question

粒粒皆辛苦

on ne voit bien qu'avec le cœur

analysis

Εὕρηκα

the mitochondria is the powerhouse of the cell

data

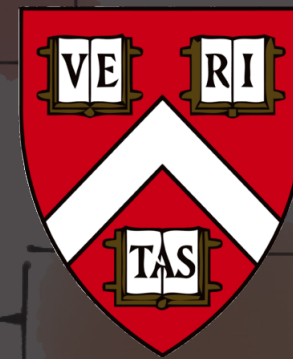big ideas

If people did not love one another, I really don't see what use there would be in having any spring.

experiment

HARVARD COLLEGE   PRISE   BLISS   PRIMO   SHARP   SURGH   ABSTRACT BOOK 2016

VE RI TAS

HARVARD COLLEGE
PBPSS
PRISE | BLISS | PRIMO | SHARP | SURGH

ABSTRACT BOOK 2016

PBPSS 2016
Abstract Book

Table of Contents

Director's Letter

Dear reader,

On behalf of Rakesh Khurana, the Danoff Dean of Harvard College, and Jay Harris, the Dean of Undergraduate Education, I am pleased and privileged to once again introduce the research of the Harvard College Summer Undergraduate Research Village community, which comprises PRISE, BLISS, PRIMO, SHARP, and SURGH. These five programs together form a robust, active, and interdisciplinary residential community of scholars focused on research with Harvard faculty in formative and substantive projects over ten weeks in the summer.

2016 marks the eleventh year since the genesis of PRISE, and the Summer Undergraduate Research Village concept that continues to evolve and provide exceptional opportunities for our fellows to explore intellectual questions of merit (and personal interest!). These questions, to a high degree, inform their undergraduate and postgraduate scholarship trajectories. Beyond research, the engagement of the fellows is evident in the broad, creative (and fun!) range of activities they organize themselves at Leverett House that essentially define each of the program cohorts and our community more broadly.

The abstracts included herein speak for themselves: impressive in their considerable interdisciplinary range, the projects are a testimony to the fellows' diligent effort and compelling sense of purpose. With gratitude to our faculty hosts, the abstracts further underscore the truly inspiring research happening across the Harvard universe.

I wish you—the fellows in PRISE, BLISS, PRIMO, SHARP, and SURGH—the very best of success in your academic endeavors going forward. I suspect the important relationships you have cultivated over the summer will continue long after the fact, and I am deeply gratified to have witnessed your enthusiasm, inclusivity, and heartfelt bonhomie.

Yours truly,

Gregory A. Llacer

Director, *Harvard College Office of Undergraduate Research and Fellowships* (URAF)
Director, *Harvard College Program for Research in Science and Engineering* (PRISE)

Editors' Letter

Dear PBPSS Community,

It is unbelievable how quickly the summer has passed by. In just ten short weeks, we have not only conducted amazing research, but also developed a close-knit community. We've heard from scholars at the forefront of their fields through the Distinguished Speaker Series and Faculty Chats, and witnessed the incredible talent present in our own community through performances at Coffeehouse. We've gotten motion sickness from both the waves of the Atlantic Ocean and the roller coasters at Six Flags. And throughout all of this, we've bonded with each other, whether over the complexities of research, the sweltering heat, or just life in general. These experiences are what makes the research village so unique.

But of course, a fundamental part of the research village is...research. This summer, our research spanned unique disciplines and showcased the diversity of academic interests in the village. From the darkest corners of distant galaxies, to the depths of the self; from communities across the globe, to those right here in Boston, our research has remarkable breadth, and it has been amazing to learn about each other's research while putting together the Abstract Book. Serving on the Editorial Board has been an incredible honor.

It would not have been possible to compile the abstract book without the help of many people. We'd like to thank Felipe Flores for all the time and effort he put into compiling the book, and Gurbani Kaur for the encouraging words and emotional support; furthermore, we owe tremendous thanks to the peer editors, the proctors and PAs, the PBPSS staff, all the research advisors and mentors, and most of all, to the PBPSS fellows. We are thrilled to have an abstract from every single one of you published in this book.

We hope this summer was as wonderful for you as it was for us. Wherever we've gotten to know you, it's been a pleasure, and we wish you the best in your future endeavors.

Sincerely,

PBPSS 2016 Editorial Board

Editors-in-Chief:
Audrey Effenberger, Leah Marsh, Leah Rosen, Soumyaa Mazumder

Editorial Board:
Audrey Effenberger, Elgin Korkmazhan Gulpinar, Ellen Zhang, Leah Marsh, Leah Rosen, Miruna Cristuş, Soumyaa Mazumder, Suproteem Sarkar, Yong (Daniel) Shen

Peer Editors:
Adam Frim, Alexander Jin, Andre Sanchez, Audrey Effenberger, Carl Denton, Christopher Li, Daniel Um, Elgin Korkmazhan Gulpinar, Ellen Zhang, Felipe Flores, Gerardo Castillo, Gui Zhen Chen, Holly Breuer, Jacob Scherba, Jason Li, Jessica Kim, Jiafeng Chen, Julianna Lee, Kristen Fang, Krystal Phu, Kushi Mallikarjun, Leah Marsh, Leah Rosen, Lingjun Xia, Madeleine Granovetter, Mara Coyan, Margaret Panetta, Michael Giles, Michael Liu, Miruna Cristuş, Quinn Sluzenski, Rachel Oshiro, Sarah Rahman, Soumyaa Mazumder, Suproteem Sarkar, Theresa Clark, Thomas Culp, Yong (Daniel) Shen

Design Ninja:
Felipe Flores

Design Team:
Audrey Effenberger, Soumyaa Mazumder, Yong (Daniel) Shen

Cover Design:
Audrey Effenberger

Program Overview

The Harvard Summer Undergraduate Research Village was initiated in response to a recommendation of the 2005 Harvard University Task Force on Women in Science and Engineering (WISE). The first program, initiated in the summer of 2006, was the Harvard College Program in Science and Engineering (PRISE). It was charged with developing a diverse, residential community of Harvard College undergraduate scientists conducting research with faculty over ten weeks during the summer.

Five years later, at the recommendation of the Dean of Social Science in the Faculty of Arts and Sciences and the Dean of Harvard College, PRISE was joined by the Behavioral Laboratory in the Social Sciences (BLISS), a joint project of Harvard College and the FAS Division of Social Sciences. The Program for Research in Markets and Organizations (PRIMO), a collaboration of the College and the Harvard Business School, was also founded in 2011.

2013 marked the first year of the Summer Humanities and Arts Undergraduate Research Program (SHARP), launched with the support of the Dean of Arts and Humanities in the Faculty of Arts and Sciences and the Dean of Harvard College. Most recently, in 2015, SURGH (Summer Undergraduate Research program in Global Health) joined the Research Village, filling out the complete interdisciplinary structure of the Summer Undergraduate Research Village.

Today, PRISE, BLISS, PRIMO, SHARP, and SURGH form a vibrant community of student researchers engaged in the pursuit of knowledge. We are proud to present their outstanding work in this abstract book.

PRISE

The Program for Research in Science and Engineering aims to build community and stimulate creativity among Harvard undergraduate researchers in the life, physical/natural, engineering and applied sciences. Fellows possess both a strong dedication to academic interest and excellence in scientific research and an enthusiasm for participating in a diverse residential community of scholars. PRISE fellows independently coordinate their research positions and work on projects with Harvard-affiliated researchers.

BLISS

The Harvard College Behavioral Laboratory in the Social Sciences aims to provide a formative and substantive research experience, to build community, and to stimulate creativity among a small cohort of Harvard undergraduate researchers in the social sciences. BLISS fellows work on projects designed by Harvard-affiliated faculty and researchers.

PRIMO

The Program for Research in Markets and Organizations aims to build community and stimulate creativity among Harvard undergraduate researchers in business and related fields. PRIMO fellows are selected to work on one of the pre-designated PRIMO research projects which span diverse topics (e.g. finance, organizational behavior, marketing), disciplines (psychology, economics, sociology), and methods (quantitative and qualitative). Fellows work on projects with faculty at Harvard Business School.

SHARP

The Summer Humanities and Arts Research Program aims to build community and stimulate creativity among a small cohort of Harvard undergraduate researchers in the humanities and arts, providing students with diverse research opportunities in an exciting range of research settings. SHARP fellows contribute to the rich, interdisciplinary intellectual, social, and residential environment in activities that include roundtable lunch talks with distinguished faculty speakers, pre-professional seminars, and opportunities to explore Harvard and the New England region more broadly. SHARP fellows work on research projects with Harvard-affiliated faculty and researchers, and library and museum administrators.

SURGH

The Summer Undergraduate Research in Global Health program aims to build community and stimulate creativity among a small cohort of Harvard undergraduate researchers in global health, and offers Harvard undergraduates the opportunity to research critical issues in global health under the direction of a Harvard faculty or affiliate mentor. SURGH fellows work on research projects with Harvard-affiliated faculty and researchers.

Adapted from the URAF website.

BLISS

BLISS

Intervention to Turn Social Networks into Supports for Student Success

Anna Zannetos
Sociology, 2018

Harvard Kennedy School of Government

Advisor: Todd Rogers, PhD

Mentor: Kim Bohling, EdM

Many students have strong relationships with supportive adults, including mentors, family friends, and extended family members. Much research indicates that strong relationships with non-parental adults are associated with greater success for students in school and other domains. However, the Family Educational Rights and Privacy Act (FERPA) prohibits schools from communicating with these adults about students' performance in school. The My Student's Team pilot project is a randomized controlled trial of a low-cost, scalable approach to inviting and empowering non-parental adults in high school and middle school students' lives to support their success in school. This project is grounded in the "nudge" approach, an emerging branch of psychology and economics that translates the science of human behavior into simple strategies for motivating specific targeted behaviors.

One hundred ninety-one parents and guardians of students attending school in mid-sized city outside of Boston identified a non-parental adult who cares about their child and authorized the district to communicate directly with this named "supporter." These supporters then received communications about the student's educational activities, including attendance and grades, during the 2015–2016 school year. Supporters also received actionable information about how to support and engage with the student and his or her school performance. Those in the randomized treatment group received more frequent, actionable, and personalized communications than those in the control group.

This intervention investigated whether communicating with non-parental adults improved students' educational success. It was hypothesized that students in the treatment group will see greater improvement in grades and attendance than students in the control group. It also investigated whether and how communications from the school affected non-parental adults' perceived responsibility, permission, and ability to provide educational support to the student.

I conducted in-person and phone interviews with students, parents, and supporters in both the control and treatment groups at the end of the 2015–2016 school year. Participants were asked questions about their relationships and interactions with each other and the perceived impact of the program. These interviews gave insight into the mechanisms driving the intervention's impact. Data from the interviews supplemented quantitative data about student performance. Data were used to improve the pilot program in preparation for implementation in approximately four school districts in the 2016–2017 school year.

New Dynamics in Urban Politics

Anthony Chen
Statistics, 2018

Harvard University

Advisor: Ryan Enos, PhD

Mentor: Peter Bucchianeri

The skylines of our country's largest cities have grown and changed tremendously. However, the effects of an increasingly urbanized and concentrated population on government and democratic life are less obvious. Professor Ryan Enos's team is studying how the cityscape is affecting politics and policies, examining a wide range of data, from commuter/traffic statistics to mayoral speeches. My main role this summer has been devising data collection strategies, as well as conducting secondary research.

One of the largest projects I have undertaken so far is parsing city council meeting minutes to pull out individual bills and their corresponding vote tallies. We are looking to use this data to create adjacency matrices, which are essentially graphs that visually show the political relationships and "networks" of city councilmembers, based on their voting history. Similar studies, called network theory studies, have been conducted on a national level with Congress and have been very revealing, showing increased party polarization over the past few decades. Network theory has yet to be applied to local government, which is most likely due to the fact that city council voting records are not nearly as organized as Congressional voting records. For the actual data parsing, I used the statistical programming language R, utilizing a wide variety of new functions and packages including regular expressions to create "wildcard" search patterns.

A second project looked into the State of the City

speeches of mayors around the country. State of the City speeches, similar to the State of the Union speech given by the President, serve as important updates to citizens about the legislative on-goings and plans of the government. I wrote an R script and the Selenium module to crawl through Youtube and pull captions from the website to create a transcript of the speeches. Selenium allowed me to automate web browsing and programmatically click links, making the daunting task of downloading hundreds of speeches into a simple program. We can use these transcripts of mayoral speeches to create visual displays based on word frequency and analyze the vocabulary used. In addition to these two larger projects, I created a variety of datasets on topics ranging from racial demographics of elected officials to inter-state migration tables. While many of these projects are very much in their preliminary stages, the exciting work done here is sure to reveal new discoveries about the impacts of urban life on government.

Partner Choice and the Pedagogical Function of Punishment

Chace Shaw
Psychology, 2017

Harvard University

Advisor: Fiery Cushman, PhD

Mentor: Justin Martin

When a partner defects on us, whether in business or personal life, we can respond in a number of ways. One common response is to engage in some form of punishment. In research focusing on cooperative behavior, punishment falls under the broader heading of partner control. In partner control, individuals are motivated to change the behavior of their partner. This is accomplished through punishing the offending partner for transgressing. Evolution favors the tendency to punish offenders, because this ensures more beneficial relationships with others. In this way, the threat of punishment serves as a deterrent of future uncooperative behavior. One prediction of this hypothesis is that successful punishment will change the offender's attitude and behavior and, fulfilling its pedagogical function, will leave the victim satisfied with the interaction. Thus, after punishing the offender, the victim will be sensitive to indications of the successfulness of the punishment, such as post-punishment feedback from the offender. Previous research has established that victims are most satisfied after punishment when the offender gives feedback to the victim, indicating that he or she understood that the victim was intending to punish and expressing a moral change of attitude and future behavior.

However, punishment is not the only option we have when responding to bad behavior. An alternative to punishing an uncooperative partner is deciding not to interact with the offending partner again. In the literature, this is called partner choice. Characterized by voluntary cooperative relationships with alternative partners readily available, partner choice differs from partner control in that one's response to defection is motivated by the ability to leave the current partner and choose new partners. As such, the goal of partner choice is not to change a partner's behavior but rather to predict whether that partner will be cooperative in the future, and if not, to end this relationship and start a new one. Thus, those engaging in partner choice should be sensitive to different signals compared to those engaging in partner control. While those who engage in partner control should care about whether punishment leads to a change in their partner's behavior, those engaging in partner choice should be indifferent to whether their previous partner's behavior was changed as a result of ostracizing them, and if anything, may be less satisfied when ostracism causes someone to change (because of missing the opportunity to interact with a now-cooperative partner).

In an online experiment, two participants will play a game in which one decides how to divide a sum of money between the two, while the other chooses how to respond. In the partner control condition, the player will respond by deducting money from the first player, while in the partner choice condition, the player will respond by choosing whether to continue playing with the same player or to leave and play with a new one. The responder will then report his or her satisfaction with the interaction after receiving a message from the other player indicating either a change in behavior or no change in behavior.

In this study, we aim to further highlight the unique pedagogical function of punishment by contrasting punishment decisions with partner choice decisions. Following the evolutionary functions of these behaviors, we focus on their differential sensitivity to feedback from individuals who have either been punished or ostracized. Specifically, we predict that post-ostracism feedback indicating a moral in behavior and attitude will actually result in a reduced level of victim satisfaction in the partner choice condition, but that post-punishment feedback indicating a change will result in higher levels of victim satisfaction in the partner control condition.

Suicide Risk Factors and the RDoC Matrix

Charlotte Anrig
English &
Psychology, 2018

Harvard University

Advisor: Matthew K. Nock, PhD

Mentor: Heather S. Pixley, MBA

This summer, I have been working on a study that examines past research on suicidality in the context of the Research Domain Criteria. The matrix, commonly known as RDoC, represents an attempt to reform the present classification system for mental illnesses: the Diagnostic and Statistical Manual (DSM) for Mental Disorders, published by the American Psychiatric Association. The DSM treats mental disorders like somatic disorders, framing each as a distinct, consistent entity that can be easily identified through observable signs and symptoms—and thus each patient can be treated the same way as everyone else who has ever exhibited the same behaviors.

In reality, though, brains are complex systems, and psychological disorders tend to be incredibly heterogeneous. Two people demonstrating roughly similar symptoms may actually have completely different neurological problems, and they might need entirely different treatments as a result. The RDoC attempts to account for this complexity by rejecting the idea of discrete disorders altogether, opting instead to arrange a matrix of mental functioning as a whole. One axis represents different domains of human psychological ability while the other axis identifies levels of biological functioning from genes to self-report, each point on the matrix representing a different, highly specific piece of mental functionality that could potentially break down. This structure allows for more specific diagnoses, and its focus on underlying biology rather than subjective observation creates a pathway towards more targeted and empirically-supported treatments in the future.

The Nock lab aims to identify RDoC elements that act as risk factors for suicidal behavior. By examining, translating, and organizing virtually all relevant research, we hope to produce a comprehensive database and data visualization tool to examine suicide risk within the RDoC framework. Thus, the research depends on two basic sets of questions: firstly, how can existing research about suicide be adapted for the RDoC? Secondly, what are the most robustly justified risk factors for suicide, and what are the gaps in the existing research? The answers to these questions should have significant implications for large conceptual questions about classification in the field as well as important clinical utility.

To create the database, a team of abstract coders select from a broad literature articles that seem to handle RDoC elements and outcomes related to self-injurious thoughts and behaviors (SITBs). After being identified and classified into an RDoC category (genes, self-report, etc.), the articles are sent to the article coders. These coders—myself and two other interns, as well as several other coders at other sites—read, analyze, and input the papers into Qualtrics, a survey platform that has been adapted to use a series of predefined questions to allow for consistent data entry for this project; as a team, we decide which findings are relevant and figure out how to translate them for the RDoC system. So far, we have been able to successfully build the beginnings of a database. Data entry will hopefully be finished in six months, with some sections potentially completed by the end of the summer.

The Effect of Emotions in Anti-Smoking Public Service Announcements

Gloria Yu
History and Science, 2018

Harvard Kennedy School of Government

Advisor: Jennifer Lerner, PhD

Mentor: Joowon "Jo" Kim, MA

Anti-smoking public service announcements (PSAs) often use a variety of methods to discourage tobacco use. Some outline the benefits of quitting, others illustrate the horrendous health ramifications of smoking, and still others provide tear-jerking narratives of the impact of smoking on loved ones. Many of them also rely heavily on emotion, both positive and negative, to motivate changes in smoking behavior. Recent research by Dr. Lerner, however, indicates that the use of certain emotions in PSAs could backfire. Based on the Appraisal Tendency Framework (ATF), a multidimensional theoretical framework for linking specific emotions to specific judgment and decision-making outcomes, certain emotions lead to impatience and a higher desire to attain products. Hence, evoking certain emotions in PSAs could ironically lead to increased smoking behavior.

Dr. Lerner's research combines theories of public health, psychology, behavioral economics, and psychoneuroendocrinology to examine the mechanisms behind emotionally-evocative PSAs and their effectiveness. We are in the beginning stages of this tentative five-year research project. This summer, we have been working on analyzing survey and written data from the pilot study on anti-smoking PSAs

and how effectively they elicit various emotions. Using its findings we can fine-tune target stimuli and test parameters for the second round of studies. Ultimately, this program of research will provide theoretical advances in understanding the effects of emotion, and practical advances in designing effective PSAs.

Matching Market Algorithms

Hansenard Piou
Applied
Mathematics, 2018

Harvard Kennedy
School of
Government

Advisor: Maciej
Kotowski, PhD

As opposed to usual markets where consumers have no preferences over what products they consume—there is no egregious difference between two apples, for example—matching markets contain individuals who have preferences of what they consume. Moreover, while the former markets can be cleared by just having all participants buy and sell products at the optimal price, matching markets can only be cleared if certain conditions are met, namely if no participants can create a better matching for themselves than the one they were given; e.g., a set of men each with preferences over a set of women whom they want to marry and vice versa.

During the early stages of my research, I learned about and coded in Python the two archetypal algorithms to clear a matching market and result in a stable matching: the deferred acceptance (DA) and top trading cycles (TTC) algorithms. The first algorithm is for instances when two groups of participants have preferences on who they want to be matched with. For this algorithm, members of one group "propose" to their top choice of those in the other group. Each "acceptor" then rejects all except their top choice of those who proposed to them, although they have the option of rejecting all of them and remaining solo. All of the rejected proposers then propose their next top choice including themselves, repeating the process until all proposers have been matched or have decided to go solo.

