary statistics to different populations.

Recently, Gradient boosted regression tree and LD adjusted (GraBLD) heuristic was proposed to improve the performance of prediction for height and BMI by using GIANT GWAS meta-analysis based on European populations. Given that the meta-analysis has been adjusted for variance in genomic structure, this study proposed that the prediction accuracy would be improved when applying the model, together with the GIANT summary statistics, to African-American populations. Height and genotype data was collected from Fragile Family and Child Wellbeing Study. This new application is believed to make a better prediction for populations on which GWAS were less done.

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

Interactions between enhancers in the regulation of even-skipped transcription in *Drosophila* embryos

Jelani Lyda, Philippe Batut, **Michael Levine**

The segmented body plan of Drosophila melanogaster is specified during early embryogenesis by a transcriptional regulatory cascade – a network of genes that encode transcription factors and regulate each other's expression. This regulatory cascade, initiated by maternally deposited factors, leads to the establishment of precise gene expression patterns along the anterior-posterior axis. We are interested in the even-skipped gene, which is expressed in seven stripes along the embryo shortly before the onset of gastrulation. Its transcription is under the control of enhancers in the different stripes. These enhancers have been thought to work independently from each other to create their respective stripes; however recent research seems to contradict this. We have engineered the endogenous even-skipped locus using the CRISPR/Cas9 system to introduce an array of MS2 loop sequences into the 3' UTR of the gene. These loops are incorporated into the mRNA during transcription of the target gene and induce the formation of RNA stem-loop structures, which bind specifically to MCP-GFP fluorescent molecules expressed in the embryo. This produces a bright fluorescent dot at the site of transcription, allowing transcriptional activity to be measured in real time using a fluorescent confocal microscope. This project aims to delete stripe 1 and stripe 2 enhancers separately to examine the effect of a missing enhancer on the other stripe's expression. The deletions were created using CRISPR Cas9 at the endogenous locus in which the MS2 loops were previously inserted. An analysis of the intensities and location of the gene expression of each stripe with the 2 deletions compared to the wild type will be carried out in the future. It is expected that there should be an expansion of the stripe's location or decrease in expression with the deletions.

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

The role of Purkinje cells in Cognitive and Affective Development

Christina Matl, Jessica L. Verpeut, **Samuel S.-H. Wang**

The cerebellum, most widely known for its role in motor functions, has also been associated with cognitive and affective functions of distant brain regions. This association has been substantiated with discovery of reciprocal loops between the cerebellum and neocortex. While the functional significance of these connections remains unknown, abnormalities in the loops are thought to be involved in the pathogenesis of Autism Spectrum Disorder (ASD). Of particular interest is the consistent loss of Purkinje cells (PCs) in cases of ASD. We hypothesize that manipulation of PCs during proposed sensitive periods of development will result in irreversible changes in the distant neocortical circuits manifested as atypical cognitive and affective behaviors in adulthood. Acute disruption of PCs in adult mice was accomplished using Designer Receptors Exclusively Activated by Designer Drugs (DREADDs) via lobule-specific injections to Lobule VI and Crus I. DREADD agonist clozapine-N-oxide (CNO) was then administered (i.p.) prior to behavioral testing, i.e. three-chamber social test and water Y-maze. Significant differences (Cohen's $d = 1.97$; $p<0.05$) in social preference (+26%) was found between Lobule VI perturbed mice and control mice. Additionally, Y-maze reversal learning was significantly impaired in Lobule VI ($p<0.001$) and Crus I ($p<0.001$) perturbed mice. Next, we will examine the effects PC disruption during development. This study will help elucidate the role of cerebellar involvement in distant brain formation and its role in ASD pathogenesis.

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

Simulation of *Myxococcus xanthus* Motility in a Wetted Environment

John McEnany, Farzan Beroz, Ricard Alert-Zenón, Chenyi Fei, **Ned Wingreen**

Myxococcus xanthus are rod-shaped bacteria characterized by their directional “gliding motility” and their numerous forms of multicellular aggregation, including roving swarms, multilayered terraces, and fruiting bodies – towers of cells that extend far into three-dimensional space. Since these cells can achieve multiple forms of large-scale order despite each bacterium only having access to local information, *Myxococcus* is an intriguing model for self-organization of active agents. Researchers at Princeton and elsewhere are conducting experimental and theoretical studies on *Myxococcus* organization in an effort to determine the factors that lead to larger-scale behavior such as layer formation. While several critical parameters have been identified, including cell velocity, reversal rate, and a surrounding water layer, it is not clear whether these are sufficient to explain the emergent behavior. An agent-based physical simulation offers unique potential to test the importance of these factors by precisely defining the forces acting on each cell and the resulting behavior of the bacteria, both on the collective and the individual level – in essence, simplifying the complex realities of biology into comparatively simple forces and interactions. We present a physical simulation of *Myxoccocus* based on one that has already shown success with the bacteria *Vibrio cholerae*, where each bacterium is represented as a rod that experiences elastic, adhesive, and active forces through contact with a substrate or to other cells. In addition, we introduce the water-mediated forces of surface tension and osmotic pressure, which are modeled in a novel manner through a convex hull surrounding the cells; these forces suggest that the interaction between cells and water plays an important role in the large-scale dynamics of *Myxococcus*. Eventually, we hope that this simulation sheds light on how selection pressures operated on properties of individual *Myxococcus* cells in order to produce complex and beneficial collective behavior.

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

The Role of Emotion Dynamics in Mental Models for Social Prediction

Isabel Medlock, Mark Thornton, **Diana Tamir**

How can people interact successfully with other people? To do so, they must understand others’ emotions and predict how they will change over time. The capacity to grasp the dynamics of mental states allows people to generate accurate mental models of how emotions change in category and intensity. To understand how people think about emotion dynamics we developed and deployed an online task to elicit people's’ perceptions of how emotions change in intensity over time. This task presents sixty mental states, ranging from friendliness to laziness to embarrassment to awe. Participants draw a graph of the intensity of each mental state for seventeen time points on a logarithmic time scale, starting at zero seconds and ending at two months. The resulting data will reveal how people think about the evolution of thoughts and feelings over time, and inform our understanding of social prediction.

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

Investigating the therapeutic potential of low-dose ketamine in an *in utero* valproic acid mouse model of Autism Spectrum Disorder

Kennedy S. Miller, Camden J. MacDowell, **Timothy J. Buschman**

Autism Spectrum Disorder (ASD) is an increasingly prevalent neurodevelopmental disorder estimated to affect up to 1 in 45 children in the United States. ASD is characterized by two core symptoms: deficits in social communication and restricted, repetitive behaviors. Therapeutic agents for ASD are limited; currently, there are no FDA-approved medications that alleviate these core symptoms. However, increasing research in both animal models and patients with ASD suggests that glutamate antagonists, including amantadine, memantine, and agmatine, may improve social deficits and stereotyped behavior. In this experiment, we test whether the glutamate antagonist ketamine, in subanesthetic doses (e.g. 'low-dose'), rescues impaired social interaction and repetitive motor behaviors in an *in utero* valproic acid (VPA) model of ASD. Experiments followed a two-by-two design and took place across two consecutive days in each animal. On day one, animals from either a VPA-exposed or saline-exposed (control) group underwent behavioral testing to assess social interactions (three-chambered social approach and social novelty assays) and repetitive motor behaviors (self-grooming and marble burying assays). Post-assessment, each animal was given an injection of either ketamine (subanesthetic dose: 50 mg/kg) or equivalent volume of vehicle (sterile saline). On day two, behavioral assays were repeated. Comparisons between groups were made with Mann-Whitney and ANOVA statistical tests. We observed limited ketamine-specific effects in either VPA-exposed or saline-control groups. These results suggest that low-dose ketamine has minimal-to-no therapeutic potential for ameliorating core ASD symptoms. Furthermore, these data conflict with previous studies demonstrating the efficacy of glutamate antagonists in rescuing behavioral deficits in animal models of ASD.

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

Leabra7: A Python Package for Hippo-cortical Interactions in Learning

Noam Miller, Daniel Greenidge, Qihong Lu, **Kenneth Norman**

There is burgeoning interest in the Complementary Learning System (CLS) hypothesis (McClelland, McNaughton, & O'Reilly, 1995), which posits that two complementary memory systems, neocortex (responsible for gradual learning about statistical structure) and hippocampus (responsible for "one shot" memorization of specific events), work together in learning. The body of neuroscientific literature supporting this hypothesis has been growing both in the clinical realm and the modeling realm, but models instantiating the CLS hypothesis have thus far been constructed using nonstandard software packages, making the verification of model results and the augmentation of models by other scholars difficult. This paper demonstrates a new Python software library that makes the popular biologically-based modeling software suite Emergent more easily understood and adaptable for modeling needs. Emergent uses an appealing neural activation and learning algorithm called Leabra (Local, Error-driven and Associative, Biologically Realistic Algorithm). Leabra incorporates numerous constraints from neurobiology and neurophysiology (e.g., separate populations of excitatory and inhibitory neurons) making it especially well-suited for modeling neural and behavioral data. Our "leabra7" Python package can easily be installed on any computer, and models can be shared through interactive Jupyter notebooks that can be run locally or remotely hosted on servers through Binder. In this paper, we demonstrate that a relatively complex model for hippo-cortical interactions (see Schapiro et. al., 2016) can be replicated in this framework. Though *leabra7* presently lacks the GUI from Emergent that helps in the construction and debugging of models, it offers many other valuable assets, including readability, share-ability, and adaptability, making it a useful platform for collaborative science.

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

The Statistical Physics of Intrinsically Disordered Proteins

Janelle Nelson, Benjamin Weiner, **Ned S. Wingreen**

Intrinsically Disordered Proteins (IDPs) are proteins that have no unique folded state but nevertheless perform many important functions inside cells. In particular, IDPs drive phase separation in membrane-less organelles, which store and process biomolecules. Such phase separation by IDPs depends on their thermodynamic properties, which are governed by an individual protein's sequence of amino-acid residues. Stretches of amino acids comprise domains that can fold and interact by forming bonds.

For this reason, we created model sequences of these interacting domains to compare their conformational statistics using a generalization of the self-avoiding walk. The conformation of a protein without bonds can be mapped to a self-avoiding walk (SAW) - a sequence of moves on a lattice of vertices that does not visit the same vertex more than once. We used a Recursive Algorithm to enumerate all random walks in which a walk can visit a vertex twice, corresponding to a bond.

The purpose of our project is to understand the statistical properties of the spatial conformations adopted by distinct sequences. For each sequence, each random walk has an energy depending on the number of bonds. Our hypothesis is that as the entropy increases, the free energy of a sequence linearly decreases. Using our enumerated states, we calculated thermodynamic quantities like the entropy and free energy to better understand the relationship between sequences, individual protein properties, and phase separation. These quantities will help us to further understand how IDPs drive phase separation to form membrane-less organelles.

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

Structural Determination of Microtubule Nucleation Factors

Phoebe Nelson, Brian Mahon, **Sabine Petry**

Microtubule (MT) nucleation during spindle formation is a process essential to eukaryotic cell division. MTs are composed of α/β tubulin heterodimers and form via nucleation at chromosomes, centrosomes, and preexisting MTs. Nucleation from preexisting MTs, referred to as branching, allows for exponential MT growth that is necessary for efficient formation and organization of the mitotic spindle. MT branching is facilitated by the recruitment of the γ-tubulin ring complex (γ-TuRC) to a preexisting MT where it acts as a template from which a new MT can nucleate and form. Recent evidence has shown that nucleation factors, such as the mitotic-spindle organizing protein associated with a ring of γ-tubulin (Mozart 1) and the targeting protein for Xklp2 (TPX2), interact directly with γ-TuRC to mediate MT branching. However, the structures of Mzt1 and TPX2 are currently unknown, resulting in poor understanding of their interactions with γ-TuRC at the MT. In this study, we will address this lack of structural information regarding Mzt1 and TPX2 in order to gain insight into how they interact with γ-TuRC at the branching site. To do this, we will use X-ray crystallography to determine atomic resolution 3D structures of recombinantly expressed and purified constructs of both Mzt1 and TPX2. From these high-resolution structures, we will be able to understand the morphological implications of their roles in the branching process. Overall, this will bring us a step closer to understanding the fundamentals of cell division in molecular detail.

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

How do cells create enhancers?
Inducing H3K4me1–H3K27ac crosstalk at enhancers in live cells with protein trans-splicing

Eva Parisi, Antony Burton, Michael Haugbro, **Tom Muir**

Histone post-translational modifications (hPTMs) mediate the functional organization of chromatin and are critical for epigenetic regulation. Increasing evidence suggests that aberrant gene expression patterns associated with cancer may result from chromatin dysregulation at the hPTM level. Moreover, the disruption of regulatory DNA elements like enhancers has been implicated in numerous diseases. Recent studies have identified H3K4me1 and H3K27ac as the epigenetic signature of active enhancers. However, the specific sequence of events that creates this signature remains unknown, while available technologies to investigate hPTMs are limited. The proposed study will deliver specific hPTMs associated with enhancers to live cells, and elucidate the mechanism by which active enhancers are established. We hypothesize that increased H3K4me1 triggers the installation of H3K27ac to build active genomic enhancers. Preliminary findings have characterized an ultra-fast split intein, which will be used to perform protein trans-splicing (PTS) to selectively install H3K4me1 in live cells. Thus far, we have used PTS to deliver multiple modification states of H3K4 to nuclei isolated from mammalian cells. We are currently adapting this method for use in intact live cells. We will then use pull-down assays, Western blotting and ChIP-qPCR to measure the downstream effects of H3K4me1 installation. The expected results will yield valuable insights into the process of enhancer creation and will provide a chemically precise tool for customizing hPTMs in physiologically relevant settings.

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

Optogenetic Regulation of Biofiuel Pathways: Control of Isopentanol and 2-methyl-1-butanol Biosynthesis with Blue Light

Olivia Parker, Sarah Hammer, **José Avalos**

Concerns about energy security and climate change have motivated efforts to identify renewable sources of fuel. Engineering the yeast *Saccharomyces cerevisiae* for biofuel production is a potential solution. In the Avalos lab, we use metabolic engineering to modify yeast for the production of renewable fuels at high yields and titers. Although ethanol is the main fuel naturally produced in yeast, the cells also produce branched-chain higher alcohols (BCHAs) in trace amounts through amino acid degradation. These BCHAs are of interest because of their higher energy density and lower corrosivity as compared to ethanol. Previous research in our lab has shown that (a) localizing amino acid degradation pathways to the mitochondria boosts BCHA production, and (b) optogenetics can be used to regulate metabolic pathways, specifically by exposing cells to blue light. In this project, we use the OptoEXP circuit (activates pathways in blue light) and the OptoINVRT7 circuit (activates pathways in the dark) to control isopentanol and 2-methyl-1-butanol production, respectively. The aim of my project is to control the ratios of isopentanol and 2-methyl-1-butanol generated in one yeast strain with blue light and in the absence of light.