The TTC algorithm is used when participants want to trade unique items amongst themselves. Each "agent" points to their top "house" of choice and each "house" to its owner. At least one cycle will form where a "preference" chain will begin and end with the same agent. We clear all cycles by giving each agent their top pick, and the remaining agents point to their top remaining choice. We repeat this process and keep clearing cycles until there are no more agents in the market.

While versions of these algorithms see plenty of use matching organ donors and patients, college applicants and colleges, and hospitals and medical school students, our research has focused on the aforementioned condition of stability, that no coalition of participants could create a matching that would be more preferable for at least one of the members. Maciej Kotowski and Ivan Balbuzanov proposed a stricter definition of stability, where no coalition could strongarm other participants into giving its members more preferred "houses" by threatening not to sell others their houses.

In order to find properties of matchings under this definition, I coded a modified TTC algorithm which creates a larger set of matchings than the original TTC and coded functions which check whether a matching satisfies this strongarm condition. While we are very much working in the realm of theory, determining the properties of stability from exclusion and algorithms which produce such stable matching will give future market designers the tools to create more robust markets.

Memory and Imagination: Using the Episodic Specificity Induction

Harrison Tanzola

Psychology, 2019

Center for Brain
Science, Harvard
University

Advisor: Daniel
Schacter, PhD

Mentors: Kevin
Madore; Preston
Thakral, PhD

This study looks to demonstrate that an episodic specificity induction can enhance imagination of future events. Prior studies have shown that episodic memory plays an important role in imagining the future. This act of imagining future events, also known as episodic simulation, is theorized to be a result of a process by which past memories are recombined in unique ways to produce an entirely new event.

An episodic specificity induction is a task that has been demonstrated to enhance episodic memory by increasing internal details, details pertinent to an event that is being remembered. Even more striking is the fact that the induction does not increase external details, merely semantic or metacognitive statements; rather, it singularly affects internal details, the details coming from episodic memory. And because episodic memory is said to have a much larger contribution to imagining the future than is semantic memory, this study theorizes that having subjects perform an episodic specificity induction as opposed to a control induction will result in en-

hanced details in imagined events.

For this study we used a within-subject control, so each subject received both an episodic specificity induction and a control induction, followed by math problems to serve as a control, and then a set of five cues for each. The subjects were asked to base a simulated future event on these cues. Both the inductions and the cues were counterbalanced between subjects. Furthermore, all responses to cues were self-timed and included a standardized set of general probing to get as much detail as possible from the subject, without having the experimenter implant details that were not there.

After all subjects have been run, and their interviews transcribed, transcripts of each subject will be scored using a standard autobiographical interview (AI) scoring protocol. If the hypothesis is correct, the AI scores will reveal a statistically significant increase in the average number of internal details in cue response following the episodic specificity induction, as opposed to those following the control induction. Also, there should be little to no change in the number of external details between inductions.

Moral Judgments and the Effortful Nature of Forgiving Accidental Harm

Irene Limb
Psychology, 2018

Harvard University

Advisor: Fiery Cushman, PhD

Mentor: Justin Martin

In cases of moral judgment, people largely care about two factors: who caused the harm and whether the harm was intentional. However, we care about these factors to different extents depending on the type of judgment we make. Decisions to *punish* the harm-doer depend upon both the outcome and the intent of the harm-doer, whereas judgments of *wrongness* only depend on the intent.

While Cushman (2008) articulated this idea at a coarse level by creating a two-process model that draws connections between the inputs—causation and intention—and the judgments they influence, recent studies have hinted at further delineations that may be involved in our moral judgments. Specifically, it seems that people can seamlessly incorporate causation into their moral judgments, but taking intentions into consideration seems to require additional cognitive effort (Buon et al., 2013). That is, under normal circumstances, we judge someone with bad intent much more harshly than someone with neutral intent, and we forgive the person who

accidentally causes harm and judge them nearly identically to a person who caused no harm. However, when our cognitive capacities are occupied, we have difficulty incorporating information about a harm-doer's mental state into moral judgment. We fail to inculpate agents with bad intent and fail to exculpate agents with good intent, judging them only on their causal connection to harm.

These results have obvious implications for Cushman's two-process model: they suggest that there is some step beyond the mere *recognition* of a harm-doer's causal role and intent, and that this additional step requires cognitive effort; what Cushman's model lacks is information on the point at which this additional step occurs in our cognitive processes.

There are two ways we can interpret this. The first is that under cognitive load—when our mental resources are taxed—people are sensitive to intentions but cannot seem to incorporate them into judgments of preference and wrongness. That is, participants are still aware that the accidental agent did not intend harm and that the intentional agent did, but somehow, because their cognitive capacities are occupied, they fail to include this information into the final judgments made. The second interpretation is that outcomes *do* have an influence on judgments of wrongness, and normally, we are capable of inhibiting this influence. But under cognitive load, we are unable to inhibit this influence and we judge perpetrators of negative outcomes to be morally wrong. The reason why previous results had shown an influence of cognitive load on intent-based judgment is not because one's ability to incorporate intentions was impaired, but because one's ability to inhibit the influence of outcomes was impaired.

To test this, we are in the process of running experiments aimed to distinguish people's judgments of punishment from their judgments of wrongness. Participants in our study will watch several video clips showing a moral agent either accidentally or intentionally harming a second agent, either under cognitive load through a verbal shadowing task, or not under load. We then ask participants such questions as, "Whose behavior was wrong?" and "Should he be punished?" in order to explicitly compare judgments of wrongness and punishment. The pilot data has supported the second interpretation, that wrongness judgments occur because people have difficulty inhibiting them when they observe a harmful action. We predict that we will continue to find a significant impact of cognitive load on judgments of wrongness, such that participants will be unable to inhibit the influence of a negative outcome on their judgments of how wrongly an accidental agent behaved.

The Use of Historical Analogues in Foreign Policy Decisions

Krysianna Papadakis
Social Studies, 2017

Harvard Kennedy School of Government

Advisor: Graham Allison, PhD

Mentors: Adam Siegel, MA; John Masco

This question is part of the Applied History project, led by Graham Allison, Douglas Dillon Professor of Government at Kennedy School of Government. Diplomats and public officials frequently make use of historical precedents in order to evaluate possible courses of action, especially in cases where there is little information about a current problem. However, the use of historical analogies can be dangerous: it is often motivated by emotional responses to past experiences and tends to make people blind to the facts that distinguish present cases of conflict from those in the past. Therefore, the Applied History project is partly aimed at encouraging a systematic approach to the use of historical analogues in present decision-making. My role in this project is mainly to provide a broad literature review of scholarship on the uses of historical examples in (primarily) 20th century policy in international diplomacy.

This project began with reading the biographies of major statesmen like John F. Kennedy, whose biography was written by Kennedy's "resident historian" at the White House, Arthur M. Schlesinger. Historians like Schlesinger are frequently present in the rooms where foreign policy decisions are made, such at the planning of the Bay of Pigs invasion and the nuclear escalations of the Cold War. Many of these historians then write retrospective analyses about how historical analogies influenced the decisions in these cases. My role was to find and summarize the work of such historians on how history has been used in the past and how it can or should be used in the future. My specific focus within this literature review is how the concept of "historical inevitability" is used, exploring how it often conceals implicit biases about what historical events are considered "good" or not, building off the work of E.H. Carr in his response to Sir Isaiah Berlin.

Finally, I have been assisting in the review of the appendix of Professor Allison's forthcoming book, *Destined for War: America, China, and Thucydides's Trap.* Specifically, I fact-checked the sixteen cases of stand-offs between rising and ruling powers in history, called the "Thucydides Trap Case File."

Third Party Intervention in Children

Liana Yamin
Human Evolutionary Biology, 2017

Harvard Lab for Developmental Studies

Advisor: Felix Warneken, PhD

Mentor: Young-eun Lee

Past research has demonstrated that uninvolved third parties who witness an unfair act are willing to intervene to restore justice, even when doing so is personally costly. Restoration of justice can generally occur in two ways: punishing the perpetrator or compensating the victim. It has been shown that adults reward third-party compensators more favorably than third-party punishers in unfair distribution scenarios, but it is unknown whether young children share this preference for compensators.

This summer, I assisted with a study that evaluated children's perceptions of third party compensation versus third party punishment. We presented children ages five to nine with a story and visual aids describing a group of children at a summer camp. One child, the "decider," does not share candy with a peer. Two "watchers" witness this, and one decides to take away candy from the decider while the other chooses to give candy to the peer who received none. We asked participants questions about each bystander to determine preference and overall perception. Then, we provided them with an opportunity to punish, compensate, or do nothing in response to the same unfair division of candy that they saw in the story. Across all ages tested, children preferred compensators over punishers. When given a chance to intervene, a significant majority chose to compensate at a cost to themselves, rather than punish. These findings illuminate that a preference for promoting justice through compensation emerges during childhood.

Political Implications of Class-in-Race Inequality

Mara Roth
Social Studies, 2019

Center for Government and International Studies

Advisor: Jennifer Hochschild, PhD

Recent years have brought increasing class disparities within racial groups. These disparities are particularly interesting to study, as from them arises the conflict between maintaining group solidarity and pursuing individual interests. Because a person's socio-economic status tends to determine where they live, the quality of their life, and their political beliefs and preferences, this increasing class-in-race inequality has the potential to create significant political divisions within racial groups. So, when does an individual choose his or her own interests over standing in solidarity with their racial group, and why?

This research project attempts to resolve these questions by investigating policy disputes in four different U.S. metro-regions. The examined policies and cities include pension reform in Chicago, educational policy in Los Angeles, re-gentrification in Atlanta, and policing in New York City. Because all of these policy disputes have both class and race dimensions—some clearer than others—they provide the perfect window through which to examine class-in-race dynamics.

The research methods for this large scale project include reading news reports, analyzing legal proceedings and census data, contacting subject experts, and traveling to the cities to conduct event/site visits and interviews with union members, political figures, academic experts, and community organizations. The project is ongoing and this information is being used to analyze and explain patterns in the behaviors of African American, Hispanic, White, and Asian American people belonging to different socio-economic classes.

While the current evidence seems to point to the idea that the well-off are more likely to abandon racial solidarity in favor of pursuing self-interest, more research remains to be conducted to further expand upon the complexities of class-in-race dynamics. The results of this investigation and further inquiry into class tensions within racial groups are important for understanding inclusivity in America, and the future of economically disadvantaged Americans of all races.

Emotions and Morals

Quinn Sluzenski
Sociology, 2018

Harvard Kennedy School of Government

Advisor: Jennifer Lerner, PhD

Mentor: Charles Dorison

There is a trolley hurtling down the tracks, brakes broken and unable to stop, about to smash into five workmen who are stuck on the rails. You are standing next to the track and see a switch—if you hit it, the trolley will divert to the side track, saving the five workmen. . . and killing the one workman on the other track. What do you do?

Most people choose to pull the switch in this classic philosophical question. However, when you change the scenario to you standing on a footbridge above the track, when the only way to save the five workmen is to push a man over the side of the bridge to collide with the trolley, most people say that they wouldn't push him.

What would you do? Would you choose to be utilitarian—calculating the greatest net value and choosing five lives over one—or deontological—valuing the inherent ethical obligation to not cause direct harm? It may be that your emotional state will impact your answer. Psychologists have argued that emotion stops you from saving five lives in the face of one, whereas deliberative reasoning leads you to save the five and act in a utilitarian manner.

That research has been extended to show that inducing positive emotions makes people more likely to be utilitarian, while inducing negative emotions makes people more likely to be deontological. However, as Jennifer Lerner's lab has shown again and again, different emotions have differing psychological effects, even if both are "negative" or "positive." This summer, we studied how sadness and anger can affect our moral decision-making in scenarios like the trolley dilemma, in which we must choose whether to cause harm or let greater harm occur. In order to study this question, we are surveying a wide variety of people, inducing sadness or anger, and asking them to judge the morality of a variety of these kind of dilemmas. We analyze their responses and their thought processes to build an idea of how people respond to these kind of moral situations.

Ultimately, we hope to determine what effect that sadness and anger have on these decisions, potentially presenting evidence that sadness and anger have divergent effects. This exploration of morality

can help us understand how and why we make decisions, and could someday even have implications for public policy—as self-driving cars take to the streets, the trolley dilemma transforms from a thought experiment to a real programming question—should we program utilitarian cars, or cars that protect their drivers? What is the appropriate trade-off? This study hopes to further the field of emotions in moral decision-making. When we better understand how we make decisions, we can make better decisions.

Political Legitimacy and Nation Building

Sebastian Reyes
Social Studies, 2019

Edmond J. Safra
Center for Ethics

Advisor: Danielle
Allen, PhD

Mentors: Jess Miner,
PhD; Tomer Perry

The Edmond J. Safra Center, currently led by Professor Danielle Allen of the Government Department, is one of the world's preeminent centers for research on ethics and its practical applications. The center brings together undergraduates, graduates, postgraduates, and others in academia in a unique collaborative setting.

At the time of its founding in 1986, then known as the Program in Ethics and the Professions, academic inquiry into practical ethics was relatively rare. Since that time, the center has become world-renowned, and those associated with it are among the top scholars in a variety of disciplines, ranging from political theory to bioethics.

As the 30th anniversary of the Center nears, part of my work this summer involved going through the center's extensive archives, composed of hundreds of hours of public lectures, symposia, seminars, conferences, and similar events recorded on cassette tapes, CDs, VHS, DVD, and online. My first task was to develop a cataloguing system for all these materials. Subsequently, I delved into the lectures by taking notes, writing down key quotes, and identifying "gems," lectures that stood apart as exceptional in some regard.

After this initial phase was completed, I began work on a research paper based on some of the material I discovered in the archive. In particular, I am focusing on a lecture given in 2005 by Professor Noah Feldman of Harvard Law School entitled "Iraq and the Ethics of Nation Building." I am exploring the concept of political legitimacy, both how Professor Feldman understands it and how it functions in nation-building more generally. In addition, I will be putting Professor Feldman's lecture and his related book *What We Owe Iraq: War and the Ethics of Nation Building* in conversation with other academic texts of a similar subject matter.

Subsequent to my work on this research paper, I am producing five multimedia blog posts on a variety of topics on which the center has focused throughout the years: equality, governmental corruption, ethics and the medical profession, critiques of capitalism, and diversity.

Representations of Personal Identity in the Political Space

Susan Wang
Social Studies, 2017

Edmond J. Safra
Center for Ethics

Advisor: Danielle
Allen, PhD

Mentors: Jess Miner,
PhD; Tomer Perry

The Edmond J. Safra Center focuses on teaching and researching ethical issues as they relate to a variety of disciplines ranging from bioethics to the theory of ethics. As research fellows with the Center, our job has two components— archival work and independently pursued research projects.

The Center is approaching its 30th anniversary, with this year's focus being "Diversity, Justice, and Democracy." As part of the archival portion of our research, our research team is looking at the past thirty years of recorded events hosted by the Center to record highlights from past lectures that can be used for anniversary celebrations, and to create a usable research database for future work. This involves listening to and sorting through all past lectures, conferences, and symposiums to pick out particularly interesting excerpts. These recordings are then marked with tags by topic, speaker, and date, and entered into a catalogue so future researchers can access the Center's materials. Highlights include talks from Amartya Sen on identity, Richard Rorty on justice, Henry Louis Gates, Jr. on ethnicity, and Frances Kamm on high and low theory, to name a few.

We are also working on independent research projects based in the Center's archival materials. I will be compiling a data visualization of the Center's work over its history, looking at metrics such as demographics of speakers, topics of focus, and change in language use, among other things, and compiling a comprehensive quantitative look at the Center's past work. The aim of this project is to produce an easily digestible quantitative overview of the nature of the Center's work and focus, and how that has changed

over time.

I am also working on a separate, longer research paper that examines representations of identity within the political sphere, and questions of how notions of personal identity translate into political stances in developed, well-functioning, liberal democracies. Using Amartya Sen's lecture on identity as a starting point, I examine the transition of personal identity into political identity. In particular, how do aspects of our personal identity—such as ethnic background, religion, gender—translate into coherent political positions? When those identity-based political positions oppose each other, how can we successfully resolve conflict and reach compromise? The aim of this project is to propose a way to respect plurality in politics while still being able to reach effective consensus.

Thus, our summer research aims at helping the Center effectively catalogue its variety of resources for future research, as well as developing specific projects that fit within the Center's 30th anniversary theme.

Domestic and International Affinities of American Elites

Will Strang
History, 2019

Harvard Kennedy School of Government

Advisor: John Donahue, PhD

The United States has long given rise to some form of elite upper class, whether this group be defined by their wealth, education, or occupational prestige. Many of these individuals devote much of their energy to improving conditions for others who are less fortunate through public service, philanthropy, and other means. With the growth of globalization and the increasing interdependence of the United States with other nations, many American elites, especially the younger cohort, seem to have eschewed nationalist principles for a more cosmopolitan ideal. Even as these Americans seem to remain roughly similar in terms of their desire to improve the lives of others, they now seem to be broadening their ambitions internationally. Anecdotal evidence seems to bear out this hypothesis, as it seems that more students at top institutions such as Harvard are going into occupational fields with a more global focus. If this phenomenon can be observed on a broad scale, it will have far-reaching implications for the future development of American society. Shifts in opinion amongst political decision-makers may result in a more expansive and inclusive role for the United States in international issues at the expense of the American middle class. Similarly,

many who relied on support from the upper class to stimulate the local economy may find their prospects adversely affected as funding is diverted overseas.

This project has attempted to provide empirical evidence to support this claim. One aspect of the project I have worked on looked at the public service work of American Rhodes Scholars, who have long been held as symbols of the elite of the United States' education system. Using LinkedIn, a prior registry of Rhodes Scholars, and other sources, two distinct decades' worth of scholars, 320 each from the 1950s and the 1980s, were examined to compare how many years of their careers were devoted to public service, non-profit and charitable organizations, residence abroad, and a variety of other categorizations. Another of my studies examined the philanthropic tendencies of the largest American non-profit foundations to observe how charitable giving's focus has evolved domestically and internationally over time. The fifteen largest organizations by asset size in 1969 and 2014 had their grants from those years broken down by international or domestic focus, as well as through their specific fields of impact. Given the longstanding tradition of wealthy Americans giving back to the community through private foundations embodied by men such as Andrew Carnegie and John Rockefeller, it will be interesting to see if this tendency has continued or whether global initiatives spearheaded by organizations such as the Bill and Melinda Gates Foundation have drowned them out. More analysis in other topics, potentially including study abroad, political opinions, and business practices, should help to continue to clarify the trend towards cosmopolitanism amongst American elites if a trend even exists at all.
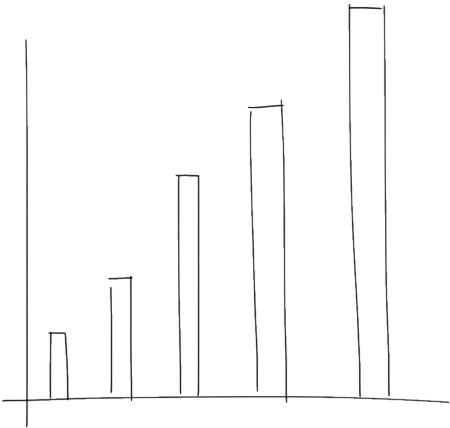
PRIMO

An Exploratory Study of Retirement

Alexander Rohe
Psychology, 2017

Harvard Business School

Advisor: Teresa M. Amabile, PhD

Mentors: Jeff Steiner; Andrew Brodsky

According to the Pew Research Center, approximately 10,000 baby boomers will turn 65 every day until the year 2030. In other words, large numbers of people are reaching the age at which people traditionally retire in the United States. However, the existing literature regarding people's retirement experiences is surprisingly sparse; researchers have not yet deeply examined the ways in which various factors affect people's retirement decisions and retirement transitions. In order to address this gap in the literature and to gain new insights into the nature of retirement, Professor Teresa Amabile and her colleagues in the Retirement Research Group (RRG) are conducting an exploratory study in which they are collecting both survey and interview data from participants at several companies. Participant groups include millennials, employees approaching retirement, those going through the transition, and people who retired from the company in the past five years.

This summer, I assisted Professor Amabile by analyzing qualitative data and by working with Research Associate Jeff Steiner to develop preliminary research presentations and reports to participating companies and individuals. I also assisted doctoral candidate Andrew Brodsky, who is collaborating with Professor Amabile, with his projects regarding online communications and idle time at work.