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 Novel Mutation, *rosette*, as a Potential Global Cue in the Core Planar Cell Polarity Pathway

Brooke Phillips, Maureen Cetera, **Danelle Devenport**

Morphogenesis, the shaping of an animal during development, relies in part on the Planar Cell Polarity (PCP) Pathway. PCP is defined as the coordinated organization of cell polarity across a tissue plane. Mutations in this pathway can have detrimental effects on neural tube closure, inner ear patterning, and hair follicle orientation in the epidermis. While the core proteins involved in PCP have been well studied, the global cues that allow this pathway to regulate polarity across an entire tissue system remain elusive. A novel mutation, *rosette*, causes a unique posterior-restricted hair follicle orientation defect in mice as well as neural tube closure abnormalities at a low frequency. Analysis of core PCP protein localization in the epidermis of *rosette* mutants shows a maintenance of local polarity, but a lack of global polarity. This mutation's ability to disrupt global polarity makes its potential function as a global cue promising. To elucidate *rosette*'s possible interaction with known core PCP genes – Vang-Like 2 (Vangl2) and Frizzled6 (Fz6) – we produced double heterozygous animals with one copy of the *rosette* mutation and one copy of a mutation in a core PCP gene. We then studied how the combination of these mutations affected hair follicle orientation in the mouse epidermis. We found that *rosette* genetically interacts with Fz6 to cause a hair follicle orientation defect similar to that observed in *rosette* homozygous mutants. This genetic interaction suggests that *rosette* plays a crucial role in coordinating PCP across the epidermis. Future experiments will investigate *rosette*'s potential genetic interaction with core PCP genes in other PCP-dependent tissue systems. A better understanding of this potential global cue could contribute to a more complete model of the PCP pathway.

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

Investigating the Role of PARP1 in the Context of Oncohistone Mutations

Anagha Prasanna, Krupa S. Jani, Katharine Diehl, Glen Liszczak, **Tom W. Muir**

Poly(ADP-ribose) polymerase (PARP) is a histone modifier known to play a role in DNA damage repair by PARylating serines on histone H3 tails within an RKS sequence motif. On the canonical histone H3, the primary substrate is S10. Acetylation on the "-1" lysine (H3K9) has been observed to inhibit PARP activity, raising the question of what other changes to the sequence motif can affect PARP. Lysine-to-methionine (K-to-M) mutations on key regulatory lysines H3.3K27 and H3.3K36 have been identified as probable driver mutations in several pediatric cancers. This project aims to understand how the ability of PARP1 to PARylate histones is altered in the context of K-to-M histone mutants. PARP1 activity was therefore measured on H3.3 histones containing following mutations: H3.3K9M, H3.3K27M, and H3.3K36M. Preliminary results suggest that H3.3K9M is associated with decreased PARP1 activity. The cancer-associated K-to-M mutations are primarily found on histone variant H3.3, which contains an additional serine residue at position 31 compared to canonical H3. This project also aims to assess whether this S31 residue is a substrate for PARP1 to better understand the relationship between PARP and oncohistone mutations.

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

Identification of Novel *Caulobacter crescentus* Width and Length Genetic Determinants Using a Transposon Library Screen

Justin Ramos, Gabriel Moore, **Zemer Gitai**

Bacteria come in a diversity of cell shapes that facilitate their ability to thrive in a variety of environments. Specifically, the length and width of a given bacterium are important for maintaining a favorable surface area to volume ratio, which aids the bacterium in nutrient intake. Additionally, factors which regulate width and length serve as antibiotic targets because they are unique to bacteria. An excellent model organism to study cell shape is *Caulobacter crescentus*, an aquatic, rod-shaped bacterium. *Caulobacter's* shape determinants have been the subject of research for many years. However, no study thus far has comprehensively surveyed all possible *Caulobacter* mutants for defects in width and length. My project aims to assemble an ordered transposon library, which contains at least one mutant in every non-essential *Caulobacter* gene and identify mutants with defects in cell length and width. I performed a qualitative microscopy screen of over 4000 mutants to assemble a list of *Caulobacter* genes related to the two shape dimensions. From the screen, I have identified over fifty candidates with visible shape defects. Many of these factors are uncharacterized and conserved among other Gram-negative bacteria. Their conservation indicates they may play a fundamental role in cell elongation, division, and life cycle regulation. These novel factors also provide an opportunity to elucidate the full mechanism of shape determination and grant insight into new antibiotic targets.

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

Requirement for Heparin Interactions with Fibronectin Module III-13 during Polymerization and Cell Association

Ellie Randolph, Katherine Hill, **Jean Schwarzbauer**

The extracellular matrix (ECM) provides a dynamic environment and structure for many tissues within the body. The glycoprotein fibronectin (FN) is a vital component of the ECM, creating a foundation for other secreted proteins and directly interacting with cells. FN-FN interactions facilitate matrix formation, and FN-cell interactions trigger cytoskeletal organization and formation of focal adhesions. Aberrant FN matrix formation can lead to diseases such as fibrosis. FN type III repeat 13 (III-13) is part of the binding domain for glycosaminoglycan heparan sulfate (HS). This module is also hypothesized to participate in FN-FN interactions via β-strand exchange. These dual functions may overlap, as binding to heparan sulfate could loosen III-13 hydrophobic interactions thereby facilitating β-strand swapping and polymerization. To test this hypothesis, the III-13 repeat was cloned into a vector containing a GST purification tag. The protein was expressed in dh5alpha *E.coli* and then purified. The GST-III-13 construct will be used to study β-strand exchange in the presence and absence of HS using fluorescence measurements, dye binding assays, and electron microscopy. These experiments will establish a new role for heparan sulfate in the β-strand exchange model.

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

Investigating the proteome partitioning between the nucleus and cytoplasm in using DNA coupled magnetic beads

Reyes, José, Nguyen, Thao, **Martin Wühr**

Despite the central role of nucleus in multicellular biology, its composition in many cell types and the molecular mechanism leading to such establishment are not fully understood. Recent study has shown that a large fraction of nuclear trafficking through the nuclear pore complex is via passive diffusion, suggesting nuclear retention via protein interaction with DNA is crucial to nuclear proteome establishment. Furthermore, not same-sized proteins have the same nuclear composition; hence we hypothesize that the differential of protein affinity to DNA would dictate the dynamicity in nuclear composition in different states and conditions. The classical *Xenopus Laevis* egg extract and the recent mass spectrometry technology provide a powerful biochemical tool to investigate this hypothesis on a system-wide scale. The extract system is essentially undiluted cytoplasm which can recapitulate many cellular processes *in vitro*, including the formation of nuclei. Here, using DNA coupled magnetic beads, we will vary the nuclei concentration, rapidly isolate and subject the nuclei to a mass spectrometry-based proteomics experiment. We aim to quantify the relative affinity of proteins to DNA under various conditions and hope to elucidate the regulatory mechanism of nuclear organization in eukaryotic systems.

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

Dendritic Disasters : Discovering the role of Muscleblind in Dendritic Arbor Formation

Denay Richards, Rebecca Alizzi, **Elizabeth Gavis**

Dendritic branching is crucial for the ability of neurons to receive and process information. When branching is disrupted, peripheral neuropathies can result, as demonstrated by the muscle wasting disease myotonic dystrophy (MyoD). Our lab uses the *Drosophila* larval class IV dendritic arborization (da) neurons to study the role of RNA-binding proteins in regulating dendrite branching and morphology. The class IV da neurons are a group of highly-branched sensory neurons that innervate the epidermis and respond to thermal and mechanical stimuli. These neurons form a sheet in between the epidermis and the extracellular matrix (ECM) and can occasionally become enclosed in the surrounding epidermal cells. Enclosure relies on the localization of the cell adhesion protein Coracle (Cora) to the dendrite – epidermal interface. Proper morphology of these neurons is dependent on both the interactions between the dendrites, ECM and epidermis, as well as the correct balance of stable microtubules and actin filaments within the dendrite branches. Here we show that the RNA binding protein Muscleblind (Mbl) is critical for proper development of class IV da neurons. *mbl* knockdown using RNAi causes a reduction in arbor coverage, along with an increase in terminal branch density. Nociception tests showed that *mbl* knockdown leads to a hyperactive response to a global heat stimulus, suggesting that branch morphology could affect larval response to thermal stimuli. Based on previous work with Mbl from other labs, we hypothesize that Mbl regulates *coracle* in order to control dendrite-epidermal interactions, and the actin crosslinker *α-actinin* in order to alter the cytoskeletal composition of the arbor. By regulating these transcripts, Mbl could function to promote proper branching in class IV da neurons. It is our hope that our findings can help us to better understand the mechanism by which Mbl knockdown leads to peripheral neuropathies, as seen in MyoD.

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

Genetic and Biochemical Analysis of MAB21 Proteins in Neuroblastoma Proliferation

Joseph Ryu, **Chandra Theesfeld**

Neuroblastoma is the most prevalent and deadly form of childhood cancer with limited treatment options. To facilitate study of this rare disease and hundreds of others, we developed a novel computational approach, URSA^{HD}, that compares gene expression signatures of one disease to over 300 other diseases and normal tissues. The URSA^{HD} neuroblastoma model identified MAB21L1 and MAB21L2 as unique to neuroblastoma. The MAB21 proteins play a critical role in limb and nervous tissue development across animal models and humans, but there is no documented link between the MAB21 proteins and neuroblastoma, nor is there clear evidence for biochemical function of the MAB21 proteins. Our initial studies indicated that MAB21L1 was essential for growth while MAB21L2 might have activity countervailing MAB21L1. Our study aims to test this model through siRNA knockdown experiments and MAB21L2 overexpression experiments. Structural similarity to the DNA-sensor, cGAS, suggests the MAB21 proteins may function through binding nucleic acid. To test this, we will perform in vitro DNA/RNA binding assays for MAB21L2, as well as cell fractionation experiments to clarify its functional location. Clinical MAB21L2 mutations lead to ocular and skeletal abnormalities but the molecular defects of the protein are uncertain, and it is unknown if these mutations are relevant to neuroblastoma proliferation. We will clone these mutations and characterize the effects of the mutations in vivo and in vitro using genetic and biochemical assays, respectively. Together, these experiments will shed light on the genetic and molecular function of the MAB21 proteins and their role in neuroblastoma.

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

Trust but Verify: Analysis of Flawed Trust Beliefs for Tor and Optimal Trust Modeling

Nicholas Schmeller, Yushan Liu, **Prateek Mittal**

The Tor Project is the most widely used anonymous communications and ensures the online privacy of journalists, political dissidents, and the public. Adversaries can break this privacy through deanonymization attacks by controlling relays on the network. Trust-aware path selection algorithms attempt to route traffic away from compromised relays, but these algorithms rely on a policy to establish the trustworthiness of each router. Some new policies allow users of the network to establish their own trust policies based on strong trust beliefs. We characterize the effectiveness of routing other users by selecting a malicious trust belief. We then present a novel, generalizable trust model based on work by Jaggard et al., an implementation of the most promising model, and significant privacy-improvement results after a simulated deanonymization attack. Additionally, we make recommendations for the implementation of trust models for The Tor Project in general.

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

Developing a Tool to Organize and Annotate Gene Expression Metadata

Yoni Schoenberg, Adam Oberstein. **Thomas Shenk**

p53 and Retinoblastoma (Rb) are cell cycle regulatory genes whose inhibition can contribute to tumor formation. Adenovirus 5 (Ad5) and Human cytomegalovirus (HCMV) both produce proteins that interact with and inhibit p53 and Rb function. Ad5, however, has been shown to induce transformation in cells, while HCMV has not. Gene expression dataset analysis from studies of Ad5 and HCMV infected cells can elucidate why inhibition of these tumor suppressors transforms cells after infection of one virus but not the other. In order to determine differences in the expression of p53 and Rb downstream genes, numerous studies must be analyzed and general trends of gene expression by the viruses must be determined. However, there is currently no straightforward way to easily and efficiently analyze gene expression data across multiple studies. Here, we describe a novel approach to organizing and categorizing studies of virally infected cells. Raw metadata of all studies with sequencing data were taken from the NCBI database. From there, the metadata were processed using the R programming language. After processing and annotating this metadata for the desired studies, the gene expression data from these studies can be used to form gene clusters of over and under-regulated genes. Conclusions from these analyses will help to better elucidate the mechanism in which inhibition of p53 and Rb leads to tumor formation.

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

Investigating the relationship between mRNA localization and fertility in *Drosophila*

Sara Shatela, Matthew G. Niepielko, Caroline A. Doherty, **Elizabeth R. Gavis**

RNA localization is a highly conserved strategy used to generate polarized cell function in a variety of cell types. In the *Drosophila* embryo, *nanos (nos)* and *polar granule component (pgc)* mRNAs localize to the germ plasm, a highly specialized cytoplasm at the embryo posterior that is both necessary and sufficient to induce the formation of primordial germ cells. During the localization process, *nos* and *pgc* incorporate into higher order structures called germ granules. Within these granules, these RNAs organize into spatially distinct clusters, called homotypic clusters, each containing multiple copies of the same RNA. Although the complete loss of *nos* localization is associated with *Drosophila* female sterility, the relationship between the sizes of homotypic clusters and fertility robustness remains unclear. Using single molecule Fluorescence In Situ Hybridization (smFISH) to determine cluster size and genetic techniques to reduce the sizes of *nos* and *pgc* homotypic clusters, we investigated the role homotypic clustering has in fertility. Our data demonstrate that reducing *pgc* cluster size increases primordial germ cell number and the number of eggs laid while reducing *nos* cluster size had little effect. Our future studies will focus on reducing *nos* homotypic cluster size further to determine if there is a clustering threshold that can confer robust egg laying. In addition, we will investigate the mechanisms by which *pgc* reduction generates more primordial germ cells.