The Resurgence of Independent Bookstores

Angela Yang
Economics, 2018

Harvard Business School

Advisor: Ryan Raffaelli, PhD

Mentor: Kathy Qu, EdM

The independent bookstore has met with many threats to its survival in recent years. The rise of major chain bookstores like Barnes and Noble and Borders, the founding and rapid growth of Amazon, and the advent of the e-book have all led to many analysts' prediction that the independent bookstore would soon become obsolete. However, in recent years, the number of independent bookstores has increased substantially, despite all reasons to expect the contrary. This unexpected recovery gives rise to the questions of how independent bookstores can manage to survive, or even thrive, in the midst of these unfavorable conditions, and what factors contribute to their resurgence.

This summer, I have been participating in mixed methods research with Harvard Business School professor Ryan Raffaelli to examine the resurgence of independent bookstores and, more broadly, to explore how organizations and industries redefine themselves in response to technological disruptions. The qualitative component of the research includes literature review and rigorous analysis of articles, interviews with bookstore owners, notes, and other sources. The quantitative component involves data collection and analysis of macroeconomic trends in the past few decades, such as the performance of competitors like Barnes and Noble and Amazon. Similar phenomena have been observed in such industries as the watch industry, where Swiss watchmakers were forced to adapt to a technology shock in the form of the quartz watch. More recently, similar trends may be observed in the increasing popularity of streetcars, fountain pens, and vinyl records.

The case of the independent bookstore has implications for many other industries, especially with the rapid advancement of modern technology. The ability of an industry to change its identity to adapt to changing economic conditions or of social movements like localism to drive significant market change may be deeply relevant wherever an industry may be confronted with the need to adapt.

Improving the Value of Health Care Delivery

Antonia Chan
Economics, 2018

Harvard Business School

Advisor: Robert S. Kaplan, PhD

Mentors: Derek Haas, MBA; Mahek Shah, MD

As the payment structure of the American health care system changes from fee-for-service to value-based payment plans, such as shared-savings capitation and bundled payments, providers increasingly need to understand and control their costs. Current methodologies, such as using cost-to-charge ratios and relative value units, produce cost estimates that have little relationship to the cost of the resources actually used to deliver care to patients. Time-driven activity-based costing (TDABC), based on bottoms-up measurement of actual resource consumption over a patient's cycle of care, offers a more accurate form of cost measurement.

TDABC requires that providers know how long and how much of each resource—whether personnel, equipment, space, supplies, or technology—is required during a cycle of care for a variety of medical conditions. Our research centers on the use of TDABC to help healthcare providers better measure and improve the value of care they deliver (outcomes relative to cost). A key component of this research is working with clinicians to understand and then redesign care delivery processes to improve their efficiency and effectiveness. In a project with Massachusetts Eye and Ear Infirmary, I observed a same-day optometry clinic's operations and designed a process map to better understand the breakdown of work, resource utilization, and time. These observations will facilitate standardized data collection for further TDABC research in same-day optometry. Our ultimate goal is to understand how to optimize cost and quality when delivering health care, translating our research into better outcomes for patients and providers, and reducing the cost to payers.

Employees and Startup Success

Benjamin Cohen
Statistics & Computer Science, 2019

Harvard Business School

Advisor: Rory M. McDonald, MBA, PhD

Mentor: Alicia DeSantola

As startups grow and prosper, or wither and die, how do the roles and responsibilities of employees change? How do employees affect firm performance? In traditional startup folklore, employees—especially early employees—and founders are considered to be vital to startup success. Traditionally these roles have been evaluated using qualitative methods, making patterns that only exist in the aggregate hard or impossible to identify.

Our research seeks to remedy this problem by applying quantitative methods to a novel data set from a widely used career networking site. The dataset includes an entry for employees of recently founded American startups and their experiences before and after working at a startup. After using natural language processing techniques and other traditional cleaning methods that translate employees' input into standardized forms, quantitative methods will be used to systematically identify patterns to understand how employee selection relates to startup success. By aggregating the data on the startup level, a timeline of the firm's employees can be created, helping answer questions about firm evolution and venture capital pressures. By analyzing these startup timelines, we can understand how different roles change over time within a startup. Finally, when this timeline is combined with data on venture capital funding from crunchbase, we can determine how startups react and adjust to the venture capital process. Understanding the answer to these questions will contribute to a growing discussion about why some startups succeed while others fail.

Private Equity Today

Brian Cornyn
Economics, 2018

Harvard Business
School

Advisor: Paul
Gompers, PhD

Mentor: Lauren
Cohen, MBA, PhD

The private equity industry is one of the relatively newer aspects of the financial world. However, much has changed in recent years around many of the largest private equity firms. Returns have fallen dramatically from fabled levels, with no rise seemingly in sight. Obviously, a myriad of stimuli and factors are influential in this outcome; however, the reduction of information asymmetry in the industry seems a strong driving force. Regardless of these recent falling returns, private equity firms still offer a higher return than intelligent investors can find elsewhere. For this reason, many important funds, most notably pensions, are heavily invested in private equity firms. This phenomenon has led to a large increase in cash on hand for private equity firms as they struggle to find places to invest it. I am going to attempt to collect some data to shed more light on how private equity firms have ended up in this situation and what options they are exploring to expand their targets.

Activist Investing

Carolina Brettler
Economics, 2018

Harvard Business
School

Advisor: Suraj
Srinivasan, MBA,
PhD

My research with Professor Suraj Srinivasan focuses on activist investing and its role in corporate governance in Japan. Activist investing refers to situations when an individual or a group of shareholders of a company exercise their rights as owners and attempt to actively influence a corporation's behavior or management. The goal of shareholder activism is to reform a company's policy or management in order to increase shareholder value. Hedge fund-driven shareholder activism is an important and controversial development in the U.S. business world, but is much less common abroad.

Despite Japan being the third largest economy in the world, it has been experiencing stagnation and declining corporate productivity for the past twenty years. I am researching and writing a background note on corporate governance reforms in Japan, focusing on recent changes in corporate governance and their effects on corporate practice. I have been examining official governance codes, analyst reports, news articles, and academic papers to address the following questions: What are the issues in corporate governance in Japan that have motivated reform? How are the reforms being enacted? How have reforms changed corporate behavior and practices?

This background information is essential to better understand individual cases of reform and activism in Japanese companies. I have also been researching and helping write a case study on the activist American hedge fund, Third Point LLC, and its involvement with the Japanese retail company Seven & i Holdings Inc. I have collected information including news articles, analyst reports, interview transcripts, and official company documents, to examine the effect of the hedge fund's involvement on Seven & i's company practices. Seven & i manages the Seven-Eleven convenience stores, Denny's Japan restaurant chain, and the Ito-Yokado supermarkets, among many other subsidiaries. Despite it being the most prominent retail company in Japan, the company, according to Third Point, suffers from inefficient capital structure and offers a low dividend for shareholders. My research examines the internal management struggle of Seven & i and the influence Third Point has had on improving profitability and company management practices. The case study should be used to address larger questions on corporate governance and shareholder activism: How does an activist investor deal with company management and vice versa? What practices in corporate governance are effective?

Digital Disruption

Christina Chiu
Economics, 2018
Wellesley College

Harvard Business
School

Advisor: Sunil
Gupta, PhD

In its simplest form, marketing is the process by which a company creates and keeps a customer. Classic marketing formulas dictate that as long as a brand has the right product, placement, price, and promotion, then any good product will sell itself. However, the rapid growth of technology has transformed how people live—and by extension—shop. Social media platforms have changed the way that companies communicate with their customers. Native advertising engines, such as BuzzFeed and YouTube, allow firms to produce valuable content for their target audiences. The wealth of consumer data available with the use of search engines allow firms to customize their advertising to target their desired audience at precisely the right moment. However, though firms are get-

ting "smarter" with the amount of consumer data available, consumer trends show that people are too. Consumers have become more keen on brand authenticity and corporate social responsibility, and generally have shorter attention spans regarding advertisements in an on-demand world.

With Professor Sunil Gupta, I am studying the changing relationship between firms and consumers in the digital age. Using both quantitative and qualitative aspects of marketing, we aim to determine how marketing strategies differ between startups and incumbent firms, across different industries, and between different countries. We aim to understand how companies are evolving their customer management and marketing strategies as they move to digital. Our methodology involves reviewing industry examples, case studies, and academic literature to form an analysis of current and future marketing trends. In addition to creating a great, usable product, filling a void in consumer demand, and setting a fair price, the shift towards digital technology ensures that firms must embrace an online presence that provides an additional valuable service to its target customer base.

Case Study on Kids & Company's Expansion from Canada to United States Childcare Markets

Christina Qiu
Applied Mathematics
& Economics, 2019

Harvard Business
School

Advisor: Boris
Groysberg, DBA

Mentor: Kate
Connally

This case study aims to examine the expansion of Canadian childcare company Kids & Company into the United States childcare market. Founded in 2002 by CEO Victoria Sopik and CFO Jennifer Nashmi, Kids & Company conceptualized itself as corporate childcare, functioning as an introduction of American backup or emergency childcare models into the Canadian market. While emergency and backup childcare models specifically target the needs of working parents who accommodate unexpected interruptions in childcare plans by taking sick days or leaving their jobs to care for their children when space in other facilities is unavailable, the Kids & Company business model remains "corporate" for other reasons as well. The company's main clients are large Canadian employers or multinationals, who for $5,000 a year can purchase a company membership that allows its employees to access Kids & Company facilities across Canada. Parents can only access these facilities if their employers have paid

this yearly fee. These facilities currently, in addition to emergency childcare, include nighttime and daily childcare, elder care, and extra meals. In a business world where work-life conflicts cost Canadian companies an estimated $4.5 to 5 billion a year and in a childcare market plagued with long waiting lists and inflexible options, Kids & Company, which ensures adequate space for children, high-quality care, and no waiting list, is an attractive option for both employers and working parents.

In addition to fulfilling working families' needs concerning childcare, Kids & Company has become a remarkably profitable and fast-growing company through aggressive growth strategies, which Sopik describes as a "greenfield build." The effect of this strategy was larger visibility within working family groups in Canada, which in turn attracted more corporate clients. This strategy of expanding where the company sensed a need for childcare, however, also resulted in expansion into less strategic areas or overestimation of centers' capabilities. Regardless, aggressive expansion allowed Kids & Company to receive much recognition as a fast-growing business, winning titles such as Best New Business from Winnipeg Biz in 2013 and the prestigious Ernst & Young Entrepreneur of the Year Award in 2012. In 2014, Sopik announced the company had "basically finished building [its] footprint in Canada" and opened its first non-Canadian center in Chicago.

Expansion into the United States childcare market may provide challenges for the company, as the United States industry may differ from Canada's in terms of quality regulations, parental preferences and needs, alternative childcare available, and government subsidies. This case study aims to identify these challenges with regard to Kids & Company's expansion and assess the company's decisions in addressing these challenges.

Multi-Sided Platform Strategies

**Daniel
Nightingale**
Economics, 2018

Harvard Business
School

Advisor: David
Yoffie, PhD

Platform businesses are rapidly transforming the world economy in a variety of business sectors by bringing together different market participants, enabling easy interactions and organized innovation in the process. Examples include large tech giants, such as Facebook, Google, Amazon, and Apple; unicorn startups, such as Uber, Airbnb, and Snapchat; credit card companies; dating websites; OS developers; and online marketplaces,

among many others. Uber, for example, works as a platform by matching drivers with riders. While these companies have many attractive qualities, such as the potential for rapid growth through the power of network effects, they also face many challenges. Platforms often struggle with clearing hurdles to obtaining critical mass, developing pricing strategies that promote growth, and effectively governing all platform participants.

This summer, I am working with David Yoffie, the Max and Doris Starr Professor of International Business Administration at Harvard Business School, as well as with Michael Cusumano, the Sloan Management Review Distinguished Professor of Management at the MIT Sloan School of Management, and Annabelle Gawer, the Professor of Digital Economy at Surrey Business School. Yoffie, Cusumano, and Gawer are currently writing a book on platforms in order to bring some balance and nuance into the discussion of platform companies. With the hype for platforms in recent years, they are attempting to dispel some of the current misconceptions related to platforms, stress the challenges of creating a successful platform, and illustrate the significant overlap between the operation of platforms and traditional businesses. They also hope to classify platforms into categories of similar business models, as current discussions of platforms combine companies that function quite differently.

I began my research this summer by focusing on the financials of platform companies. We are trying to answer basic questions about the profitability, value, and growth patterns of platform companies and the different factors that affect these metrics. I am using a variety of methods in order to parse the data and look for significant findings. Additionally, we are beginning to look more closely at the development of hybrid platforms, platforms that combine elements of transaction and innovation platforms. This has involved gathering detailed financial information on Amazon, Google, Facebook, and Apple. I might eventually help with case studies for specific chapters of the book as well.

Managing Patient Experience, Medical Outcomes, and Cost in Healthcare Operations

Grant Zhao
Chemical and Physical Biology, 2018

Harvard Business School

Advisor: Ananth Raman, MBA, PhD

Patient experience plays a fundamentally important role in healthcare. A patient expects not only to receive proper treatment and positive medical outcomes, but also to be treated with respect, to have their concerns listened to with meaningful communication, and to have their wait times, room cleanliness, and noise levels be managed appropriately. However, despite this importance, hospitals have historically only emphasized medical outcomes as a priority, often at the expense of positive patient experience. Furthermore, previous fee-for-service payment systems incentivized higher costs and volumes of patients over improving quality of care and patient experience.

In recent years, hospitals have faced increasing pressure to improve their patient experience. Most notably, the Center for Medicare and Medicaid Services (CMS) has adopted a pay-for-performance or value-based purchasing (VBP) payment system, basing hospital reimbursements partly on patient experience scores. By 2017, hospitals with low patient experience scores may have up to 2% of their annual reimbursements deducted, a critical policy in a time of significant financial hardship for the healthcare industry. Furthermore, with increasing competition between hospitals, patients are placing experience scores at a higher priority when choosing hospitals and doctors.

This research project aims to explore the drivers of patient experience, medical outcomes, and costs in healthcare operations. Qualitatively, I look at cases such as the Cleveland Clinic, which made significant efforts to improve patient experience and jumped from the 55th to the 92nd percentile in overall satisfaction from 2008 to 2012. Quantitatively, I study publicly available data from the CMS, particularly in the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS), which is the patient satisfaction survey required and used by CMS to determine hospital reimbursements. I create specific datasets from the survey information, and by analyzing them, I look for trends in patient experience, both overall and in particular distributions by geography, hospital type, hospital size, etc. Lastly, I hope to perform analyses on individual

patient-specific data to discover patterns in patient experience based on illness type and severity, treatment duration, age, sex, race, or other factors.

By analyzing trends in patient experience, we may gain a deeper understanding of its drivers and their relation to medical outcomes and costs in healthcare operations. These findings may provide broader insights to hospital management and healthcare providers into how best to improve and maintain a high quality of care while keeping lower costs.

Migration, Diasporas, and Business Networks

Hellary Zhang
Applied Mathematics & Economics, 2018

Harvard Business School

Advisor: Meg Rithmire, PhD

Overseas Chinese have had a profound influence on the economic development of China, even after they have left their home country. Although existing academic research has examined the influence and importance of the *huaqiao* (overseas Chinese) community in specific geographic areas, and although the Chinese government has kept extensive records of the volume of emigration from China, no one has compiled this information capturing the Chinese diaspora. Many questions are unanswered: What does the entire Chinese diaspora look like on a world map? How far-reaching are their movements? How can we visualize the transactions between *huaqiao* communities and their hometowns? Professor Rithmire's research seeks to build an open-source website that maps the movements of members' hometowns to Southeast Asia, the United States, South America, and beyond. We will be compiling data from a wide range of sources at varying administrative levels—from local gazetteers on overseas Chinese to national statistical yearbooks collected from census data—to create a visualization of the magnitude of the Chinese diaspora and its economic influence on China through foreign direct investment (FDI), remittances, and other flows of capital and resources. The hope is that this website will be easily accessible to academics and the general public as an educational tool, as well as serve as a portal through which researchers can easily contribute their data and findings on the *huaqiao* community. We also hope that, upon building this database, we will be able to draw conclusions about the Chinese diaspora's business networks.

Pharmaceutical Innovation: Economics, Incentives, and Regulation

Holly Breuer
Molecular and Cellular Biology, 2018

Harvard Business School

Advisors: Ariel D. Stern, PhD; Amitabh Chandra, PhD

As advances in medical technology have allowed for the discovery and development of new drugs, millions of lives have been improved and even saved. However, pharmaceutical innovation has come with very high costs. Recent estimates of the cost of development of a new drug have been pegged at $2.6 billion, while new precision medicines that promise to personalize treatments and therapies for many diseases are expected to be even more costly. In the face of such high costs, our healthcare system now faces an industry landscape vastly different from, and more challenging to navigate than, that of ten years ago. In this context, the research of Professors Ariel Stern and Amitabh Chandra focuses on the economic incentives that drive biopharmaceutical innovation and the development of new drugs, including regulatory dynamics like FDA programs that make it easier and less costly for certain kinds of drugs to receive approval.

In order to become acquainted with the current biotech landscape, I have been researching and begun writing two HBS cases for use in Professor Stern's class "Demystifying Data: Managing with Analytics." Each case examines an up-and-coming player in the healthcare space: the first, Genepeeks—a Cambridge-founded company that seeks to provide an innovative, analytical approach to carrier screening for genetic diseases; the second, PillPack—a mail-order pharmacy providing new packaging methods to help the growing number of American patients managing multiple medications. Through company interviews and further literature research later this year, we'll be able to focus each case on a central operational, analytical, or technological issue that will allow Harvard Business School students to take on the role of decision-making company executives in the classroom. I've also used my background in biology to assist Professor Chandra in the preparation of a presentation explaining the dynamics of Hepatitis C virus and translating the biology of current drug treatments for a non-science audience.

In another project this summer, we'll examine the role of public research funding in recent biopharmaceutical innovation. We are currently in the data collection phase of the project, which involves match-

ing clinical trials for new drugs with any NIH grants that played a role in their discovery and development with information from the Thompson Reuters Cortellis Competitive Intelligence Database. From understanding current industry innovations and incentives to examining healthcare's regulatory landscape and high costs, our work seeks to shed light on the challenges that have come with improved medical outcomes in the hope of understanding how modern society can adapt to modern medicine.

Government Spending, Income Inequality, and Optimal Taxation

Jiafeng Chen
Applied
Mathematics, 2019

Harvard Business
School

Advisor: Matthew
Weinzierl, PhD

Income inequality has generated much interest for both academic and policy economists. Since governments often redistribute with taxation, optimal tax theory— the study of taxes that maximize welfare gains and minimize efficiency losses—is especially relevant. Marrying income inequality and optimal taxation, however, has two difficulties. On one hand, the literature on income inequality usually analyzes summary statistics of income distributions, such as the 90/10 ratio (the ratio of income at the 90th percentile to that at the 10th) or the Gini coefficient, but summary statistics do not easily lend to numerical simulations, an important analysis tool for optimal tax theorists. On the other hand, the literature on optimal tax traditionally treats taxpayers' income-earning abilities (abilities for short)—both innate and acquired—as exogenous. However, realistically, the trend of widening income inequality may partly come from widening inequality in the distribution of abilities; thus, treating ability as exogenous leaves a large hole in tax theorists' models. Indeed, factors like government policy and technological change can conceivably influence the distribution of income-earning abilities, hinting that ability may be endogenously determined.

Our research attempts to attack these two problems. We fit income quantile data (U.S. 1979–2013) to a probability distribution that hybrids the lognormal and Pareto distributions, with properties consistent with empirical observations in income. The fitting allows us to generate simulated individual income data. We then feed the income data through a simple labor-supply model and back out the ability distribution. Once we simulate a sample of ability from a sample of income, we fit the ability data to a probability distribution, which allows us to capture the over-time shifts in the distribution of income-earning abilities with a few parameters. We would like the parameters to be endogenously determined by factors like government spending and technological progress—to be determined by some sort of "ability production function." Empirical analyses, preliminary and speculative at this point, would then allow us to uncover how income-earning abilities change with respect to these explanatory variables. These empirical evaluations would then enlighten optimal tax theorists to sensible, endogenous-ability models that better attack income inequality problems.