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

Characterizing the mechanism of the Bam complex during outer membrane protein assembly

Jessica Sheng, **Thomas J. Silhavy**

Gram negative bacteria are especially resistant to antibiotics due to an additional outer membrane that surrounds their peptidoglycan wall. Outer membrane proteins (OMPs) are transmembrane proteins found in this outer membrane. Their synthesis and assembly are mediated by the Bam complex, which is a five-subunit complex in the outer membrane that takes unfolded OMPs and assembles them for insertion into the outer membrane. The sigma(E) (σ^E) stress response of *Escherichia coli* (*E. coli*) is a pathway that is activated by accumulation of unfolded OMPs, and it relieves this toxic buildup by decreasing flux through the OMP biogenesis pathway and cleaving unfolded OMPs for disposal. Understanding the Bam complex and σ^E mechanisms would help develop potential antibiotics that target OMP biogenesis. Our lab has generated the temperature-sensitive *lamB*$_{G439D}$Δ*degP* double mutation that causes cell death due to accumulation of the unfolded OMP LamB. Screening is ongoing for suppressors of this double mutation; changes in the Bam complex that restore LamB folding will shed light on its mechanism. To study the Bam complex from another angle, Merck & Co., Inc. is providing our lab with compounds that potentially target the Bam complex. Currently, I have identified a mutation in BamA, *bamA*$_{E470K}$, that confers resistance to the compound L-494. The next steps will be to characterize how it affects the structure and mechanism of the Bam complex. This research will contribute to our understanding of the mechanism of OMP biogenesis and to development of potential antibiotics that target this pathway.

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

PAC-Bayes Control: Synthesizing Controllers that Provably Generalize to Novel Environments

Anoopkumar Sonar, Maxwell Goldstein, **Anirudha Majumdar**

Our goal is to synthesize controllers for robots that provably generalize well to novel environments given a dataset of example environments. The key technical idea behind our approach is to leverage tools from generalization theory in machine learning by exploiting a precise analogy (which we present in the form of a reduction) between robustness of controllers to novel environments and generalization of hypotheses in supervised learning. In particular, we utilize the Probably Approximately Correct (PAC)-Bayes framework, which allows us to obtain upper bounds (that hold with high probability) on the expected cost of (stochastic) controllers across novel environments. We propose control synthesis algorithms that explicitly seek to minimize this upper bound. The corresponding optimization problem can be solved using convex optimization (Relative Entropy Programming in particular) in the setting where we are optimizing over a finite control policy space. In the more general setting of continuously parameterized controllers, we minimize this upper bound using stochastic gradient descent. We present examples of our approach in the context of obstacle avoidance control with depth measurements for a GhostMinitaur Quadruped. Our simulated example of the Minitaur inside the Stanford 3D Indoor Spaces Database demonstrates the potential of our approach to provide strong generalization guarantees on controllers for robotic systems with continuous state and action spaces, complicated (e.g., nonlinear) dynamics, and rich sensory inputs (e.g., depth measurements).

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

Mapping the adhesion profile of hair follicles during placode polarization

Yanilka Y. Soto Muñiz, Liliya Leybova, **Danelle Devenport**

Self-sorting through different cadherin expression is a common mechanism used in the developing embryo to prevent cell mixing. Surprisingly, cell mixing is also prevented in the mammalian skin during hair follicle morphogenesis. Hair follicles develop from multicellular placodes that emerge from the embryonic epidermis with two cell types, inner and outer cells. Shortly after placode formation, inner and outer cells undergo opposing directional movements coordinated in a counter-rotational pattern of cell flow. However, the two populations never mix with one another nor with the surrounding interfollicular epidermis, but the reason remains unclear. Ultimately, these rearrangements lead to the polarization of the placode where it obtains an anterior tilt. Previous experiments have shown that in the absence of outer cells, such as in Shh mutants, proper polarization does not occur. These mutants present an irregular arrangement of cells, where the inner cells are poorly segregated from their neighbors. Differential expression of cadherins play a role in organizing tissues and preliminary data suggest that inner cells express high levels of P-Cadherin and low levels of E-Cadherin, while in the outer and IFE cells display inverse levels. Using immunofluorescence and confocal imaging, we have measured the fluorescence intensity of E-Cadherin and P-Cadherin to determine the mechanism that prevents cell mixing within the placode. Our data revealed that these proteins are expressed in a gradient within the placode. We hypothesize that differential cadherin expression compartmentalizes the placode into two distinct lineages that do not mix with each other or the surrounding IFE and in the absence of outer cells adhesion differences are not as pronounced allowing cell mixing to occur.

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

Characterizing the functions of motifs A, B, and C and identifying the nuclear localization signal in *Saccharomyces cerevisiae* Pif1 helicase

Ashley Stone, Carly Geronimo, **Virginia Zakian**

Pif1 is a multifunctional DNA helicase with both nuclear and mitochondrial functions that promote genome integrity. Pif1 is highly conserved, including among *Saccharomyces cerevisiae*, bacteria, and humans, and is part of the helicase Superfamily I whose members share motifs in the helicase domain. Various functions in *S. cerevisiae* Pif1 (ScPif1) have been well-studied, but the functions of motifs A, B, and C, which it shares with *E. coli* RecD helicase, have not been determined. The first aim of this study is to identify the functions of motifs A, B, and C by mutating conserved residues and assaying for mitochondrial function and for nuclear functions including telomerase inhibition and Okazaki fragment processing. In addition, the nuclear localization signal (NLS) of ScPif1 has not yet been characterized, so the second aim of this study is to identify the NLS. The NLS will be identified by mutating conserved residues in the predicted sequence (KKRK) near the carboxy terminal end and visualizing ScPif1-GFP fusion proteins by confocal microscopy. The findings of this study would identify the NLS and determine the functions of motifs A, B, and C, contributing to the field's understanding of ScPif1 functions.

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

Autoagglutination of O1 biovar El Tor *Vibrio cholerae* strain

Dominique Summerville, Matthew Jemielita, **Bonnie Bassler**

Vibrio cholerae, the pathogen responsible for the disease cholera, relies on the formation of multicellular-communities, such as biofilms, for disease progression. The formation of surface biofilms is regulated by quorum sensing (QS), a method in which bacteria communicate with one another. Through QS, low cell density (LCD) environments promote biofilm formation while high cell density promotes biofilm dispersion. Previous work in the Bassler lab identified formation of suspended cell aggregates of *V. cholerae* at high cell density QS state. This regulates opposite the known process of surface biofilm formation. Activator genes of aggregation were identified through screening. I successfully saturated a genetic screen for repressor genes of aggregation in LCD locked QS state *V. cholerae*. Specifically, I found 3 genes that repress autoagglutination, defined as cells sedimenting in liquid medium. While natural, virulent strains of *V. cholerae* naturally autoagglutinate, the current, disease causing strain I work with, O1 biovar El Tor, does not. I plan to create mutants of the *tcpA* pili gene, an essential virulence factor responsible for autoagglutination in the other virulent strains, and other pili genes to determine whether the autoagglutination phenotype is caused by the same factors. I created a mutant of the *mshA* pili gene that caused no change in the autoagglutination phenotype. I plan to research the change over time of cell network structures involved in autoagglutination. This research on the phenotype of autoagglutination will possibly shed light on the mechanism of HCD aggregate formation and increase understanding of *V. cholerae*'s life cycle.