Decision Making and Behavioral Economics

**Kevon
Edmondson**
Psychology, 2018

Harvard Business
School

Advisors: Alison
Wood Brooks, PhD;
Ryan W. Buell,
DBA; Francesca
Gino, PhD; Leslie K.
John, PhD; Michael
I. Norton, PhD

Mentors: Grant
Donnelly; Oliver
Hauser, PhD;
Ximena
Garcia-Rada, MBA;
Tami Kim; Ovul
Sezer

Economic theories generally employ the rational human—*Homo economicus*—who strives to maximize utility. This common, albeit idealized, convention allows economists to create overarching theories in a complex world. It should come as no surprise that our decisions are not always systematic and deliberate, as these theories often necessitate. This summer, I worked alongside researchers in the Harvard Business School NerdLab. This lab delves into the irrationality in human decisions and judgments via psychological and economic frameworks. In one project, we explored the role of chance encounters in the dictator game, where there is a dictator and a recipient. The dictator is tasked with allocating a sum of money between him or herself and the recipient. In our study, participants were assigned to either the "recipient" or the "other" condition. In both conditions, there was a chance that the participants would have to reveal their allocations. Dictators in the "recipient" condition would reveal their allocation to the recipient, while dictators in the "other" condition would reveal to non-recipients. In the experiment, dictators were prompted to allocate \$5 multiple times between themselves and a recipient, each time the probability of meeting the recipient or "other" increased. In the end, one of the allocations was randomly selected and the dictators were told whether they needed to disclose that allocation.The aggregated allocations of the dictators do not appear linear. The amount given to the recipient seemingly approached an upper bound

that was lower than $5, notwithstanding the increasing probability of a chance encounter. In a rational world, one would expect keeping all $5 to yield more utility when compared to only allocating $3 to oneself. However, the potential for receiving negative utility from keeping all $5 exists. People do not make decisions completely independent of others; instead, the actors may consider how another person may feel as a result of their decision. It is possible that being selfish (keeping all $5) could result in negative utility, as the dictator would be viewed as selfish by his or her peers. This could lead a dictator to allocate him or herself less than $5.

Strategy and Innovation in Nascent Markets

Lauren Nicholson
Sociology & Government, 2018

Harvard Business School

Advisor: Rory M. McDonald, MBA, PhD

Mentor: Cheng Gao

Nascent markets are defined by what is perhaps one of the most unsettling words in business—uncertainty. There are many variables that are difficult to forecast and, as such, they make it difficult for firms to plan for success. When competing in a market where everything from profit opportunities to regulations are unclear and unpredictable, understanding how firms can achieve competitive advantage in these fledgling fields is of paramount importance.

For my project with Professor Rory McDonald, we study competition and strategy in the nascent biotech field of direct-to-consumer (DTC) genetics testing industry. We collect data through archival research and interviews to create timelines for the top firms in the DTC genetics testing industry so that we can compare their methods and performance against our success metrics to help determine what exactly it is that separates the companies that thrive and ultimately dominate the market from those that ultimately flounder. We then use the timelines to engage in comparative case analysis to build theory on strategic actions firms can take to win in nascent industries. In addition to the DTC genetic testing industry, we also hope to examine these competitive dynamics in other nascent industries, such as fantasy sports.

The main application for this research is derived from the wildly popular idea of disruptive innovation, a term coined by Clayton Christensen which "describes a process by which a product or service takes root initially in simple applications at the bottom of a market and then relentlessly moves up market, eventually displacing established competitors." By examining what factors separate the firms that stay at the bottom of the market from those that rise and disrupt the traditional order and hierarchy, our research can hopefully help firms discern which rising companies in these nascent markets are more likely destined for success (and should thus perhaps be bought) and those more likely destined for failure (which they can leave to fizzle out on their own).

The Economics and Psychology of Consumer Behavior

Lingjun Xia
Economics & Psychology, 2018
Wellesley College

Harvard Business School

Advisors: Alison Wood Brooks, PhD; Ryan W. Buell, DBA; Francesca Gino, PhD; Leslie K. John, PhD; Michael I. Norton, PhD

Mentors: Tami Kim; Ximena Garcia-Rada, MBA

The growing field of behavioral economics differs from traditional economics by introducing a heavy dose of psychology. Indeed, the research in the NerdLab investigates psychological mechanisms underlying human behavior, oftentimes explaining irrational behaviors that cannot be fully accounted for by traditional economics. This psychological framework is particularly essential in fields such as marketing and consumer behavior. Two of the projects I participated in this summer study the decision-making process from the perspectives of consumers.

One project studies joint-consumption decisions. According to Gorlin and Dhar (2012), consumer decisions can be roughly sectioned into four quadrants. Along one dimension, one could make a decision for oneself (single-decision) or with others (joint-decision); along the other dimension, one could consume the product for oneself (single-consumption) or with others (joint-consumption). While most consumer behavior research focuses on single-decision, single-consumption scenarios, our project examines single-decisions that consumers make for joint-consumption occasions. Through an online survey, our results show that single-decision, joint-consumption occasions are ubiquitous in real life and thus it is important to understand the dynamics behind these choices. We will further analyze the data to look for consumption patterns, and follow-up lab experiments will dive into understanding the psychological mechanisms that shape these decisions.

The other project examines the influence of gender-

related typecasting on consumers' willingness to purchase certain products. This project shows that consumers tend to avoid products that evoke gender labeling: women may reject a pink pen labeled "pink for women," whereas they may have purchased the pen had there not been a label. Within this project, we conducted a field study simulating a political campaign: we asked female participants to choose between a sticker and a button, both having slogans supporting Hillary Clinton. We manipulated the slogan on buttons to sometimes say "candidate for women," and sometimes say "candidate for U.S." We hypothesize that participants will forego quality to avoid choosing the gender-labeled product. The result is close to a marginally significant level; however, since field studies involve many uncontrollable variables, usually resulting in weak statistical significance, the result is actually rather satisfactory. Going forward, we will conduct further studies along this path to examine under what conditions people will avoid gender-related typecasting. This study has applications in marketing, providing firms a smarter way to label their products.

Mental Models and Strategic Interaction

Osama Arif
Economics, 2018

Harvard Business School

Advisor: Dennis Yao, MBA, PhD

Mental models are simplified representations of the world. Many economic models assume that all firms have the same mental model of the world. In our project we use computational methods to relax this assumption and explore strategic interaction between firms with different mental models. Our line of research specifically involves studying competitive interaction between firms with different mental models of the world. In this case, different mental models entail different views of the demand function. We use a regression framework for mental models and a series of simulations in Python to conduct our analysis. We focus on Cournot competition in the case of a duopoly, in which each of the two firms chooses a quantity to produce and the market price is determined by the total quantity produced by the two firms and the demand function. The true demand structure in our simulations is $P\_t = \alpha + [\beta_1 x_1 + \beta_3 x_3]Q_t$. We introduce variation in the mental models of the firms by varying the "x's" in this structure, which are random shock variables that affect demand. One firm holds the correct mental model and has the same shock variables as the true demand. The other has an incorrect mental model and has a shock variable "x_2," which has no deterministic power for true demand versus the correct "x_1." Both firms have the correct "x_3" shock variable in their respective mental models.

The first step in our simulations involves generating a pre-history for each firm by randomly generating quantity choices and using each firm's mental model to generate a series of price-quantity data points for the initial calibration of the parameters of each firm's mental model. We explore several combinations of neutral, optimistic, and pessimistic priors for the mental models in this process. An optimistic prior entails a firm assuming demand to be inaccurately flatter than in reality and thereby producing an unprofitably large quantity, and a pessimistic prior entails the opposite. Next we play out the competitive interaction. This step involves randomly generating values for all shock variables in each firm's respective mental model and each firm subsequently choosing a quantity by maximizing profit with respect to the parameters of its mental model, the observed shock variables, and the competing firm's quantity—which the firm assumes to be the same as its own quantity. A market-clearing price is then generated based on the the true market demand function. This process is repeated for twenty periods, and the parameters of each firm's mental model are re-calibrated based on its history of quantity-price points at the start of each new period as per our regression framework. The simulations are repeated over 10,000 iterations for consistency and statistical power.

One of our initial findings is that the firm with relatively optimistic priors generally generates larger profits, but this outperformance only persists in the long run when the firm with the incorrect mental model has these relatively optimistic priors. This finding can be explained by the fact that the regression framework allows the firm with the correct mental model to eventually adjust its parameters to the correct ones found in the true demand, but the incorrect mental model does not allow for such an adjustment. Actors' representations of the world are a crucial element of all types of strategic interaction and our project seeks to shed light on how varying these mental models held by firms would affect competition.

Understanding Task Specialization Versus Task Diversification Through Real-Life Examples

Rohan Reddy
Applied
Mathematics, 2017

Harvard Business
School

Advisor: Mihir
Desai, MBA, PhD

On March 19, 2012, Apple disclosed the initiation of dividend and share repurchasing program: "We have used some of our cash to make great investments in our business through increased research and development, acquisitions, new retail store openings, strategic prepayments and capital expenditures in our supply chain, and building out our infrastructure. You'll see more of all of these in the future," said Tim Cook, Apple's CEO. "Even with these investments, we can maintain a war chest for strategic opportunities and have plenty of cash to run our business. So we are going to initiate a dividend and share repurchase program." In 2013, Professor Mihir I. Desai initially detailed the execution of the capital return program in a Harvard Business School case study titled "Financial Policy at Apple, 2013 (B)." Since that time, there has been a series of important additions to Apple's capital return initiative that may provide insight into the company's internal operations. The purpose of this study is to analyze the additions to the initiative from 2013 to 2016 and determine overall changes to Apple's shifting capitalization structure, balance sheet position, operational performance, and stock performance. This information will be then used to update the original case study.

Family Influence in Venture Capital

Steven Tan
Statistics, 2018

Harvard Business
School

Advisor: Paul
Gompers, PhD

Mentors: Lauren
Cohen, MBA, PhD;
Kanyuan Huang

The importance of venture capital remained relatively nonexistent until the late 1990s and the rise of the Internet bubble. Since then, venture capital has become a driving force in providing entrepreneurs the ability to pursue startups and business endeavors. Among established venture capital firms, one can see the increase of family firms where the family members become co-founders and CEOs. Because one has the closest relationships with family members in most circumstances, it is understandable for entrepreneurs to entrust their siblings or children to continue leading the firm. However, studies done in Denmark have shown that family successions have a negative causal impact on firm performance as operating profitability on assets fall during CEO transitions. Moreover, employers tend to hire employees not only for ability-based characteristics (e.g. graduating from a top university) but also for affinity-based characteristics (e.g. ethnicity), despite the fact that such decision-making decreases overall productivity and success in the firm. From my research, I hope to use new venture capital data to specifically address the increasing percentage of family firms. I will distinguish family history by matching surnames and information provided online. Afterwards, I will run regressions and statistical analyses to determine whether family firms are performing as well as non-family firms. I will compare differences in the number of trade deals and IPOs to measure performance and determine a difference in performance between family versus non-family firms.

Decision Making and Behavioral Economics

Yilin Chen
Economics, 2018
Wellesley College

Harvard Business
School

Advisors: Alison
Wood Brooks, PhD;
Ryan W. Buell,
DBA; Francesca
Gino, PhD; Leslie K.
John, PhD; Michael
I. Norton, PhD

Mentors: Karen
Huang; Ovul Sezer

The study of decision making and behavioral economics lies between the fields of psychology and economics. This summer I worked with the NERD lab, which consists of a group of professors and doctoral students from different academic fields including marketing, organizational behavior, and management. The research at NERD uses fundamental psychological and sociological theories to investigate human judgments and decisions. I worked on several projects related to these fields throughout the summer.

One of these projects analyzes how people engage in conversation. Conversation is a pervasive human experience, one that is necessary to pursue intrapersonal and interpersonal goals across myriad contexts, relationships, and modes of communication (e.g. written, spoken). In a conversation, people make a series of decisions about what to say and how to respond. The researchers collected data both in lab and through MTurk (Amazon Mechanical Turk), an online platform. We paired participants into dyads to chat for fifteen minutes. All conversations were captured on ChatPlat, a chat software. For each set of conversations, we analyzed the turn-

by-turn text in the conversation, as well as overall
impressions of the people who engaged in conversa-
tions. Specifically, we analyzed the number of types
of questions (e.g. follow-up, switch, and rhetorical),
and the degree to which each chatter felt heard and
listened to. Investigating these questions will help
us as researchers better understand the psycholog-
ical and linguistic mechanisms underlying human
conversational behavior.

SHARP

SHARP

"Reframing Reality": Exploring Artistic Representation in the Harvard Art Museums

Aaron Suduiko
Philosophy, 2017

Harvard Art
Museums

Advisor: David Odo,
PhD

Mentors: Correna
Cohen, EdM; Erin
Northington, MTS

How can art lead us to see the world differently? I aimed to answer this question in my tour, "Reframing Reality," at the Harvard Art Museums. I explored various ways in which art can represent the world as different from our ordinary perception of the world, thereby encouraging museum-goers to challenge their preconceived notions of what the world is like.

The first object I selected for my tour was *Endless Surface in the Form of a Column* (1958), by Max Bill. The piece is a roughly nine-foot-tall, metallic Möbius strip mounted on a wood-block base. Bill was a pioneer of "concrete art," an artistic movement in which artworks neither reference nor represent the natural world. Concrete art thus generates an aesthetic based solely on the art object's intrinsic form and the viewer's spatial and optical relation to that form. The Möbius strip, a surface with only one side, epitomizes this artistic aim: as a form not found in nature, the strip challenges viewers to reconsider their relationship with space.

The second object I selected was *Still Life with Inkwell* (1911–1912), by Pablo Picasso. The fragmented, geometric painting, typical of Picasso's analytic cubist period, challenges the viewer to locate the objects it purports to represent: an inkwell, a book, a collection of papers, and some writing instruments. Around the turn of the twentieth century, mathematicians such as Charles Howard Hinton popularized the idea of a fourth spatial dimension, which humans, existing only in three dimensions, could not perceive. Cubist painters sought to represent this higher-order space by painting objects from multiple perspectives at once, as Picasso did in this painting. One can see this especially well in the bottom right area of the painting, where we see both the outside of a book and pages from inside the book. Though this fusion of perspectives aims to capture a greater reality about the represented objects, it also renders the objects virtually unrecognizable, thereby challenging viewers' belief that their perception reflects the world as it really is.

The third and final object on my tour was a scroll painting of *Courtesan and Attendant*, by Miyagawa Chōki (active c. 1716–1751). The *ukiyo-e* scroll captures an instance of the real world influencing art, which in turn influences the real world. Eighteenth-century Edo was known for its pleasure district, a result of wealthy merchants being trapped at the bottom of Japan's class system, socially immobile and in need of ways to spend their money. Artists at this time began painting in a style designed to capture their impression of reality, rather than a mimetic representation of the world around them. These artistic impressions then influenced the standards by which real people lived their lives. Thus, an artist would paint his idea of a courtesan and thereby generate standards of the feminine ideal to which real courtesans would aspire. This particular artwork especially evokes the idea that the artist represents a type of person rather than a specific person by representing the courtesan and her attendant against a blank background, divorcing them from any broader context. I ended my tour with this object to suggest that art doesn't just provide us with new ways of seeing the world: precisely by changing people's perspectives on the world, art also changes the world itself.

Accessible Podcasts at Poetry in America

Brieanna Martin
English, 2018

Poetry in America,
edX

Advisor: Elisa New,
PhD

Mentor: Leah Reis
Dennis

My research this summer has been with Professor Elisa New's multiplatform digital project Poetry in America, which aims to show poetry's relevance and accessibility to a wider audience through conversations with diverse guests, from Bill Clinton to high school students to Nas. I've been working with her team as they develop and produce content for several several online courses, K–12 teacher-training courses, and a public television series in production with WGBH Boston.

I think of my work this summer as helping answer the following questions: How do we make poetry and conversations about poetry interesting, accessible, and relevant to people they wouldn't otherwise touch? How can digital media be used to make the range of poetry readers as diverse as the voices speaking through the medium?

My work so far with Poetry in America has consisted largely in developing podcasts from educational content that was previously available only through classes on the edX platform and within the Harvard Extension School. I watch the raw conversations between Lisa and guests, then edit and rearrange them so that they follow an intuitive narrative arc, and make for thought-provoking, enjoyable listening for poetry enthusiasts and those with no experience alike. I've been learning the technical skills necessary to edit and record audio, using Final Cut and learning the proper use of audio equipment, as well as applying the skills I've learned in two years of English seminar classes to decide which content to feature and what to include.

I've also been assisting the team here in developing content for the 1945 to the Present module of one of their online courses, Poetry in America for Teachers: "The City" from Whitman to Hip Hop. I've spent some time in Lamont's Woodberry poetry room searching for poems related to one another and brainstorming people whose lives they might relate to so we could bring them in for interviews. In this work, I'm exploring how poems can be accessible and relevant without sacrificing complexity, beauty, or freshness of expression.

By the end of the summer, I will have produced a cohesive and interesting series of podcasts that are accessible to anyone in both content and form. I will also have a few answers to the question of how poetry can be a mode of expression and communication on a more basic human level, enjoyed and appreciated by many people outside the academic world.

Femmes Littéraires: Une Histoire Culturelle

Caleb Shelburne
History and
Literature, 2018

Harvard University

Advisor: Christie
McDonald, PhD

Mentor: Kylie Sago

In Enlightenment-era France, women were widely considered naturally subordinate to men in intellect, emotional resilience, and imagination. This bias influenced even the most "enlightened" thinkers, whose medical articles, philosophical treatises, and fictional endeavors perpetuated uncritical understandings of women's physiology and mental capacity. Nevertheless, despite biases in education, restrictions on publishing, and conventions that disdained female authorship, women did write, producing works in all genres that complemented and challenged the philosophical explorations and artistic productions of their male contemporaries. Many of these women achieved great fame in their own lifetimes, but all but a few were forgotten by the 19th century, when men such as Voltaire, Diderot, and Rousseau were becoming almost mythical figures. Only in the second half of the twentieth century did feminist scholars begin to resurrect these brave women: Émilie du Châtelet, Françoise de Graffigny, and Olympe de Gouges finally appeared on university syllabi and in critical editions next to the men they had debated. Even today, however, many women, especially those in genres like theater and journalism, remain to be discovered.

Foucault, reinterpreting Kant, describes the Enlightenment as a philosophical perspective rather than a period—"a permanent critique of our historical era." In this spirit, our project aims to uncover the long-neglected women writers of the eighteenth century. Who were these women? What unique institutional and social prejudices constrained their ability to write and publish their work in any genre? Which philosophical issues did women writers raise or engage with? How did women writers navigate a society that considered them destined for the private sphere and incapable of creativity? A survey of these women, both the greats and the practically unknown, is much needed as a reminder of their presence and as an attempt to identify the tactics they used to approach an unfriendly society.

Preliminary findings reveal that women did write in all genres: epic poetry and metaphysics to vaudeville and republican journals. While many women wrote without challenging the status quo, others advocated for greater gender equality, relying on subtle allusions or subversive genre shifts to contest the literary and social establishment. Even if they were overlooked by the literary historians of the nineteenth centuries, these women contributed to French literature and the Enlightenment with just the sort of critical mentality that Foucault identifies. *Femmes littéraires: une histoire culturelle* will finally provide the recognition they deserve and, in doing so, will complicate and extend our understandings of the period and its people.

Investigating Wuzhou Networks with CBDB

Edith Claire Enright
Comparative Literature, 2018

Fairbank Center of Chinese Studies

Advisor: Peter Bol, PhD

Mentor: Hongsu Wang, MA

The China Biographical Database (CBDB) is a computer database that aims to catalog all recorded biographical and relational information from pre-modern China; currently, the database contains almost 400,000 individual biographical entries from throughout Chinese history, and is an especially invaluable resource for the quantitative investigation of scholarly, official, and familial links in middle-period China (960–1400 CE). The China Biographical Database project at Harvard is headed by Professor Peter Bol and Wang Hongsu, and I am conducting research using the database for the first time this summer.

This SHARP project entails using the information recorded in this database to investigate the degree of local interconnectedness of intellectual, literary, and kinship ties between members of the literary elite of Wuzhou prefecture in China's Southern Song (1127–1279) and Yuan (1271–1368) dynasties, a period coinciding with the emergence and spread of Neo-Confucian philosophy. By mapping these networks, we can determine the response of these elite communities to the dynastic shift that occurred in the late thirteenth century when the Mongol empire conquered China's Song dynasty and founded the Yuan dynasty. With the help of the geographical imaging software QGIS and the social network mapping software Gephi, we have mapped the social connections of Wuzhou residents across China and analyzed the extensive social and kinship ties between members of the Wuzhou elite in order to determine the degree of local connectedness as well as the most important actors in the network—often members of major literati families who are related to imperial degree holders.

We are now investigating the social networks of other Song and Yuan dynasty prefectures for purposes of comparing their extent and demographic makeup with the Wuzhou network. Moving forwards, we plan to construct database queries that will allow us to determine the local origins of authors of major texts, as another way of mapping the geographic distribution of literary culture. We intend for the quantitative data yielded by CBDB to be useful in evaluating the hypothesis that Chinese social and literary connections became locally denser and richer

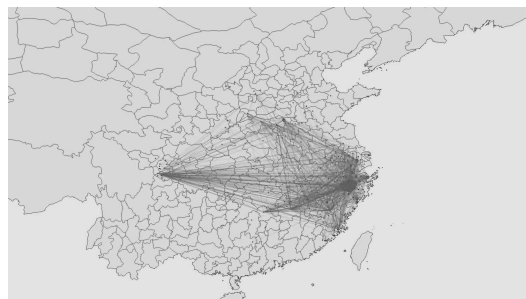alongside the flourishing of a Neo-Confucian ideology during the Yuan dynasty.

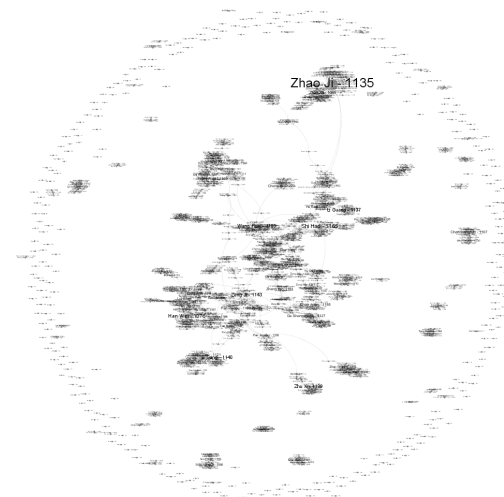Figure 1: Network connections of Wuzhou residents in the Southern Song dynasty.

Figure 2: Section of Wuzhou kinship network in the Southern Song.

Femmes Littéraires: Une Histoire Culturelle

Henry Shreffler
History and Literature, 2018

Harvard University

Advisor: Christie McDonald, PhD

Mentor: Kylie Sago

This project aims to explore the literary productions of French women of the 18th century. While historically neglected by scholars until the second half of the twentieth century when scholars began to recover the works of some of these writers, women of the period created literary masterpieces in all genres. Women of all classes wrote plays, novels, philosophical treatises, and more, engaging with men on complex issues of the time, all the while facing significant structural barriers to their participation in the literary sphere. Men and women alike perpetuated harmful falsehoods about the female condition, asserting that

women were unable to reason, should read but not write, and were best occupied distracting themselves with their makeup and idle nothings at their toilettes. While men were expected to be involved with the public sphere, it was widely understood that women were to remain in the private sphere. Works such as Rousseau's *Émile* only accentuated this prejudice, lauding women for their role as mothers.

Nonetheless, some women wrote, and achieved success and fame through their literary works. However, even as published authors, women often wrote anonymously, and often faced having their work misattributed to the men in their lives. *Femmes littéraires: une histoire culturelle* aims to highlight the important contributions of French women of letters in their respective fields, rewriting a male-dominant historical narrative to better reflect the historical reality.

My Work of a Different Nature: John James Audubon and Nineteenth-Century America

Jess Clay
History and
Literature, 2017

Houghton Library

Advisors: Emilie
Hardman, MA, MLS;
Heather Cole, MA,
MLIS

In the first half of the 19th century, John James Audubon traveled across America to paint the birds that would make him famous. His experiences took him from the Everglades to Acadia, from riverboat journeys to transatlantic passages, and from the backwoods of Kentucky to the White House dining room. He became the most celebrated naturalist in America, and arguably one of the nation's best-loved artists. My research engages with the life and work of Audubon as a reflection and fulcrum of the early republic. As a Houghton Library–SHARP fellow, I undertake self-directed research in a wide range of fields pertaining to Audubon. Beyond his enormous contributions as an American artist and scientist, I study the more literary and political aspects of Audubon and his work. In particular, I explore Audubon's influences and those he influenced, his involvement with Romanticism and Transcendentalism, and his contributions to subjects ranging from scientific advancement to national identity.

My primary resource is Houghton Library's John James Audubon collection, featuring his artwork, letters, journals, manuscripts, and published works. Among these are his early drawings and mundane business transactions, as well as first editions of his magnificent Birds of America and a letter to Daniel Webster. The collection also features documents from family members which pertain to Audubon,

along with early biographies and readers. Other collections within Houghton prove equally rewarding. The John Keats collection, for instance, not only offers a window into Romanticism, but also into the tumultuous personal relationship between Audubon and the Keats family. I also work with secondary sources in other libraries, as well as the more scientific holdings of Harvard's Museum of Comparative Zoology. Such collections mirror the variety and richness of Audubon's own experience. In both my personal research and writings, as well as a planned public exhibition, I put various figures and movements in a dialogue with Audubon at the center.

Decolonization and the Constitution of the International Order

Lev Asimow
Social Studies, 2018

Harvard Law School

Advisor: Samuel
Moyn, JD, PhD

The wave of decolonization that swept the globe between the 1940s and 1970s brought to fruition the ideal of self-determination for upwards of a billion people across the Global South. But beyond redrawing the map, the collapse of Empire in Africa, in Asia, and in Latin America rendered uncertain the foundational precepts of the post-War international economic and political order. While this paradigmatic break with the structuring logics of Empire spurred the birth of nation-states throughout the formerly-colonized realm, this result was by no means foreordained.

The first aspect of this inquiry, thus, is into the victory of nationalism as the dominant principle for political self-determination, and the fading of alternative organizing models such as federalism and departmentalization. Understanding this dynamic, by reference to the writings and speeches of anticolonial leaders, can provide unique insight into both the development of economic autarky, or self-sufficiency, and the modern conception of the nation-state that shaped policies such as 1974's New International Economic Order.

Secondly, this project seeks to conceptualize the notions of equality, development, and social betterment mobilized by diverse actors in the formerly colonial world, and to investigate the varied politico-economic methodologies proposed for actualizing these ideals. It is thus jointly an exercise in political and intellectual history. Given the complex confluences of Soviet-style socialism and American capitalist globalization that exerted influence on decision makers in the Global South, casting light on the ideological conflicts and compromises that motivated

political change in the era of decolonization necessitates a better understanding of such phenomena as Léopold Sédar Senghor, Mamadou Dia, and Sékou Touré's African Socialism; Kwame Nkrumah's Pan-Africanism; and Julius Nyerere's *Ujamaa.* Economic growth was a paramount concern for the historically impoverished nations of the Global South, who collectively desired a greater international parity with their former colonizers. But certain socialist sensibilities motivated these fledgling nations to adopt the more ambitious goal they termed "development," denoting an intra-nationally equitable contribution to and division of the spoils of growth. When compared to the welfare state model that, in this period, committed the governments of the metropolitan Global North to establishing basic social minima domestically, the economic policy of these developing nations can provide valuable insight into the politics and ethics of international governance, economic aid, and political intervention.

Third and finally, interrogating the politics of decolonization provides a unique vantage point for theorizing about human rights and conceptions of global justice that persist today. As decolonization stretched the Western category of "the human" to its breaking point, motivated by theories of race, existentialism, and Marxism, the governing logics of rights, privileges, and distributive justice similarly were tested, expanded, and teased apart. By viewing decolonization as a theoretical and philosophical intervention, this project aims to more capaciously conceive the origins of the purportedly neutral model of liberal internationalist human rights. By bringing together these economic, political, and philosophical problematics, we hope to make a modest contribution to the genealogy of our decolonized world.

Remembrance and Renewal at the Harvard Art Museums

Rob Hopkirk
English & Religion, 2018

Harvard Art Museums

Advisors: Jessica Martinez, PhD; David Odo, PhD

Mentor: Correna Cohen, EdM

In November of 2014, the Harvard Art Museums reopened to great fanfare after an extensive six-year renovation. It was the culmination of a long effort to revitalize the Museums as an integral institution of teaching and learning for Harvard and the wider community. After that initial burst of publicity, however, it was clear that hidden within this apparent ending was another project just beginning to take shape. The Museums were reno-

vated, refurbished, and relit, but they could not merely function as a beautiful building. The space had to be activated with programming and outreach in order to become a dynamic environment for curators, students, faculty, and the public to interact with each other and the vast resources at their fingertips. The Museums are still in a process of introspection and self-evaluation. Now that they have moved past their renovation project, they will only continue to transform and grow.

This summer, I had the pleasure to play a small role in this growth as a guide in the Division of Academic and Public Programs. Through conversations with staff and independent research in the curatorial files, I was able to craft a tour that sampled from the diverse collections featured in the Museums. Each object on my tour, Remembrance and Renewal, is the work of an artist who, in the process of creation, looked to the past for inspiration while simultaneously reinventing it for new generations. I emphasize evolution over revolution. In both museum galleries and the public consciousness, the history of art is usually a narrative of innovation and supersession. My hope is that visitors come away from my tour with a more critical eye and a willingness to see artworks not as autonomous creations, but rather as links of continuity between the cultural traditions that precede and follow them. My variegated tour focuses on a fresco depicting police brutality during the Great Depression, a portrait mask made for a mummy in Roman Egypt, and a monumental altarpiece from Renaissance Venice dedicated to lost love and heavenly hope.

My overarching concern is the story of the renovated Museums, caught at the crossroads of tradition and possibility. On one hand, I hope to demonstrate that the Museums offer a portal the past, with objects that conjure up vanished cultures and faded artistic movements. I also hope, however, to show that the Museums are a gateway to the future, with new research and initiatives, such as the Student Guide program, to help make it a space that is more accessible and inclusive for visitors of all ages and backgrounds. Like the objects they house, the Museums act as mediator between the past and the future, as well as between the artist and the viewer. Each tour I give is a new opportunity to bridge the gap that prohibits visitors from deeply engaging with the art. The renovation may have entered into the realm of memory, but the renewal of the ever-evolving Harvard Art Museums begins again each morning.

An Undiscovered Contribution to Egyptian Archaeology

Sarah Judd
Statistics, 2019

Harvard University

Advisor: Peter Der Manuelian, PhD

From 1905 to 1947, the Harvard University–Boston Museum of Fine Arts Expedition enjoyed spectacular success at twenty-three archaeological sites along the Nile in both Egypt and Nubia, now modern Sudan. This success can be seen in the wondrous collection of Boston's Museum of Fine Arts, at Harvard University's Peabody Museum, and in museums in Cairo and Khartoum. One key source for presenting the important story of the Expedition and its leader, Harvard archaeologist George Reisner, is the Ashton Sanborn papers. The papers are a collection of hundreds of letters by Ashton Sanborn (1882–1970), a member of the HU–MFA Expedition, and lay undiscovered in the archives of the Schlesinger Library until just this past year. This collection includes descriptions of work at the sites of Dendera and Memphis for the University of Pennsylvania Expedition; famous discoveries at the Giza Pyramids and at Meroe (Sudan), as well as comments on Reisner's archaeological field methods for the HU–MFA Expedition. There are also mentions of many great Egyptologists, important world events and visitors to the sites, as well as descriptions of urban life in Cairo and rural life further up the Nile. Other Sanborn letters, written from the Museum of Fine Arts, Boston in 1925 to his wife temporarily living in Bryn Mawr, illuminate the history of the MFA and the HU–MFA Expedition from the Boston perspective. The collection is made up almost entirely of letters from Ashton Sanborn to his mother, Lillian, and his wife Agnes, as well as their responses, along with newspaper and journal clippings from this same period. The collection presents a unique window into the history of archaeology, early 20th century Egyptian politics, international relations, and the many exciting discoveries of the Harvard University–Boston Museum of Fine Arts Expedition. It is the goal of this research project to gather and organize the collection, to study these newly discovered papers, and to process them in a filemaker database, enabling the preparation of scholarly and popular publications—and perhaps even museum exhibitions—as a result.

Quine's Edits

Theresa Clark
Philosophy, 2018

Houghton Library

Advisors: Emilie Hardman, MA, MLS; Heather Cole, MA, MLIS

This summer, I was lucky to work with the archive of W.V. Quine at Houghton Library through the Houghton–SHARP Fellowship. W.V. Quine is one of the foremost philosophers of the 20th century; he spent several decades at Harvard, from his time as a graduate student to his passing in 2000. Broadly, the goal of my research was to use the archive to gain a fuller picture of the development of Quine's philosophical commitments, with a specific focus on understanding the rationale behind Quine's edits in some of his famous papers. More specifically, the process of my research fell into three interrelated categories.

First, using Versioning Machine, a textual display interface, and Text Encoding Initiative (TEI), I encoded several of Quine's papers, including "Epistemology Naturalized" and "Natural Kinds." The encodings of these documents, which are to be published online, will be available to researchers for the purposes of comparing different iterations of Quine's papers, thereby tracking the development of his philosophical ideas. A critical compilation of Quine's edits has yet to be published; therefore, this encoding offers a unique opportunity for scholars to directly compare drafts of Quine's works.

Second, the encoding of these drafts has allowed me to evaluate the significance of Quine's edits—from single word changes to the insertion or deletion of entire paragraphs. Using Quine's notes and his correspondence with other philosophers—housed both at Houghton and at the Harvard University Archives—I was able to match parts of Quine's editorial changes to the rationales described in his notes and correspondences I mapped out the evolution of his papers, thereby gaining insight into the development of Quine's ideas over time.

Finally, I have focused on exploring points of confusion regarding Quine's writings, by comparing Quine's drafts with later philosophical analyses of his work. Quine's writings, and in particular his seemingly singular views, have led to difficulties in interpreting his position. By comparing points of confusion cited in these later analyses to a compilation of Quine's edits, I hope to gain a better understanding of Quine's philosophical commitments and to resolve some of these points of contention.

The techniques I have uses this summer represent the vanguard of modern humanities research; as Versioning Machine and TEI have thus far been used mainly for older text and poetry, I hope that my work will encourage their use in a broader range of humanistic disciplines, and enable philosophical writing to be accessible to a variety of audiences.

Edward Sheldon, George Pierce Baker, and the Origins of the American Theatre at Harvard

Thomas Peterson
History and Literature & Theater, Dance, and Media, 2018

Houghton Library

Advisors: Emilie Hardman, MA, MLS; Heather Cole, MA, MLIS

As one of SHARP's Houghton Library fellows, I am pursuing an independent research project in the library, which holds the Harvard Theatre Collection among its many treasures. I began my project with an interest in investigating Harvard's theatrical history, both curricular and extracurricular, with the goal of producing both an academic essay and a play inspired by my findings. The logical starting point for this endeavor was with the papers of George Pierce Baker, the first Harvard professor to offer practical courses in theatre. In examining Baker's papers, which take up eighty-two boxes in Houghton's collection, I discovered that he maintained extensive correspondence with a remarkable range of colleagues and students.

The first name to catch my eye among his archived correspondents was that of Eugene O'Neill, the Nobel laureate and four time Pulitzer Prize-winning playwright who is widely regarded as the first great American writer of plays and one of the inventors of a truly American drama. The folder containing O'Neill's correspondence with Baker begins with O'Neill's request to join Baker's playwriting class, entitled English 47. O'Neill had received no prior theatre training, but only a few years after his graduation from Baker's class he had garnered mainstream success. In examining Baker's papers, I have discovered that Baker's English 47 class, which began at Radcliffe in 1904 and was introduced at Harvard the following year, trained many in the first generation of notable American playwrights—a group instrumental in introducing the gritty realism and social consciousness that would define the American stage in the twentieth century.

One name that consistently appears in Baker's papers is that of Edward Sheldon, an alumnus of Baker's first English 47 class at Harvard. Intrigued by his early success as a playwright, I delved into Sheldon's papers, which are also housed in Houghton Library, and found that he lived an extraordinary life and left a remarkable artistic and social legacy despite becoming totally paralyzed and blind by his mid-thirties due to a rare form of arthritis. I realized that Sheldon's life would provide fascinating dramatic material, so I have begun the process of writing a play inspired by the life of Edward Sheldon. In this process, I am utilizing not only Sheldon's correspondence with friends ranging from playwright Thornton Wilder to the philosopher Krishnamurti, but also his personal papers, his library, and the materials compiled by his biographer. As Sheldon played an essential role in blazing the trail for American realism, influencing many of his successors in Baker's class, this artistic endeavor goes hand in hand with the academic paper I will produce. The paper will focus on Baker, Sheldon, and O'Neill in explaining Harvard's important role in the development of the American theatre.

Animation and Realism

Zachary Wong
Visual and Environmental Studies, 2017

Harvard University

Advisor: Ruth Lingford, MA

For Ruth Lingford's film *A Girl in Question*, I have been tasked with crafting several key scenes through the time-honored technique of hand-drawn animation. This process of producing hundreds of drawings to create an illusion of continuous movement is as much constructive as it is analytical. Hand-drawn animation forces the animator to analyze and internalize the appearance and physicality of real-world movement: gravity, inertia, acceleration, etc. It demands a hyperawareness from the animator, as each individual frame is both a continuation of its predecessor and anticipation of the drawings to come.

Hand-drawn animation's deep connection to timing and real-world physicality prompted the fundamental question of my research for Professor Lingford. While working on this project, I wondered if animation's connection to live-action at the fundamental level of movement could be extrapolated to the broader components of an animated film: composition, design, and editing. Professor Lingford encouraged me to ask myself if the cinematic language and aesthetics of live-action films could be translated into an animated film, and if so, what its impact would be on a viewer.

Aside from my daily animation responsibilities, I research the films of Roman Polanski, Stanley Kubrick, and Alfred Hitchcock, specifically those that deal with psychological horror. I have divided my time between animating at my desk and viewing live-action films at the Sever Film Studies Library. What I am beginning to realize is that these films seem to suppress the naturalism of their subjects in an attempt to create an atmosphere of fantastic horror. The symmetry of John Alcott's cinematography in Stanley Kubrick's *The Shining* upholds a visually stunning geometric order that seems to overwhelm the psychological chaos of the film's characters. Robert Burks' use of green light in Hitchcock's *Vertigo* seems to anticipate not the lighting logic of naturalistic photography but rather the fantastical color palette of an animated film or Mary Blair color study. The interplay between live-action film and fantastical aesthetics allows the films to heighten the abnormality of the filmed characters' circumstances.
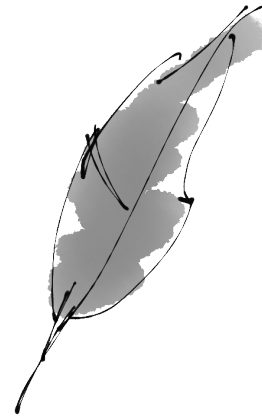
These examples of live-action filmmaking emulating fantasy qualify my original research question. If hand-drawn animation is inherently unrealistic, in what ways can it be influenced by live-action horror, a medium that in many instances seeks to suppress its own realism? An answer to this question would arise in the form of a fully animated scenes that have been inspired by live-action cinema. I hope that through my exploration of live-action filmmaking and hand-drawn animation, I can uncover new ways to extend the verisimilitude of an animated film.
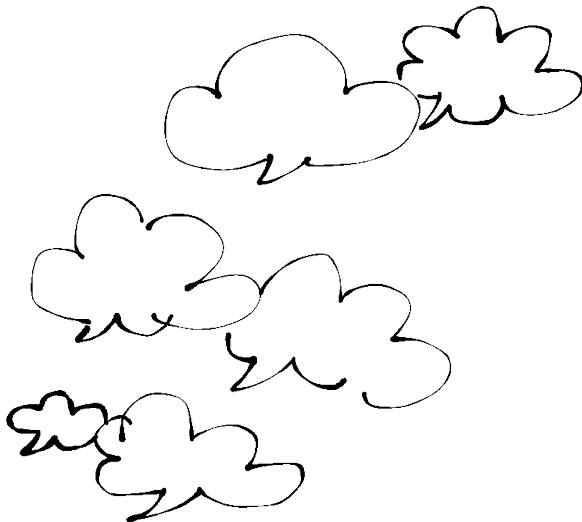
SURGH

Understanding Attachment Styles of Children of Female Former Child Soldiers and their Mothers in Sierra Leone

Gita Bhattacharya
Cognitive Neuroscience and Evolutionary Psychology, December 2016

Harvard T.H. Chan School of Public Health

Advisor: Theresa Betancourt, ScD

Mentors: Katharine Collet, MPH; Elyse Thulin, MSc; Tia McGill Rogers, PhD

During the Sierra Leone Civil War (1991–2002), nearly 20,000 children were forced into militia groups such as the Revolutionary United Front (RUF) to oppose the existing leadership's corruption. These children were often forced to commit violence towards their community members as initiation into the RUF, and were also victims and observers of violence. Following the Civil War, thousands of children were displaced, and the question of their reintegration into society remained. To investigate the negative health consequences of war on these former child soldiers, the factors contributing to their mental health outcomes, and protective processes and forms of psychological resilience, the Research Program on Children and Global Adversity (RPCGA) began a longitudinal investigation in 2002 with data collection points in 2002, 2004, and 2008. A mixed methods approach was applied over the course of the study to first gain a qualitative understanding of how former child soldiers were behaving, feeling, and socially interacting, that would inform quantitative assessments of mental health outcomes. Two hundred sixty male and female former child soldiers between the ages of ten and seventeen at baseline were recruited from an NGO-run Interim Care Center in Kono District. Results from the first waves of the study show that both war experiences and post-conflict risk factors—the stigma associated with being part of the RUF, post-conflict family abuse, and the death of a parent—contributed to long-term mental health effects of former child soldiers. Post-conflict protective factors such as family reintegration mitigated negative mental health outcomes.