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

Mapping intracellular HCMV infection using a computational and proteomic approach

Samvida Venkatesh, Joel Federspiel, Pierre Jean-Beltran, **Ileana Cristea**

Human cytomegalovirus (HCMV) is among the most prevalent infectious agents in humans and is life-threatening in immunocompromised individuals. For successful infection and spread, the virus must reorganize and redirect the functions of various host cell organelles at the right times, yet molecular mechanisms of cellular infection remain to be elucidated. To see these organelle alterations during HCMV infection from a proteomic perspective, I constructed a web-accessible visualization of translocations of host and pathogen proteins with previously determined protein localization data during infection. The versatile tool is not limited to this dataset alone, as it allows users to input any proteomic data with subcellular localization information to highlight dynamic spatiotemporal localization in their own data. By visualizing the spatial and temporal proteome of HCMV-infected cells, we uncovered an important role for a less-studied organelle in infection, the peroxisome. Lipid metabolism in the peroxisome is thought to be important for enveloped viruses such as HCMV to form their host-derived lipid envelope, yet peroxisomes remain poorly characterized as they cannot be adequately resolved by biochemical fractionation. Therefore, I adapted a proximity labeling approach using APEX to identify host and viral proteins in the peroxisome throughout infection. APEX is an engineered peroxidase that biotinylates proteins in its vicinity upon the addition of biotin-phenol. The newly biotinylated proteins can then be enriched by immunoaffinity purification against the biotin tag and identified by mass spectrometry. I demonstrated that APEX constructs can be modified to include a peroxisomal targeting signal and consequently be localized to peroxisomes via transient transfection. Experiments utilizing APEX in HCMV-infected cells to identify and functionally characterize proteins localized to the peroxisome are ongoing. Through an integrated visualization tool and peroxisomal proteome-labeling method, I have provided novel techniques to illuminate the complex proteome of HCMV-infected cells.

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

Automated tracking of freely-moving mouse behavior

Xiaoting Sun, Jessica Verpeut, Mikhail Khislin, Lindsay Willmore, **Samuel S.H. Wang**

The mouse, as a model organism, is frequently used to study behavior, yet the details of their movements and postures have not been analyzed. Simpler organisms with smaller sets of movements and postures such as *C. elegans* and, recently, *Drosophila melanogaster* [1] are best suited for detailed computational analysis to quantitatively define and examine distinct behaviors. However, there is a need for the equivalent in mice, due to their importance in clinical studies. Neural perturbations relevant to humans often show subtle effects not easily captured by current methods — manual scoring is inefficient and potentially biased, while automated systems use coarse metrics as proxies for certain behaviors. Our goal is to design a quantitative, unsupervised pipeline and create a baseline behavior map for the individual mouse. Preliminary testing of LEAP, a deep learning neural network for body position estimation[2], showed that it is capable of tracking basic gait behavior. 120 hours of video, across 36 mice in an open field and linear track were recorded. A network was trained on the wide range of actions from the open field with 600 frames and moderate tracking success, but currently needs more tuning. The next steps are to develop the analysis pipeline and further characterize behavior by analyzing mouse posture in 3D, by incorporating both bottom and side-view recordings from linear track data. These recordings will create a map of stereotypical mouse behavior, which will then be applied to future studies of neural manipulations.

This research was made possible by the generous support of the Lewis-Sigler Institute for Integrative Genomics as well as guidance from Samuel S.-H. Wang, Jessica Verpeut, Mikhail Kislin, and Thomas Kellogg.

Sirtuin-3 vs. pUL37 in mitochondrial changes during human cytomegalovirus infection

Amy Tien, Xinlei Sheng, Joel Federspiel, **Ileana Cristea**

Human cytomegalovirus (HCMV) replication induces significant mitochondrial changes including fragmentation and localization to the virion assembly complex periphery. We aim to understand these mitochondrial changes from both a viral and host perspective. Mitochondrial NAD^+-dependent deacetylase sirtuin-3 (SIRT3) has been shown to be an antiviral factor against DNA viruses but the underlying mechanism remains unknown. Earlier lab work has established a dynamic mitochondrial acetylome and SIRT3 interactome during HCMV infection, suggesting deacetylase involvement in SIRT3-mediated defense. To address this, we aimed to create enzymatically inactive SIRT3-H248Y mutants and determine the consequences of diminished mitochondrial deacetylation in the context of HCMV infection via viral titer. Moreover, this project aims to investigate the interplay between viral proteins and mitochondrial host defense factors. We postulate that SIRT3 may execute its antiviral functions by counteracting mitochondrially localizing viral factors such as viral protein pUL37, whose necessity in viral replication and inhibition of apoptosis has been well characterized. Previous lab work has found metabolic and apoptotic pathway commonalities shared between the pUL37 and SIRT3 interactomes during infection. Our analysis of the pUL37 interactome revealed temporally regulated pathway associations in line with progression of viral replication needs. Normalization of pUL37 interactions to mitochondrial proteome abundances during infection found early enrichment of interactants involved in mitochondrial transport. Intermediate and late infection saw enrichment of metabolic and apoptotic pathway associated interactants, respectively. To further investigate pUL37 interaction partners of interest, we aimed to create lentiCRISPR-mediated knockouts of interactants and determine the effects of the disruption on HCMV infection. A better understanding of the SIRT3/pUL37 interaction underlying mitochondrial changes during infection will highlight pathways involved in host defense and infection progression, providing new targets for antiviral therapies.

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

Optogenetic Control of Apical Contraction in the *Drosophila* embryo

Annan Timon, Heath Johnson, **Jared Toettcher**

For *Drosophila* embryos to undergo the precise morphological changes necessary to set up a functional body plan, individual cells must first receive signals in order to initiate the process of tissue folding. In the complex and dynamic environment of the early embryo, these signals occur at specific locations and at exact times during the organism's developmental timeline. We are interested in the factors contribute to this specificity and how they fit into the tissue folding pathway. Here, we look at cellular constriction and invagination, specifically at ventral furrow (VF) and the posterior midgut (PMG) and the secreted autocrine signaling protein Folded gastrulation (Fog). While the Fog signal is necessary to induce these cellular movements, it is not sufficient to induce changes without serious spatial and temporal limitations. In fact, rearrangements such as constriction and invagination demonstrate a requirement for the colocalization of Fog and its recently discovered GPCR, the mesoderm-invagination signal transducer (Mist). Despite this, it is still relatively unknown whether there are other factors that constrain constriction in both space and time. Characterizations such as *mist* granule release hint at possible posttranscriptional regulation of the pathway. Therefore, we propose a novel approach to understanding the Fog and Mist signaling pathway using optogenetic technology and high-resolution microscopy. We will construct an optogenetic system for activating *fog* and *mist* in vivo to study tissue folding with high spatiotemporal precision and resolution. With these tools, we will be able to precisely visualize *fog* and *mist* dynamics. If successful, our approach will not only answer the question of *fog* and *mist* sufficiency for constriction, but also demonstrate the efficiency of generalized optogenetic transgene expression systems.