My current project involves the fourth wave of the RPCGA's longitudinal study, in which I am examining the intergenerational effects of war on the female former child soldiers who are now becoming mothers and forming intimate partnerships. Continuing the mixed methods approach, my current task is to clean and analyze key informant interviews that discuss the attachment styles of children of female former child soldiers. Initial read-throughs of the interviews reveal certain themes. For instance, children raised by non-biological parents following the war exhibit more antisocial behaviours compared to children who could reintegrate with their biological families post-war. According to initial read-throughs of some twelve key informant interviews, adoptive family and community members were more likely to abuse these foster children compared to their own biological children. Grounded Theory and Thematic Content Analysis will likely reveal additional reasons for this difference in parental treatment, as well as further trends in attachment style between the children of female former child soldiers and their mothers.

HIV Prevention and Education for Boston-Area Adolescents

Henna Hundal
Undeclared, 2019

Center for AIDS Research, Harvard University

Advisor: Wanda A. Allen

Mentors: Sannisha Dale, PhD, EdM; Valerie Earnshaw, PhD

Few epidemics in the U.S. have carried more stigma and incited more public fear, but also inspired more community activism, than HIV/AIDS. When the epidemic became one of the most prominent public health concerns in the 1980s, posters urging people to "Put a rubber on it!" and outreach efforts underscoring the disease's fatality flooded the American consciousness. Fast-forwarding more than three decades later, the success of HIV combination therapy in reducing AIDS-related deaths has rendered HIV/AIDS a chronic, manageable condition. Nevertheless, ample challenges continue to accompany life with the virus, including stigma, access to and affordability of healthcare, and social determinants of health. Moreover, knowing that STIs increase the likelihood of contracting HIV, and given that STI rates are on the rise in the U.S., targeted HIV outreach efforts are still much in order.

Our research aims to revise and re-implement a community-based HIV education and prevention model to better address the changing needs in HIV/AIDS awareness in the U.S. The model, Safety Net, was last updated in 1992 and bears resem-

blance to the "Tupperware model" of direct selling. The original Safety Net called for dynamic, engaged social gatherings at a community member's residence, where Safety Net educators could facilitate discussions and provide resources meant to invite individuals to take charge of their own health. The program, though launched with a focus on women, was later expanded to include couples, mixed-gender, and adolescent sessions.

Our updated Safety Net, designed primarily for adolescents aged fifteen to twenty living in the Dorchester, Roxbury, and Mattapan neighborhoods of Boston, seeks to retain critical aspects of its predecessor, including the emphasis on peer-to-peer advocacy, safe spaces, self-efficacy, and motivation to advocate for one's health. However, our updated Safety Net differs in several key areas, including accents on new prevention strategies, expanded testing options, structural and institutional players in healthcare, and greater integration of HIV education with other STI information so as to reduce stigma. The aforementioned amendments were selected based upon interviews with key informants and local health personnel as well as probes into research literature available on the subject. To this end, our updated model intends to utilize technology for the streamlined delivery of information as well as welcome active user participation to drive the sessions. We will facilitate two installments of the updated model and conduct focus groups, gauging effectiveness in meeting the revised criteria via a pretest and posttest assessment.

Driving Healthcare Innovation Through Building New Technologies and Deploying a Comprehensive Digital Health Platform

Isaac Hung
Chemical and
Physical Biology,
2018

Boston Children's
Hospital

Advisor: John
Brownstein, PhD

Mentor: Matthew
Murphy

Across the healthcare industry, there's been a growing interest in developing digital healthcare solutions for clinical, research, patient, and general consumer use. In 2015, venture capital funding for digital health companies surpassed a staggering $4.3 billion in the US.

I work with Boston Children's Hospital's Innovation & Digital Health Accelerator (IDHA), whose mission is to shape the future of healthcare by building new technologies and deploying a comprehensive digital health platform. IDHA identifies and vets high-priority health technology innovations; then, it provides the resources, funding, and momentum to accelerate development and commercialization. Of over fifty digital innovations from clinicians and researchers that the Accelerator has vetted, I am involved with multiple high potential projects.

As part of the Accelerator team, my work involves using a research-first approach to digital technology development. The method emphasizes the importance of understanding the pain point or unmet need the technology is meant to address before jumping to technology deployment. For each of the projects that I am supporting, I take into account current clinical, workflow, or process pain points as well as major business, regulatory, or clinical requirements. These considerations are critical to a project's development and serve as a starting point for further research.

In the continuum of technology development, the majority of my time is spent in the research and design phase, which will then be followed by clinical research validation to show health care impacts, or an efficiency study to alter processes and workflows. Furthermore, through extensive clinical interviews and market analysis, I am able to understand and justify a product's business value which, in health care, often comes from improving care outcomes and reducing costs. Doing so will ensure that these innovative ideas become effective solutions, widely adopted by patients and clinicians. We hope to extend the reach and impact of these healthcare solutions beyond Boston Children's Hospital.

Investigating the Role of FtsQ in Regulating Cell Division in Mycobacterium smegmatis

Jenna Zhang
Molecular and
Cellular Biology,
2017

Harvard T.H. Chan
School of Public
Health

Advisor: Eric Rubin,
MD, PhD

Mentor: Cara
Boutte, PhD

Tuberculosis (TB) is caused by the bacterium *Mycobacterium tuberculosis* (*Mtb*), and is one of the deadliest infectious diseases still present in every part of the world. The disease is passed from person to person through the air; once infected with the bacteria, individuals have a 10% lifetime risk of falling ill with TB. However, this risk increases in people with compromised immune systems such as individuals living with HIV, malnutrition, diabetes, and other long term illnesses. This leaves developing countries particularly at risk, and over 95% of cases and deaths occur in developing countries in Southeast Asia and

Africa. About one-third of the global population has latent TB. This means that they have been infected by *Mtb* but are not yet ill and cannot transmit the disease because *Mtb* cell division is a prerequisite to infection. The importance of cell division for infection makes the study of cell division processes and their regulation in *Mtb* vital for stopping the spread of TB. As such, understanding the molecular mechanisms of cell division during adaptation to infection stress is important. Many of the factors mediating cell division are largely conserved from the model organism *E. coli* to *Mtb*, although regulation of these factors is different between the two.

FtsQ is an essential cell division protein in *E. coli* that associates with other key cell division factors in a complex known as the divisome that helps link together major events of cell division. FtsQ has been shown to be phosphorylated in *Mtb*, but has otherwise not been extensively studied in mycobacteria. Phosphorylation can modulate growth either by directly controlling the level of protein activity or by affecting the interactions between proteins in a complex, and previous literature on divisome players suggest a possible role for FtsQ in growth regulation under oxidative stress. Using *Mycobacterium smegmatis*, a model organism for *Mtb*, and based on studies in *E. coli*, my work this summer uses biochemical and genetic approaches to investigate the hypothesis that FtsQ plays an important role in the division complex to regulate cell division and that phosphorylation of FtsQ helps integrate environmental stress signals into the regulation of the division complex. The techniques I am implementing include timelapse microscopy with fluorescently-tagged FtsQ, affinity pulldowns to identify interactor partners of FtsQ, and growth and kill curves to identify growth defects in mutant strains expressing constitutively phosphorylated and unphosphorylated FtsQ. In doing so, I hope to clarify the role of FtsQ in cell division and the ways in which phosphorylation affects and regulates FtsQ activity.

Understanding HIV Transmission Patterns in Botswana: Using Evolutionary Analytics to Improve Global Health

Jonathan You
Molecular and Cellular Biology, 2018

Harvard T.H. Chan School of Public Health

Advisor: Max Essex, DVM, PhD

Mentor: Vladimir Novitsky, MD, PhD

Since its emergence in 1981, HIV has infected 78 million people and claimed 39 million lives worldwide. The deadly virus, first clinically observed in Los Angeles, began ravaging Sub-Saharan Africa in the 1990s. By 2000, an unprecedented 35% of adults were infected in Botswana, a middle-income country in southern Africa.

Today, Sub-Saharan Africa remains the region most affected by HIV. No cure or vaccine has yet been found for the disease. However, global health experts, in collaboration with national ministries of health, have begun to tame the HIV epidemic with nationwide programs for affordable antiretroviral treatment (ART), a safe and effective means of restoring normalcy to the lives of HIV-infected patients.

The Essex lab is a key partner in one of these ART programs. Dr. Max Essex, Chair of the Harvard AIDS Initiative and the Botswana Harvard AIDS Initiative Partnership, has worked in Botswana to conquer HIV since 1996. Botswana, once the most HIV-affected country in the world, has adopted an extremely successful national ART program since 2002—so successful, in fact, that Botswana now leads Sub-Saharan African countries in proportional ART coverage of HIV-infected individuals.

My research expands upon an exciting foundational step of the fight against HIV/AIDS in Botswana. To establish a baseline and monitor progress of HIV prevention interventions, researchers in Botswana, including those from the Essex lab, have surveyed thousands of households over several years to determine rates of HIV incidence (new diagnoses) and prevalence (total cases), identify new infections for follow-up care, and collect blood samples from those infected.

Like all viruses, HIV-1C, the HIV subtype most prevalent in Botswana, contains a set of genes that encode instructions for building and maintaining itself. These viral genes can be sequenced from HIV-

infected blood samples collected in the household surveys above. HIV-1C evolves so rapidly that the spread of the virus through a population can be tracked by examining the accumulation of changes in its genes in HIV-infected individuals over time. The greater the extent of evolutionary relatedness between two samples of HIV captured from distinct individuals, the greater the likelihood that the two individuals were infected within the same transmission chain, a string of related viral infections.

My project analyzes genetic sequencing data from a household survey of hundreds of HIV-infected individuals in Mochudi, a village near the capital city of Botswana. By employing biostatistical and phylogenetic analyses, my research investigates the evolutionary dynamics of locally circulated HIV-1C lineages. By cross-referencing these patterns with demographic data collected in the same survey, my project can help identify vulnerable subpopulations, establish factors that affect HIV transmission, and, perhaps, inform public health interventions. For instance, treatment-as-prevention—calming active infections with ART to reduce risk of further transmission—is considered a front-line intervention in resource-limited areas, but this method is best suited for regions where HIV transmissions generally occur within an isolated set of ART-treated communities. If transmissions more frequently occur between ART-treated and ART-untreated populations, then a more comprehensive, prophylaxis-based approach may be required.

The Effect of the Host Red Blood Cell on the Malaria Parasite

Julianna Lee
Molecular and
Cellular Biology,
2019

Harvard T.H. Chan
School of Public
Health

Advisor: Manoj
Duraisingh, PhD

Mentors: Martha
Ann Clark, PhD;
Brendan Elsworth,
PhD

Malaria remains a significant life-threatening disease, affecting over two hundred million individuals yearly. It is caused by the *Plasmodium* parasite, which is transmitted from infected to uninfected individuals by mosquitoes. The blood stage of a malaria infection is responsible for all symptoms of disease. Previous studies have revealed that naturally occurring red blood cell variants in the human population such as the sickle cell trait affect the behavior of the malaria parasite. These variants change the parasite's ability to propagate and cause disease in the host, either by blocking invasion or growth within the red cell and may ultimately provide protection from the malaria infection.

Here we aim to investigate the impact of red blood cell physiology on the parasite, in particular *Plasmodium knowlesi*, one of the five species that cause human disease. Specifically, I have been working to understand why *P. knowlesi* propagation is reduced in cord blood red cells containing fetal hemoglobin (HbF) and in oxidized adult red cells. Other *in vitro* studies have shown that when *Plasmodium falciparum* is cultured in cord blood, the levels of invasion increases; however, growth is retarded. By the end of a baby's first year of life, the HbF levels drop to a few percent compared to 80% at the time of birth. For normal adult red cells it is thought that oxidative damage to the host changes the erythrocyte cytoskeleton and thus affects the cell membrane and alters normal cell mechanical function in relation to the parasite.

I have used microscopy and flow cytometry to assess *P. knowlesi* proliferation in the red cells I've been studying. Thus far I have observed that *P. knowlesi* proliferation is reduced in both cord and oxidized red cells. To identify the specific parasite factors responsible for reduced proliferation in cord blood, I have been attempting to adapt *P. knowlesi* to cord blood by continuously culturing *P. knowlesi* in cord blood. When I observe normal proliferation of *P. knowlesi* in cord blood, I will sequence the adapted parasites to identify genetic mutations responsible for normal proliferation. I have additionally been employing an osmotic fragility assay to investigate the hypothesis that *P. knowlesi* osmotic stability may be compromised in cord and oxidized red cells. Finally, to test our hypothesis that reduced growth may be caused by premature lysis of the *Plasmodium*-infected red blood cells (RBC) as a mechanism of host mediated protection from the malaria infection I have been working with the parasites' genetics. I can use NanoLuc, a reporter protein, to help monitor RBC lysis as described in a previous assay.

As malaria continues to place a large burden on public health, studies like these can help reach the potential of revealing novel mechanisms of red cell mediated protection against malaria infection. Such findings have the potential to ultimately help reach groundbreaking therapeutic interventions.

Chaperone Independent Type III Effector Secretion in *Shigella flexneri*

Kristen Fang
Molecular and Cellular Biology, 2019

Massachusetts General Hospital, Harvard Medical School

Advisor: Cammie Faith Lesser, MD, PhD

Mentors: Nadja Ernst

Shigella flexneri, a Gram-negative rod-shaped bacterium and the causative agent of bacillary dysentery or shigellosis, affects an estimated 90 million annually worldwide, with at least 100,000 fatal cases. *S. flexneri* contains a virulence plasmid (>200kb) that codes for a Type III secretion system (T3SS), a complex, syringe−like nanomachine that delivers effector proteins into the cytoplasm of human gastrointestinal cells. There are 25–30 known effector substrates of this system that together enable *S. flexneri* to manipulate host cellular processes over the course of an infection. There is currently no vaccine available to prevent shigellosis, and little is known about the secretion pathways of T3S effectors. Increased understanding of the T3SS in *S. flexneri* is a step forward in the process of developing a drug that targets the secretion system not only in *Shigella*, but in other T3SS species including *Salmonella* (typhoid fever), *Yersinia* (plague), and *Chlamydia* (sexually transmitted disease).

Past research suggests that T3S effectors are defined by an N-terminal secretion sequence. Some effectors also have a chaperone-binding domain (CBD) that is recognized by a T3S chaperone that likely plays a role in the hierarchy of effector delivery. More than 50% of *S. flexneri* effectors do not bind to a known chaperone. For these putative chaperone-independent (CI) effectors, it is unknown what defines them as T3S secreted substrates other than their N-terminal secretion sequences. Our project focuses on the secretion pathway of chaperone-independent T3S effectors. We took two strategies: first, we investigated if secretion of putative CI effectors is dependent on known chaperones; and second, we investigated whether any of the annotated virulence-plasmid encoded genes (open reading frames, ORFs) of unknown function act as a previously undiscovered chaperone.

S. flexneri strains, each of which carry a chaperone knockout ($\Delta ipgA$, $\Delta ipgE$, and $\Delta spa15$) were tested in a high-throughput assay for secretion of CI effectors. There was no significant difference in secreted levels, suggesting that the secretion of these effectors is independent of the known chaperones. To screen for any potential undiscovered chaperones, we took two approaches. First, we examined CI-effector secretion in the Minimal Type Secretion System (minT3SS) that lacks all the unknown ORFs. The minT3SS, coded for by the entry region (~31kb) on the virulence plasmid, only consists of the Type III Secretion Apparatus known chaperones, and a few effector and regulatory proteins. The entry region of the virulence plasmid was isolated onto a plasmid referred to as pSER2 for Shigella Entry Region 2. The introduction of this plasmid into *Escherichia coli* results in T3S competent *E. coli* (mT3 *E. coli*). Secondly, we obtained a collection of twenty-five *S. flexneri* strains, each with an unknown ORF removed. We ran high-throughput secretion assays on the mT3 *E. coli* strain and the *S. flexneri* ORF knockout strains. In both approaches, there was no significant difference in CI effector secretion, suggesting that there are no undiscovered chaperones. Together, these two strategies add evidence to the existence of T3S chaperone-independent effectors.

Early Child Development in Rwanda

Kristy Hong
East Asian Studies, 2017

Harvard T.H. Chan School of Public Health

Advisor: Theresa Betancourt, ScD

Mentor: Rose Wilder, MA

Early childhood development (ECD) interventions can significantly impact the lives of vulnerable children through enhancing family environments and improving nutrition- and health-related practices. The Government of Rwanda (GOR) has identified ECD services as a key pillar of the country's flagship social protection strategy, Vision 2020 Umurenge (VUP); this mandate falls under the Ministry of Local Government (MINALOC). The overall social protection policy focuses on complementing income support with cross-sectoral collaboration to break intergenerational cycles of poverty and invest in Rwanda's future growth and prosperity. The recently revised and approved ECD policy in May 2016 also states cross sectoral collaboration to ensure Rwandan children reach their full potential by focusing on the crucial years of childhood development; this policy is implemented by the Ministry of Gender and Family Promotion (MIGEPROF).

Dr. Betancourt's team has developed the Family Strengthening Intervention for Early Childhood Development (FSI-ECD), a home-visiting, parenting intervention aimed at families with young children

living in extreme poverty. The program consists of a community-based volunteer "coach" visiting a household to engage the family in a series of fifteen structured modules. These include modeling positive parenting, presenting alternatives to harsh punishment against children, promoting the involvement of fathers, and providing prescriptive knowledge about brain development, nutrition, health, and hygiene.

The team used mixed methods—qualitative and quantitative data collection—to test the program's effectiveness in select districts in Rwanda with bachelor-level home-visiting "coaches." Currently, the team is gearing up for a full scale trial to test the added value of the FSI-ECD home-visiting program with VUP public works services by evaluating both relative effectiveness and cost effectiveness. With an array of tested metrics, they plan to measure the added value of the FSI-ECD model combined with Rwanda's flagship VUP public works program compared with the public works program alone. The team is currently adapting the original manual for use by lower literacy community based volunteers in order to scale the program.

By working with the GOR and development partners, the data have the ability to influence future policy that emphasizes the importance and relative cost of incorporating ECD services in social protection policies.

Improving HIV Prevention and Sexual Health Education for Adolescents

Laura Kanji
Undeclared, 2019

Center for AIDS Research, Harvard University

Advisor: Wanda A. Allen

Mentors: Sannisha Dale, PhD, EdM; Valerie Earnshaw, PhD

Over the past two and a half decades, the HIV/AIDS landscape has changed dramatically. An HIV diagnosis is no longer a death sentence; instead, with the rise of highly active antiretroviral therapy, HIV has become a chronic and manageable condition. However, HIV is not yet a disease of the past. It still affects people of color and LGBTQ people disproportionately, and these disparities have not changed over the past thirty-five years. While HIV medication is highly effective when taken as prescribed, it can have serious side effects, and systemic issues surrounding the health care system—including poverty, education, and race—pose obstacles for adherence. Finally, there is stigma attached to sexually transmitted infections, particularly HIV, that proves detrimental to

prevention, treatment, and testing efforts. However, as the average prognosis of those living with HIV has improved, sexual health education has declined. It is frequently inadequate, outdated, or completely nonexistent. Currently, there is a chlamydia epidemic among teens in Boston, and rates of sexually transmitted infections have been rising nationally over the past decade. The increased STI rates among adolescents coupled with the disparities and stigma that still surround HIV point to a need for comprehensive sexual health education that incorporates recent changes in HIV/STI treatment, prevention, and attitudes, particularly for underserved youth.