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

Investigating Expression of Tenascin-C Isoforms in Diabetic Conditions

Awele Utomi, Maria E. Vega**, Jean E. Schwarzbauer**

Diabetic nephropathy (DN) results from excess accumulation of the mesangial extracellular matrix (ECM) proteins, which leads to end-stage renal disease. Pathological conditions caused by high glucose levels within the diabetic kidney have been shown to increase assembly of extracellular fibronectin (FN) matrix, a key component of the mesangial ECM. In our prior research, we looked for changes in gene expression induced by high glucose. We used immortalized mouse mesangial cells as an *in vitro* model of DN paired with Next Generation Sequencing and while there was no change in FN gene expression, we found 3-fold increased expression of another ECM protein, tenascin-C (TN-C). The TN-C transcript is alternatively spliced to generate at least five isoforms that vary depending on tissue type. Analysis of TN-C revealed that high glucose promotes selective expression of a large isoform, whereas a smaller isoform is prominent in normal glucose conditions as it is incorporated into the matrix. Next, we investigated the alternative splicing of TN-C so as to determine the exact isoforms that are expressed in normal and high glucose conditions. Isoform-specific primers are being used for PCR amplification of distinct splice variants which are identified by sequencing. Analysis of TN-C alternative splicing is already underway in other diseases, primarily in cancers. For the treatment of high grade glioma, some clinical trials have capitalized on the fact that glioma cells selectively express a specific TN-C isoform; antibodies targeting the alternatively spliced region are conjugated to radioactive ions, allowing for very specific delivery of radiation to cancerous cells. Overall, our study hopes to identify these alternatively spliced sites so as to potentially enable pharmaceutical manipulation of the model which would be a reasonable approach to attenuate mesangial ECM expansion.

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

Localizing the Most Highly Conserved Proteins in Photosynthetic Organisms

Kelly Van Baalen, Liangyong Wang, Yihua Xie, **Martin Jonikas**

Global food demand is projected to double by 2050. To meet this increased demand without causing further environmental degradation by putting additional land under cultivation, yields must be increased on existing agricultural land. One promising route for increasing yields is to genetically modify crop plants to have greater photosynthetic efficiency. Progress has already been made in improving the photosynthetic efficiency of some plants through genetic editing and more significant increases are possible through more complex editing. However, such improvements have been constrained by the fact that the identities and specific functions of many genes with putative roles in photosynthesis are currently unknown. Here we consider 271 proteins from the GreenCut2, a list of proteins conserved in photosynthetic organisms, as candidate proteins for involvement in photosynthetic function. We use a high throughput pipeline to produce *Chlamydomonas reinhardtii* strains expressing fluorescently tagged GreenCut2 genes. We image these strains using a confocal microscope to determine the localization of the GreenCut2 proteins. In a test batch of ten GreenCut2 proteins, two strains with fluorescently tagged proteins were successfully produced and had sufficient fluorescent signal for imaging. We observed that both of these proteins localize to the chloroplast and characterized their sub-chloroplast localizations. Our work expands our knowledge of the components of photosynthesis and provides valuable clues as to their functions, bringing us closer to the understanding necessary to engineer photosynthesis to be more efficient.

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

Determining a role for *Lmx* in neural tube formation

Nancy Wenger, Laurence Lemaire, **Michael Levine**

Neural tube formation is a highly regulated process in the timing of cell division, orientation, and movement and involves both environmental cues and genetic factors. Perturbation of its closure leads to *spina bifida,* a major birth defect in humans. However, a comprehensive mechanistic gene network has yet to be fully elucidated due to the complexity of vertebrate systems. *Ciona intestinalis* serves as a model organism for the purpose of this research because of its highly simplified neural tube and fixed embryonic cell lineage. The transcription factor *Lmx* is expressed in the anterior dorsal neural tube cells in *Ciona* and has shown an intercalation defect when down-regulated. To gain more insight into the role for *Lmx* during neural tube formation, we ectopically expressed the gene in adjacent epidermal borders using a reporter construct driven by the gene *Klf*. *Lmx* overexpression in this region shows dorsal curvature of the tail instead of ventral and in most cases, an open tube. Analysis using two-photon microscopy shows that the rate of cell division and its orientation along the anterior-posterior axis is affected, causing the dorsal epidermis to shorten rather than elongate. This preliminary data suggests that *Lmx* may be involved in cell cycle regulation. Future experiments will examine the mechanism interrupted by increased *Lmx* activity in the epidermal cell lineage. Single cell transcriptome data will identify gene activity co-localized to that of *Lmx* and provide a stepping stone for isolating its downstream targets.

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

Functional analysis of *germ cell less* during germ cell formation and specification in *Drosophila melanogaster* embryos

Lillian Wilkins, Melecia Wright, Yunah Kim, **Girish Deshpande**

Transcriptional quiescence is an evolutionarily conserved trait that distinguishes the embryonic Primordial Germ Cells (PGCs) from their somatic neighbors. In *Drosophila melanogaster* embryos, Germ cell less (Gcl) protein is required for proper pole cell formation and the establishment and/or maintenance of transcriptional quiescence in PGCs. Thus, pole cells from *germ cell less* (*gcl*) embryos mis-express several somatic genes including two X-linked numerator elements, *sisterless-a* (*sis-a*) and *sisterless-b* (*sis-b*). These two proteins are shown to activate transcription of *Sex-lethal* (*Sxl*), a sex-determination gene that orchestrates female identity in the somatic nuclei. Importantly, wild type pole cells from blastoderm stage embryos do not express *Sxl* and are naïve with respect to their sexual identity. Here we have examined the influence of loss of *gcl* on Sxl expression in the PGCs. Consistent with precocious activation of numerator elements, *Sxl* is inappropriately activated in gcl^- PGCs. Reciprocally, ectopic expression of *gcl* in the soma is sufficient to inhibit Sxl protein expression. Supporting the conclusion that Sxl is an important target of *gcl*, simultaneous removal of Sxl and *gcl* mitigates loss of PGCs observed in gcl^- embryos. Conversely, precocious expression of *Sxl* in PGCs results in a modest yet consistent reduction in total number of PGCs in early embryos, and disrupts their migration during mid-embryogenesis. These observations underscore the biological relevance of transcriptional quiescence in the embryonic PGCs and establish Sxl, the master determinant of female somatic fate, as a critical target of silencing mechanisms in PGCs.

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

Increasing CRISPR-Mediated Homologous DNA Recombination Efficiency

Brittany Williams, Danielle Isakov, Siddhartha Jena and Max Wilson, **Jared Toettcher**

CRISPR-Cas9 has emerged as a tool for genome editing capable of inducing modifications in specific genes through either Non-Homologous End joining (NHEJ) or Homologous DNA Recombination (HDR). NHEJ and HDR provide scientific and clinical applications to further test and treat gene specific diseases. However, CRISPR-mediated HDR is limited by low successful insertion rates. Genetic or chemical manipulation of the NHEJ pathway and the cell cycle has been utilized to increase HDR efficiency, but still results in low success. A recent study using the Cas9-Avidin-Biotin ssDNA (CAB) system showed a tenfold increase in HDR efficiency compared to standard procedures. However, the effects of the various HDR pathway parameters on recombination efficiency are still unknown. Using the CAB system, we sought to understand the relationship between HDR efficiency and two important variable parameters of the donor strand: insert length and homology arm (HA) length. We hypothesize a non-linear relationship between CRISPR-mediated HDR efficiency and insert length or size of HA. We built a Cas9 piggyBac transposon plasmid and transfected NIH 3T3, HeLa and 293T cells to create Cas9 expressing cell lines. Hygromycin B drug selection was used to select for Cas9 integration. Successful HDR was ascertained using inserts of different fluorescent proteins. The number of successful insertions were visualized through microscopy. By plotting the efficiency versus the insert length and HA length we expected to find the optimal settings needed for efficient HDR. Finding the optimal donor strand parameters can then lead to a genomic modification strategy that overcomes the uncertainties of HDR and eliminates the need for genomic PCR.