In response to these issues, our research aims to develop and evaluate a modernized, relatively short HIV and STI prevention intervention for Boston adolescents. Among other recent changes, the intervention addresses the development of rapid STI testing, HIV prevention medication, and prevailing societal myths, new and old, about HIV and STIs. Additionally, it aims to address mistaken perceptions that HIV is a disease of the past, is no longer worth preventing, or that certain groups are destined to contract HIV. We developed our intervention after conversations with key stakeholders in the Boston area, including public health officials, educators, community health workers, and individuals living with HIV. We also conducted a literature review on theories of health behavior change. Our intervention is based on the Social Cognitive and Information-Motivation-Behavior theories, which stress the importance of increasing a sense of self-efficacy and motivation among participants in health education. In the next few weeks, we will carry out our intervention with groups of Boston adolescents and use pretests and posttests along with focus groups to analyze its effectiveness. This evaluation will hopefully reveal whether our educational methods were effective in improving participants' knowledge, confidence in their ability to make and carry out decisions related to sexual health, and attitudes towards STI testing and HIV, and will inform future prevention efforts.

Purification of WD40

Michael Kennedy-Yoon
Visual and Environmental Studies, 2017

Harvard T.H. Chan School of Public Health

Advisor: Dyann Wirth, PhD

Mentor: Allison Demas

Artemisinins are a category of antimalarial drugs commonly thought of as frontline treatment for the disease. Their use is widespread, and generally effective at counteracting malaria. It is therefore troubling that there is an observed resistance to artemisinins in several strains of the plasmodium parasite more commonly known as malaria, and specifically in *Plasmodium falciparum*, the deadliest of the five most common subspecies. It is this concern that motivates our study of several factors that may confer immunity to artemisinins, including several single nucleotide polymorphisms (SNPs), which are one-base-pair mutations in the DNA code of an organism, in the WD40 region of the Pf Coronin protein. Coronin is thought to be an actin binding protein, and WD40 is a large region within coronin that is dubbed a "propeller region." It is our goal to synthesize and purify this WD40 region, in both wild type and mutant forms, in order to better study and understand how these SNPs may confer artemisinin resistance.

We began by creating a bacterial plasmid that contained the WD40 protein code. A plasmid is a piece of circular DNA, specific to prokaryotes, into which we can add genes to be expressed by the prokaryotes. We then inserted the plasmid into *E. coli*, in order to facilitate protein expression of the wild-type WD40. Most recently, we have attempted to induce protein expression via IPTG induction. Our current efforts are centered around the verification that protein expression has occurred. To this end, we are running gels of samples of cells taken at different times during the induction process, which will allow us to determine whether or not the protein was successfully expressed. If this process results in the affirmation of protein presence, we will attempt to purify the protein using a nickel purification column. In the protein there is a negatively charged histidine tag that will hypothetically allow for the selective extraction of the expressed protein using the positively charged nickel ions.

Given that all of these measures are successful, we will proceed to expression of the mutant WD40 proteins. The procedure here is largely the same, with transformation of the plasmid, induced expression via IPTG, and nickel column purification. After the purified protein is obtained, we will be able to engage it in further study to analyze functionality. If all goes well, by the end of this process, we will better understand if, and how, changes in WD40 confer resistance, and will be able to recommend courses of action in cases of artemisinin resistance.

Quantifying Public Sentiment of Healthcare Sector Organizations

Patrick Tan
Statistics, 2018

Boston Children's Hospital

Advisor: Carl Brinton, MBA

Since the dawn of the digital age, both the rate of information production and information collection have grown exponentially, giving organizations and individuals access to vast amounts of data. Within this context, there has been a corresponding need to develop efficient methods to systematically parse and summarize these large quantities of data. One technique of interest is the concept of "sentiment analysis," which involves the use of natural language processing and machine learning to extract qualitative information from a body of text. This technique has proved effective in a wide variety of applications, ranging from the analysis of social media posts to classifying consumer reviews. Within the healthcare sphere, where the focus is primarily on healthcare sector enterprises and startups, there is a particular interest in determining the general sentiment surrounding these healthcare organizations. Armed with this information, a host of potential interest groups in the healthcare sphere (investors, partners, customers, suppliers, etc.) would be able to make better decisions with regards to these healthcare sector organizations (which ones to invest in, which ones to support, etc.) Thus, the goal of this project is to develop an efficient, quantifiable metric to summarize this sentiment. To accomplish this, we train a machine learning algorithm to recognize positive and negative sentiment surrounding a healthcare organization of interest within news articles about the organization. We gather news articles that have preceded significant movements in the stock price as training examples, since these news articles are likely to contain the strongest positive or negative sentiments. Finally, the algorithm then takes in the wealth of recent news about a certain company and outputs a "sentiment score," which represents our desired quantifiable metric of general sentiment surrounding the organization.

Quality Improvement of Congenital Heart Surgery in Developing Countries

Sarah Rahman
Sociology, 2018

Boston Children's Hospital

Advisor: Kathy Jenkins, MD, MPH

Mentors: Kate Doherty; Jennifer Kupiec, MPH

The International Quality Improvement Collaborative for Congenital Heart Surgery in Developing World Countries (IQIC) collects surgical outcome data from forty-seven sites in twenty-two countries worldwide. Congenital heart surgery in resource-deprived areas is difficult and has high mortality rates. IQIC aims to improve surgery practices in these regions through collaboration with local sites. IQIC determined three key drivers to reduce mortality: safe perioperative practice, infection reduction, and team-based practice (nurse empowerment). I worked in many facets of IQIC, with the overall mission to implement these drivers in the participating sites and improve surgical outcomes. I prepared for monthly teaching webinars focusing on these key topics that are broadcasted to the sites and discussed among surgeons, nurses, and staff. In addition to webinars, IQIC connects with partner sites to develop regional learning sessions, which nearby sites can attend in person, participating in presentations, simulation sessions, and panel discussions on quality improvement.

IQIC conducts annual site visits to meet with the local teams, perform data source verification, and to reinforce IQIC goals. Each year, annual benchmarking reports are provided to sites, detailing their outcomes and comparisons to the overall collaborative. I contributed to these reports through describing each site visit and organizing data, which sites can then use to understand their outcomes and areas for improvement. In addition to the surgery dataset, I am assisting with new IQIC initiatives: creating a new database on cardiac catheterization aimed to gather information on a procedure that has not been well documented in developing countries, as well as developing a pilot infrastructure assessment survey designed to evaluate the equipment and physical spaces available for surgery at the sites.

Furthermore, I am working on a global development project with IQIC. I plan to link IQIC congenital heart surgery outcome data to World Bank global development indicators to understand the role of congenital heart disease in the global health agenda. The global health landscape has been changing: as poverty and infectious disease decrease, non-communicable diseases such as congenital heart disease in developing countries are accounting for a greater percentage of deaths in childhood and must be addressed in health initiatives. By comparing development indexes to the outcome data of IQIC, I seek to discover the relationship between development stage and surgical outcomes. This can help tailor quality improvement and investment strategies for the sites.

PRISE

Astronomy and Astrophysics
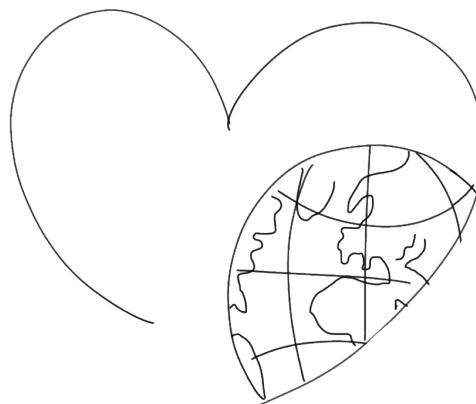
Kinematic Modeling of Oxygen-Rich Supernova Remnant N132D

Charles Law
Physics, 2017

Harvard-Smithsonian
Center for
Astrophysics

Advisor: Dan
Milisavljevic, PhD

A supernova remnant (SNR) is the resulting structure left behind after the explosion of a star in a supernova explosion. SNRs are bounded by an expanding shock wave and are composed of ejected material expanding from the explosion as well as interstellar material swept up in the process. A small subset of young SNRs, known as "oxygen-rich" SNRs, in which interior fragments of the supernova progenitor star are exposed to direct investigation, allow for an opportunity to make direct tests of stellar evolution and nucleosynthesis models of massive stars. We have used the Inamori Magellan Areal Camera and Spectrograph (IMACS) on the 6.5 m Magellan telescope to map the [O III] $\lambda\lambda$4959, 5007 Å, emission-line dynamics of the oxygen-rich SNR N132D in the Large Magellanic Cloud, a nearby satellite galaxy of the Milky Way. From the acquired spectra, we were able to reconstruct an interactive, 3D kinematic map of N132D, which allows for a detailed study of the remnant's internal structure, particularly ejected material in the form of oxygen-rich filaments and knots. While previous attempts at reconstructing the 3D structure of N132D have been hampered by insufficient data quality, our observations allow for a detailed study of the remnant's kinematics. Since the trajectory of ejected material in a supernova explosion is ballistic in nature, the spatial and kinematic configuration of heavy-element ejecta in a particular SNR provides a sensitive measure of its age. Thus, in addition to probing explosion dynamics and nucleosynthetic yields, an improved understanding of the inner structure of N132D will allow us to estimate a more accurate age for the SNR, a quantity that has, to this point, remained somewhat uncertain.

Searching for Earth-Like Exoplanets using the Radial Velocity Method

Chris Dolliff
Chemistry and
Physics, 2018

Harvard-Smithsonian
Center for
Astrophysics

Advisor: David
Phillips, PhD

The first confirmed exoplanet observation was recorded in 1988; since then there have been 3,443 confirmed observations of exoplanets. There are several ways to observe the presence of an exoplanet orbiting a distant star, but because it is extremely difficult to optically resolve these planets, almost all methods involve measuring the host star's behavior. In a solar system, all bodies are orbiting around the center of mass of the system. This means that a star with an orbiting planet will move with respect to an observer on Earth as it wobbles around the center of mass of its system. This wobble can be detected by measuring the Doppler shift in the star's emission spectrum. By measuring these oscillations, we can detect the presence of a planet and determine its mass. This technique is called the radial velocity method. This method is applicable to many planetary systems. When combined with methods that can only determine the planet's volume (most commonly the transit method) we can determine the planet's general structure (rocky, gas giant, ocean planet, etc.).

The idea of finding exoplanets by measuring radial velocities has been successful for almost thirty years. The method excels at finding heavy, quickly orbiting, and nearby exoplanets. The method struggles, however, to find planets with Earth-like mass and orbital period with the current spectroscopic technology and stellar noise theory. The best radial velocity measurements can detect signals with amplitudes as low as ~1 m/s, which is not good enough to find planets with Earth-like masses and orbits. For that, the theory predicts a ~10 cm/s signal amplitude. My lab has spent the last nine years setting up a high precision laser comb that can be used to calibrate a spectrograph to reach the required level of spectroscopic error. The lab's goal is to use its solar telescope in the Canary Islands to observe the transit of Venus around the Sun by measuring the Sun's radial velocity. We test this technique on the Sun because there are not currently good enough models of the noise in spectroscopic data from Sun-like stars, even with the best calibration, to get 1 cm/s signals from point source stars. Thus, we wish to use our knowledge of the Sun and ability to resolve its surface to minimize

the error in a way that can be applied to other stars.

My role in the project has been to analyze possible sources of error in our measurement of the solar spectra. I used Python to analyze a number of possible sources of error including the mounted guide that keeps our telescope centered on the sun, some strange behavior in images that cameras on either side of the telescope take every five minutes, and high temperature variance inside the telescope. As I do this, part of the goal of the development of these tools is to produce a more robust diagnostic infrastructure to detect malfunctions of the telescope quickly and easily as they happen.

X-Ray Emissions from L, T, and Y Type Brown Dwarfs

Deanna Emery
Physics &
Astrophysics, 2017

Harvard-Smithsonian
Center for
Astrophysics

Advisor: Akos
Bogdan, PhD

Brown dwarfs are substellar objects that are not massive enough for the nuclear fusion of hydrogen, necessary in main sequence stars; however, they are more massive than planets, exceeding at least thirteen times the mass of Jupiter. Brown dwarfs are separated into spectral types M, L, T, and Y, with M being the warmest type, and Y being the coolest. The spectral classes L, T, and Y, specifically, are cool enough that they are only supposed to emit in low energy wavelengths. However, in recent years, these brown dwarfs have been observed emitting in high-energy X-ray flares. The mechanisms that cause these brown dwarfs to flare are not yet understood. For this project, I am going through XMM-Newton and Chandra X-ray point source catalogs of large sky surveys containing 396,910 and 241,074 sources, respectively, in order to determine if any of the observed X-ray sources might be from brown dwarf flares. This is done by using Python code to correlate the coordinates of X-ray emissions with the coordinates of known brown dwarfs, while accounting for the proper motion of the brown dwarfs. I will present the results of this cross-correlation between catalogs and attempt to offer some insight on the emission of X-rays from brown dwarfs in spectral classes L, T, and Y.

Exploring the Dark Matter Content of Nearby Elliptical Galaxies

Oliver Philcox
Astrophysics, 2017
Emmanuel College

Harvard-Smithsonian
Center for
Astrophysics

Advisor: Akos
Bogdan, PhD

Exploring the evolution of galaxies throughout cosmic time is paramount in modern astrophysics. It is believed that galactic development hinges on three major components: the central supermassive black hole with a mass millions of times that of our sun, the stellar body, and the dark matter halo. Dark matter is a hypothetical type of matter that does not interact with normal matter, except through gravity, and there is strong experimental evidence to suggest that it makes up 85% of the total mass of the universe. The specific symbiosis of these three components likely results in the demography of galaxies observed across the Universe, and a study of the dark matter content in nearby elliptical galaxies can shed new light into how they are linked together, which can be used to improve models of galactic evolution. Research of this variety has been attempted before; however, it was on a much smaller scale, considering only seven galaxies rather than thirty, due to the smaller number of images previously available. Thus, its utility was far more limited.
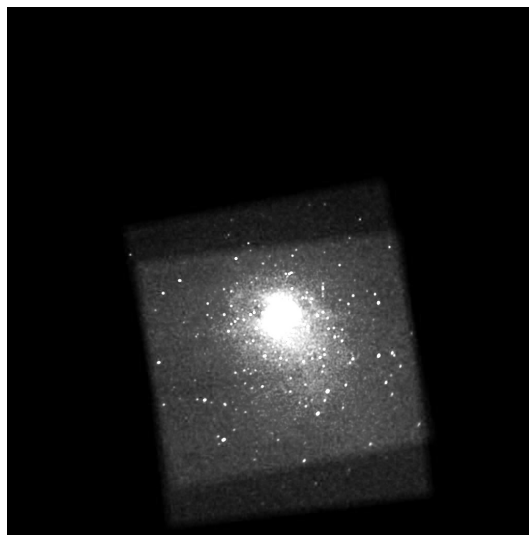

Figure 1: Galaxy M49.

This project has focused primarily on calculating the gravitating mass profiles of a large sample of nearby elliptical galaxies, through use of X-ray images from the Chandra satellite. Launched in 1999, Chandra is the most powerful X-ray telescope made to date, capable of over 100 times greater resolution than its predecessor and currently in orbit 50,000 miles above the Earth. Shown above is an X-ray image of

the galaxy M49, located 56 million light-years from the Earth, and the code used in this study was initially written to analyze this, before being extended to other galaxies.

In order to find the galactic mass profile, the images had to be analyzed using a variety of methods. Firstly, data from solar flares, which occur at periods of drastically increased solar activity, had to be removed, as did the point sources seen across the image. I was then able to extract spectra of the hot gas in the galaxies at a range of radii, and, by fitting established models to this, could create gas density and temperature profiles for the galaxy. These were then used to compute the total mass profile as a function of radius, due to both dark matter and stellar material. Having obtained these data, they could be combined with the readily available measurements of stellar mass and supermassive black hole mass in order to probe whether supermassive black holes exhibit tighter correlations with their stellar or dark matter halo masses. This project will contribute significantly to the development of astrophysical understanding of galactic evolution.

The Design and Implementation of Data Acquisition and Control Systems in the Detection of B-Mode Polarization in the Cosmic Microwave Background

Thomas Culp
Physics, 2019

Harvard-Smithsonian
Center for
Astrophysics

Advisor: John Kovac,
PhD

Mentor: Kirit
Karkare

The cosmic microwave background (CMB) is the oldest light in existence, the relic radiation from the Big Bang that birthed the universe. The relative isotropy of the CMB supports the theory of inflation, an exponential expansion of the universe during the first 10^{-32} seconds of its existence. Inflation predicts the production of gravitational waves, which would leave their trace on the CMB in the form of a B-mode (curl) polarization pattern, which could be detected if inflation took place at energy scales predicted by Grand Unified Theories.

The detection of B-mode polarization depends on instrumentation capable of measuring signals on the scale of nanokelvin, making the detectors which do this work the most sensitive of their kind. BICEP3 and the Keck Array are microwave telescopes which observe CMB radiation at the South Pole in search of this inflationary signature. Ensuring accurate measurement of the CMB requires precise measurement of possible sources of external interference as well as a variety of instrumental effects.

The first part of my research involves developing a data acquisition system to measure radio frequency interference from hand-held radios and satellite communications at the South Pole. I write software to acquire, visualize, and analyze the data from Signal-Hound spectrum analyzers which continuously monitor the radio spectrum from 100 kHz to 12.4 GHz. These devices operate simultaneously with the telescopes to help account for potential external interference. Alongside my work on the data acquisition system, I create a publicly accessible web browser which allows for the download and analysis of these time-stream spectrum data.

The second part of my research includes developing a generalized suite of Linux-based motion control software for positioning stages which are used in various calibration apparatus. The software is general enough to be used for near-field beam mapping, Fourier transform spectroscopy, and controlling a rotating polarized source. These tools allow for the processing of instrumental polarization data which are essential to the integrity of the sensitive measurement of CMB polarization.

The Role of T Cell Phenotype on Endothelial Cell Sprouting and Skeletal Muscle Stem Cell Proliferation and Differentiation

Erica Budina
Engineering Sciences, 2018

Wyss Institute for Biologically Inspired Engineering

Advisor: David J. Mooney, PhD

Mentor: Brian Kwee

Musculoskeletal diseases affect more than six million Americans annually. To improve muscle function following injury or disease, therapies must be developed that promote both angiogenesis (the formation of new blood vessels from pre-existing ones) and muscle regeneration. Muscle regeneration involves the activation of a quiescent population of skeletal muscle stem cells called satellite cells; upon injury, these cells proliferate, migrate to the site of injury, and eventually differentiate and fuse, forming new myofibers.

Previous research in the Mooney laboratory has demonstrated that the combined delivery of growth factors known to promote both angiogenesis and myogenesis has the potential to reverse muscle injury caused by tissue ischemia. However, the ability of these therapies to promote regeneration in ischemic injury has been shown to depend on the presence of immune cells. This research work investigates the role of $CD4^+$ T cells in modulating these processes. In the long term, it seeks to understand whether enhancing the number and type of T cells at sites of ischemia can dually promote angiogenesis and myogenesis. In the short term, this research work explores how the cytokines secreted by mouse T cells differentiated into Th1, Th2, Th17, and Treg phenotypes affect endothelial cell sprouting in aortic ring slices, as well as satellite cell proliferation and differentiation *in vitro*.

To address this research question, $CD4^+$ T cells were isolated from adult mice and differentiated into Th1, Th2, Th17, and Treg phenotypes. After T cell phenotype was verified via FACS, conditioned media was collected from each cell type. To study the role of T cell phenotype on endothelial cell sprouting, T cell-conditioned medias were applied to an aortic ring assay. Aortic ring slices were then fixed, fluorescently stained, and imaged via confocal microscopy. Endothelial cell networks were quantified using Angiotool software; specifically, the average vessel diameter, length, and lacunarity were determined for each condition. To study the role of T cell phenotype on muscle regeneration, T cell-conditioned medias were applied to satellite cells isolated from mouse hind limbs. Proliferation and differentiation assays were performed, followed by immunohistochemistry staining and confocal microscopy. In the proliferation assays, satellite cells were stained for EdU, a nucleotide analog that is incorporated into the DNA when cells are undergoing mitosis. In the differentiation assays, cells were stained for myosin heavy chain, a mature muscle cell differentiation marker. Proliferation was assessed by determining the percentage of EdU-positive cells and by measuring the intensity of the EdU signal. Differentiation was assessed by determining the percentage of myosin heavy chain-positive cells and by calculating the fusion index for each myosin heavy chain-positive cell.

Preliminary results from the aortic ring assays suggest that Th2 conditioned media promotes endothelial cell sprouting, while Th1 conditioned media inhibits sprouting. However, results from the skeletal muscle stem cell proliferation and differentiation assays suggest that Th1-conditioned media enhances skeletal muscle stem cell proliferation and inhibits their differentiation. Overall, these studies suggest that a combination of pro- and anti-inflammatory T cell cytokines may be necessary to dually promote angiogenesis and myogenesis.