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

The Antimicrobial Role of Candidate Biosynthetic Gene Clusters from *Actinomyces* Strains within the Oral Human Microbiome

Lucy Williamson, Jared Balaich, **Mohamed Donia**

The oral community is an essential part of the human microbiome controlled by both aerotolerant and anaerobic commensal bacteria including the *Streptococci* and *Actinomyces* genera. Recent studies have shown the antimicrobial role of commensals through production of antibiotic metabolites that can inhibit bacterial growth within the human microbiome. These secondary metabolites are produced through biosynthetic pathways regulated by biosynthetic gene clusters (BGCs), which have become a key genomic tool in identifying candidate gene clusters to test for antibiotic bioactivity. Two BGCs – a phenazine-like cluster and a thiopeptide-like cluster – from different *Actinomyces* spp. were identified using an antiSMASH similarity analysis against known homologous antimicrobial gene clusters. We hypothesized that these BGCs each produce secondary metabolites that can inhibit both pathogenic and commensal growth within the oral community. We endogenously cultured BGCs within their *Actinomyces* strains to assess the cell extracts using 1) bioassays against a panel of oral *Streptococci* commensals and indicator pathogens to test for antimicrobial activity, and 2) liquid chromatography-mass spectrometry (LC-MS) to analyze extract composition for phenazine-like and thiopeptide-like molecules. The bioassays have shown potential bioactivity at lower extract concentractions, yet further analysis of extract concentration and composition must be done to definitively identify any antimicrobial molecules. Additionally, a heterologous expression of each candidate BGC in *Streptomyces albus* will measure for antimicrobial bioactivity specifically attributed to these two clusters. This project should ultimately determine the importance of *Actinomyces* abundance within the microbiome to potentially limit both pathogenic and commensal growth within the oral community.

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

RATrak: A Novel Software Suite for the Analysis of *Drosophila melanogaster* Locomotor Data

Scott Wolf, Simon Forsberg, **Joshua Shaevitz and Julien Ayroles**

Drosophila melanogaster locomotor patterns provide keen insights into the molecular and genetic structures that control complex behavioral patterns. Furthermore, many of the movement characteristics and regulators identified in *Drosophila*, including many related to psychostimulants and environmental adaptation, can be translated to mammals. Fortunately, with the advent of modern computer vision applications, recording locomotor activity is a relatively simple endeavor and can be used to generate valuable datasets quickly. Up until now, the packages that provide strong computer vision foundations have lacked user-friendly programmatic interfaces for end users to leverage datasets. In this project, we took on the challenge of providing a generalized R package to facilitate the investigation of these novel datasets. Working specifically with locomotor data collected from high-throughput assays in which flies are tracked in segregated arenas, the RATrak package provides users with a suite of structured analysis functions and visualization tools for probing their datasets. Along with the software suite, we present a newly developed feeding assay to complement the package and allow for simultaneous measurement of locomotion and food consumption. Using these tools in unison, we recapitulated the finding of previous studies and validated our work as a novel resource for gaining further insight into the behavioral repertoire of *Drosophila melanogaster*.

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

Identification of antiviral signaling pathways restricting hepatitis delta virus persistence

Evelyn Wu, Elham Shirvani Dastgerdi, Benjamin Y. Winer, Gabriela Hrebikova, Brigitte Heller, **Alexander Ploss**

Hepatitis delta virus (HDV) is a negative-sense, single-stranded RNA satellite virus that depends on hepatitis B virus (HBV) surface proteins (HBsAg) to form infectious virions. Of the 258 million people chronically infected with HBV, 15-20 million are co-infected with HDV. HDV/HBV co-infection leads to rapid progression of liver disease, often culminating in hepatocellular carcinoma (HCC) and death. HDV can only establish robust infection in humans and chimpanzees, which poses a challenge to studying HDV dynamics and assessing the efficacy of antiviral treatments. A small animal model that accurately recapitulates HDV pathogenesis *in vivo* is thus urgently needed. The discovery of the human sodium co-transporting polypeptide (hNTCP) as the functional receptor for HBV and HDV entry has helped overcome the barrier to entry in non-susceptible cells and has led to the generation of *in vitro* and *in vivo* models. Our preliminary data show that in the absence of persistent HBsAg expression, HDV can persist in human hepatocytes for at least 8 weeks while it is cleared in hNTCP-expressing murine hepatocytes after 4 weeks. This difference in persistence indicates that additional restrictive factors are involved in defining HDV's host tropism. This study seeks to identify murine HDV sensing mechanisms that might contribute to HDV clearance by knocking out cellular nucleic acid sensors using CRISPR/Cas9 technology. sgRNA constructs targeted against the identified nucleic acid sensors have been tested *in vitro* in murine Hepa1.6 cells. Constructs that generate successful knockouts as identified by Western blot analysis will then be delivered to our novel mouse model via adeno-associated virus-mediated transduction, enabling us to determine the importance of different antiviral signaling pathways in restricting HDV replication in rodents.

This research was made possible by the generous support of the Class of 1943 Molecular Biology Senior Thesis Fund & the Crecca '46 Molecular Biology Senior Thesis Fund.

Determining the localization of uncharacterized proteins required for photosynthesis

Yihua Xie**,** Lianyong Wang, Kelly Van Baalen, **Martin Jonikas**

Plants use photosynthesis to produce the food and energy required by life on Earth, and we rely on rice and wheat for consumption especially. But due to the rapid growth in human population, a shortage of agricultural supply is an increasingly growing concern. To obtain higher yields, it is both environmentally friendly and effective to explore methods that can improve the photosynthetic efficiency of plants. To achieve this, various approaches are available for engineering plants. The more promising ones include accelerating the slow recovery in photoprotection, over-expressing enzymes which are involved naturally in the Calvin-Benson-Bassham cycle and attempting to recreate the algal photosynthetic mechanism known as the carbon concentrating mechanism (CCM) into higher crop plants. However, one of the major obstacles faced in these approaches is identifying the specific proteins and enzymes that can improve yield. While many components involved in photosynthesis have been identified and studied intensively, many others remain uncharacterized. This can clearly be seen in the list of photosynthesis-deficient *Chlamydomonas reinhardtii* mutants generated in Jonikas lab using random insertional mutagenesis. The genes responsible for this phenotype are novel and largely uncharacterized. To develop a more in-depth understanding of these genes, this experiment aims to track their protein localizations. We use a high throughput cloning protein and tagging pipeline to fluorescently tag the proteins encoded by the genes. Finally, following up with the fluorescence tagging, the *Chlamydomonas* samples are photographed with confocal microscopy. Having the ability to analyze the functions of each and one of these proteins will dramatically advance the field's understanding of what role each gene plays in photosynthesis, and this localization data is the first and vital step in this process.

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