Fabrication and Testing of a Novel Therapeutic Dental Material

Jacob Scherba
Engineering Sciences, 2018

Wyss Institute for Biologically Inspired Engineering

Advisor: David J. Mooney, PhD

Mentors: Kyle Vining, DDS; Adam Celiz, PhD

Over one billion dental fillings are performed annually, but presently-used composite resin filling materials are toxic to cells. If a filling is unsuccessful, it could necessitate more painful and more expensive procedures like root canals or extractions. The introduction of a reliable, biocompatible dental material, therefore, would be of great clinical significance. In this study, we aim to develop a material that is not only biocompatible, but also therapeutic, to restore teeth to health. *In situ* curing of bulk monomers results in a polymer that may prove able to induce differentiation of dental pulp stem cells (DPSCs) into

odontoblasts that produce dentin, the mineralized component of the tooth. DPSCs were cultured on polymer-coated tissue culture plastic (TCP) in order to determine the effect of the polymer on cell growth and viability. High-throughput screening using polymer microarrays identified these polymeric materials. The best-performing polymer facilitated cell growth levels comparable to the TCP control and was thus chosen as the lead candidate biomaterial for use in *in vivo* studies. A differentiation study was performed to evaluate the level of DPSC differentiation on our materials compared to controls. DPSCs were cultured in osteogenic differentiation media for three weeks on our material, and qPCR analysis will be used to quantify gene expression of key protein markers associated with osteogenic and odontoblastic differentiation. A cell adhesion time-course experiment was conducted to explore the mechanism of DPSC adhesion to our material. Cells were counted and imaged over a 48-hour time period, and RNA was collected. We hope to identify differentially expressed genes involved in adhesion to our materials compared to TCP-positive control and Bis-GMA, the monomer of composite resin dental fillings, as a negative control. To study the *in vivo* performance of our material, we mechanically injured rat maxillary molars in a dental pulp injury model and implanted our polymer, compared to existing filling materials. The polymer will remain in the tooth for four weeks, at which time the tooth will be extracted. We will perform histology, immunohistochemistry, and micro-computed tomography imaging of these to reveal the biocompatibility of our material as well as quantity of dentin at the injury site. This would confirm our *in vitro* findings that our therapeutic material allows for the growth and differentiation of DPSCs. Our material, therefore, could have considerable implications for therapeutic dentistry and may enable the regeneration of decayed dentin. The insights of this study unlock numerous possibilities in dental care, but so too does the mechanism of our material urge the consideration of applying similar materials in other body sites for regeneration.

Engineering Probiotic Bacteria for *In Vivo* Tracking in the Gut

Jessica Kim
Chemistry and Physics, 2019

Wyss Institute for Biologically Inspired Engineering

Advisor: Neel Joshi, PhD

Mentor: Noemie-Manuelle Dorval Courchesne, PhD

The gut microbiome is composed of bacteria that have a symbiotic relationship with the human body, helping to keep the immune system healthy. Because gut microbiomes differ for all people, studying and imaging the gut microbiome has become crucial in developing personalized medications. Thus, the goal of this project is to create engineered bacteria similar to those that already live in the gut that could be ingested and tagged, so that the bacteria can be tracked as they move throughout the gut.

In this study, a probiotic strain of *Escherichia coli* was chosen as the candidate bacteria because it can colonize and survive the harsh gut environment. Our task was to then engineer this microbe so that it could be tagged and targeted with an imaging probe. Previously, protein tags, functional peptides, or binding domains have been introduced as appendages to these bacteria. This study aims to track probiotic bacteria using a novel probing system.

Because the gut environment is highly competitive and many microbes must share a limited amount of nutrients to survive, we wanted to ensure that these engineered microbes could survive in this environment while simultaneously acting as an imaging probe. To test for fitness, we performed growth rate experiments for these engineered microbes and compared their growth to that of their non-engineered counterparts. We used doubling time, the amount of time required for the number of cells in a culture to double, as a proxy for organismal fitness. In addition, our subsequent genetic engineering efforts suggested that our engineered microbes could be selectively tagged with an appropriate functional near-infrared dye. Overall, both the normal growth of our engineered microbes, and their ability to be tagged suggest that, in the future, these engineered microbes would be prime candidates to allow for *in vivo* tracking of bacteria in the gut.

Methods for CRISPR Library Design and *In Vivo* Delivery of Gene Editing Proteins

Joshua Meier
Chemistry &
Computer Science,
2018

Broad Institute of
MIT and Harvard

Advisor: Feng
Zhang, PhD

Mentors: Neville
Sanjana, PhD; Ian
Slaymaker, PhD

The development of CRISPR (clustered regularly interspaced short palindromic repeat)-associated Cas9 nuclease has enabled new methods for interrogating gene function. With just a simple alteration of the single-guide RNA (sgRNA) or direct repeat (DR) sequence, one can reprogram a nuclease to target different sites in the genome. Genome-scale CRISPR-Cas9 knockout libraries have emerged as popular tools for unbiased, phenotypic screening, but technologies for library design in specific applications have lagged behind. Furthermore, in order to fully realize the therapeutic potential of gene editing proteins and study their function in living organisms, new methods for *in vivo* delivery are urgently needed.

Here we describe CRISPR Library Designer (CLD), a web application for automated, tissue-specific library design, and report on delivery methods that can efficiently deliver Cas9-sgRNA nuclease complexes *in vivo.* We then apply gene editing methods to study the development of the mouse brain *in vivo.*

CLD provides a complete pipeline for the design of custom sgRNA libraries for use in a wide variety of applications. It uses fast algorithms to search for genome-wide sgRNA binding sites across an assortment of genes. Furthermore, it integrates tissue expression data from the GTEx Consortium and on-target scores from Microsoft Azimuth to optimize guide selection for specific tissue-based experiments. When possible, CLD targets functional protein domains based on Pfam data, as mutagenesis of functional protein domains leads to a higher proportion of null mutations and enhanced severity of negative selection. CLD can design optimized libraries in seconds and provides intuitive customization features at each design stage for fine-tuning of the libraries. For instance, to design a library targeting 500 chromatin-specific genes with 6 guides per gene requires 8 seconds for the initial algorithmic generation, followed by an optional 10 minutes of user interfacing to fine-tune the library.

Plasmid-mediated delivery of Cas9 into cells can result in uncontrolled integration of the plasmid sequence into the host genome. Furthermore, transfection tools for mediating this delivery show little efficacy *in vivo.* Here we design several methods for *in vivo* delivery, providing a plasmid-free way to use CRISPR-Cas9 with reduced off-target effects. We envision that our methods will facilitate new applications of *in vivo* gene editing.

Thericardium: A Replenishable Reservoir for Targeted Cardiac Therapy

Keegan Mendez
Biomedical
Engineering, 2017

Wyss Institute for
Biologically Inspired
Engineering

Advisor: Conor
Walsh, PhD

Mentor: William
Whyte

Cardiovascular disease (CVD) is one of the leading causes of morbidity and mortality worldwide. Myocardial infarction, one of the main clinical presentations of CVD, is responsible for the death of over 120,000 people each year in the United States alone. Current therapies for myocardial infarction, such as coronary artery revascularization techniques, are effective at restoring lost blood flow to the heart; however, they often cause residual myocardial scarring that patients are left with permanently. Elimination of myocardial scarring and restoration of full cardiac function after a heart attack could eliminate the cascade of events that eventually lead to heart failure. At present, there is an absence of effective clinical therapies in this domain. Simple cell delivery to the infarcted heart has demonstrated promising results, especially in a laboratory setting. However, improved patient outcomes are limited by suboptimal localization of cells to the pathological site, low viability and poor engraftment efficiency. Clinical translation of cell therapy is dependent on new strategies to enable cells to survive and integrate within the harsh environment of the infarcted heart and exert therapeutic benefit over extended periods.

To address these challenges, we previously developed an advanced delivery system entitled "Thericardium," which enables the direct and controllable administration of therapeutic agents (cells, small molecules, proteins) to the heart via a therapeutic reservoir. The reservoir is implanted on the surface of the heart and can be replenished in a non-invasive manner through an implantable conduit connected to a subcutaneous port. This approach shows great promise. However, an adaptive immune response and subsequent cell death create a toxic biomaterial microenvironment, hindering further replenishments.

In the present study, we refined the therapeutic reservoir to maximize transplanted cell retention and survival, and to enable protection from the immune response. The proposed reservoir is composed of a biomaterial core with a cellular payload, surrounded by an outer impermeable soft shell, and separated from the heart tissue by a synthetic semipermeable membrane. The outer shell, made from thermoplastic polyurethane, is designed to provide structural support, while preventing cell loss caused by the mechanical disruption of a beating heart. The semipermeable polycarbonate membrane is engineered to allow for the bidirectional movement of oxygen, nutrients, and cellular proteins, but prevents the inward migration of immune cells and antibodies responsible for the immune rejection of allogeneic cells. The transplanted cells, protected against the host immune response, may then release the bioactive factors necessary for cardiac regeneration over an extended period of time.

In summary, we present a therapeutic reservoir that can mechanically protect its fragile cargo during transplantation and delivery, form an immunological barrier in the harsh inflammatory environment of the incessant beating heart, and be fully retrievable if deemed clinically necessary. We demonstrate that this system allows the targeted, replenishable, and sustained presentation of cellular therapy. The potential applications for this platform delivery system to other diseased tissues are vast.

Drug Delivery through Slippery Lubricant Infused Porous Surfaces (SLIPS)

Natasha Dhamankar
Biomedical Engineering, 2019

Harvard John A. Paulson School of Engineering and Applied Sciences

Advisor: Joanna Aizenberg, PhD

Mentors: Steffi Sunny, M.S.; Ida Pavlichenko, MS, PhD

SLIPS, slippery lubricant-infused porous surfaces, are materials that can effectively repel bacteria and fluids. With this property, they have the potential to improve the performance and safety of existing medical materials. Stents, which are tubes commonly placed in clogged arteries to allow for blood flow, are among such medical implants that could be enhanced by a SLIPS coating. In addition to providing a supportive cage inside an artery, stents release medications to prevent the narrowing of the arterial wall. Nevertheless, because a stent is a foreign object in the body, it prompts clot formation around its own mesh, which could result in stroke or heart attack. Anticoagulants are generally prescribed to prevent this, but such a therapy may cause adverse side effects for patients with other health issues, like risk for internal bleeding.

A SLIPS stent coating could eliminate these side effects by creating a slippery drug-infused layer over the stent mesh so that clots are unable to form. Previous testing of SLIPS confirms their anticoagulant properties; however, it is still unknown whether the roughened matrix of SLIPS is able to release drugs critical to preventing arterial clogging and maintaining blood flow. During my research, I will be investigating the kinetics of drug release from a SLIPS anti-coagulant coating to ensure that it can deliver the same medicinal benefits as the current generation of stents. At this initial phase, research involves developing a variety of nanoparticle layer porous films and testing their ability to both function as a SLIPS and hold and release drug over a period of time. Ultimately, these tests will help craft a better generation of stents for patients.

Developing a Predictive Model to Identify High-Risk Head/Neck Patients with Vascular Anomalies for Early Intervention

Pooja Chandrashekar
Biomedical Engineering, 2019

Boston Children's Hospital

Advisor: John Brownstein, PhD

Mentors: Jared Hawkins, MMSc, PhD; Salim Afshar, DMD, MD

Vascular anomalies are malformations or tumors that result from the abnormal development of blood vessels. Because these anomalies can often progress into complex conditions that affect multiple parts of a patient's body, they necessitate complicated treatment protocols that vary widely among patients. Thus, being able to intervene during the initial stages of patients' disease progression and identify whether a patient is high-risk is critical to minimizing the severe physical and psychological impact later-stage vascular anomalies have on patients. We therefore sought to develop a method that makes use of electronic medical record (EMR) data and physician notes to rapidly and accurately identify high-risk patients with head or neck vascular anomalies early on in their disease timeline. Additionally, we sought to develop a predictive model capable of linking current patient symptoms and diagnoses with possible outcomes, treatments that may be required later, and further complications that may arise. To do this, we worked

with data from the Vascular Anomalies Database, a previously untapped database containing all of the physician notes from all patients diagnosed with vascular anomalies at Boston Children's Hospital. By linking this data with the EMR data describing their diagnoses and outcomes, we hope to first identify key terms within the text that can be used to label a patient as high- or low-risk and then build a machine learning classifier that learns the relationship between these key terms and patient outcomes. Thus, the eventual goal is to use a physician's initial examination of a patient with head or neck vascular anomalies, their notes, and the patient's EMR data to determine if the patient is likely to require a particular procedure later on or to develop a particular condition. From this information, the goal is for the physician to be able to make better-informed, evidence-based recommendations for early intervention, treatment, or lifestyle adjustments that can lead to fewer complications and better patient outcomes.

Identifying Genes Involved in Sepsis Resistance and Tolerance Using Bioinformatics Analysis

Sai Shanthanand Rajagopal
Biomedical Engineering & Studies of Women, Gender, and Sexuality, 2019

Broad Institute of MIT and Harvard

Advisor: Jim Collins, PhD

Mentor: Sriram Chandrasekaran, PhD

Previous literature on viral genomics has shown the viability of diagnosing different viruses by measuring gene expression levels in the blood of humans. From the data of patients' viral loads and severity of symptoms, it is shown that some patients are tolerant to the virus (high viral load but low symptom severity) or are resistant (minimal viral load) to the virus. Using these preliminary data, we are interested in understanding if there is a genetic basis for viral tolerance or resistance.

Blood was drawn on healthy control pigs and infected samples (deceased and surviving) over regular intervals of time. After blood transcriptome sequencing, the lab received gene expression data from all samples, which included sequencing from multiple cell types. Using the MATLAB Bioinformatics toolkit, we pre-processed the data by filtering data with low overall expression values and low entropy. Using a t-test from the MATLAB bioinformatics toolkit, we compared the tolerant vs. sensitive, tolerant vs. resistant, and resistant vs. sensitive populations to determine differentially expressed genes in the studies.

To further filter the list of differentially expressed genes, we used the MATLAB gene ontology (annotation) package to computationally decide upon genes of relevance (e.g. typically found in viral infections/blood regulation). Using clustering analysis, we were able to visualize the differential expression results. Using the open source tool Cell type enrichment analysis applet, we then grouped the gene expression data by characterizing the different cell types from which the samples came.

We will likely find differentially expressed genes that are involved in tolerance mechanisms in specific cells. Finding genes likely responsible for tolerance as well as the cell types they are usually active in helps elucidate pathways that are involved in tolerance mechanisms. Moving from the dry lab to the wet lab then allows drugs to be developed that can take advantage of these tolerance mechanisms in order to provide resistance and tolerance to sepsis.

Developing a Higher Throughput Instrumented Skeletal Muscular Thin Film Device

Sayo Eweje
Engineering Sciences, 2019

Harvard John A. Paulson School of Engineering and Applied Sciences, Wyss Institute for Biologically Inspired Engineering

Advisor: Kevin Kit Parker, PhD

Mentor: Johan Ulrik Lind, PhD

Despite the billions of dollars invested by the pharmaceutical industry annually for bringing new drugs to market, there is a high failure rate in the late stages of drug development because of a lack of screening methods that adequately replicate a patient's physiology. Organs-on-a-chip (OOCs) are a new technology that can solve this problem by recapitulating the necessary structure-function relationships required to mimic both the healthy and diseased states of a particular organ. By doing so, OOCs can provide the pharmaceutical industry with patient-relevant drug testing models to use in pre-clinical screening that will reduce the cost and improve the efficiency of drug development.

One example of this technology is muscular thin film (MTF) devices. These devices allow the contractile strength of engineered smooth, cardiac, and skeletal muscular tissues to be determined *in vitro*. The central principle of these assays is the bending of a thin cantilever substrate by the contraction of muscle tissue attached to the cantilever surface. The cantilever

is composed of elastomeric materials such as poly-dimethylsiloxane (PDMS) or gelatin. The deflection of the cantilevers is determined by visual tracking of the cantilever edge, to give a direct measure of tissue contractility.

Recently, instrumented MTF devices that allow fast quantification of the contractility of cardiac musculature have been developed. However, current devices have not been adapted for use on skeletal muscle tissues. My goal is to address this challenge by developing an instrumented MTF platform that is compatible with engineered skeletal muscle tissues. Such a device could be used in studies of human skeletal muscle for developing treatments for diseases like Duchenne muscular dystrophy or myasthenia gravis.

Muscle responds to both topographical and chemical cues when adhering to a surface. Skeletal muscle tissue and cardiac tissue rely on different extracellular matrix (ECM) proteins to promote cell adhesion. Presently, many scientists who work with skeletal muscle thin films use microcontact printing to add patterned lines of these proteins onto substrate surface, resulting in uniaxial tissues. Micromolded grooves in the substrate surface are also used to guide the alignment and development of the cells. Experiments will be conducted to find the best width, depth, and spacing of these lines for cell adhesion, alignment, and growth, as well as to identify the proper ECM protein(s) to use. The stiffness and thickness of the substrate itself must also be optimized, as skeletal muscle tissues have been shown to have significantly lower contractility than cardiac tissues.

If the project is successful, the device can be used to study the development of skeletal muscle contractility over the course of several weeks. This device can be an important part of pharmaceutical research as a means of quickly conveying information about the contractility of the skeletal muscle tissues being tested during pre-clinical trials and drug dose studies, reducing the need for companies to rely on animals for these tests. Ultimately, this device could provide new possibilities for studying the health and disease of skeletal muscle.

Chemistry and Biochemistry
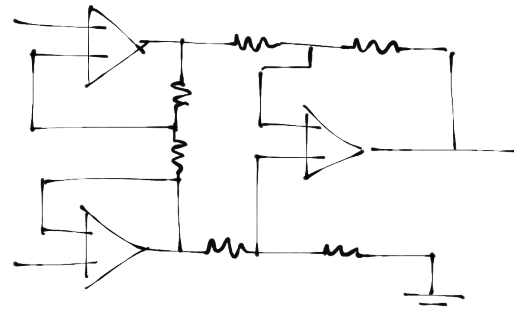
Design, Development, and Synthesis of Novel Macrolide Antibiotics as Tools Against Resistant Bacteria

Andre Sanchez
Chemistry and
Physics, 2018

Harvard University

Advisor: Andrew G.
Myers, PhD

Mentor: Ziyang
Zhang, PhD

With the rise of antibiotic resistance, the reservoir of effective drugs available for treatment of various bacterial infections and diseases has become dangerously low. For decades, chemists utilized a tool known as semisynthesis in order to combat the threat of antibiotic resistance. A semisynthetic strategy relies heavily on the isolation and fermentation of complex antibiotics from nature that are subsequently modified through chemical synthesis. Although semisynthesis has proven to be an extremely useful tool in antibiotic discovery, it is inherently limited. That is, semisynthetic methods are structurally constrained, as starting substrates are extremely complex with many competing functionalities that make certain chemical transformations difficult or nearly impossible. This limitation is further exacerbated with the growing evidence of bacterial resistance to some of the leading macrolides obtained semisynthetically, including azithromycin and solithromycin.

In an attempt to gain control over the impending crisis of antibiotic resistance, a fully synthetic route towards macrolide antibiotics was adopted. Such a technique allows for facile diversification of both appendages of current lead macrolide antibiotics and new structurally diverse scaffolds that were previously inaccessible by other means. Drawing lessons from recent crystallographic studies that elucidated the interactions of antibiotics bound to the bacterial ribosome, new chemical modifications of macrolide antibiotics can be rationalized to improve their on-target binding affinity.

Here, a new synthetic route is presented to provide scalable and rapid access to structurally diverse macrolide scaffolds. Through convergent chemical transformations from simple modular building blocks, one may build chemical complexity quite rapidly. The robustness and versatility of the synthetic strategy was evidenced by the successful construction of more than 500 new macrolides spanning a vast chemical space. The novel macrolide antibiotics were then tested against a panel of gram-negative and gram-positive bacteria in order to evaluate the biological potential of developed drug candidates. Of these newly synthesized macrolide antibiotics, several have demonstrated quite promising activity against macrolide-resistant bacterial strains, further illustrating the potential impact of this approach. More specifically, this work focuses on the diversification and construction of carbazate linkers in ketolide antibiotics, as well as structural modifications in azaketolide antibiotics. Furthermore, this project aims to develop novel ketolide and azaketolide antibiotics with increased on-target binding affinity and overall potency against macrolide-resistant bacteria.

The Molecular Basis of T Cell Exhaustion

Ashavari Ghose
Biochemistry, 2018
Emmanuel College

Brigham and
Women's Hospital

Advisor: Richard
Blumberg, MD

Mentor: Walter Kim,
MD, PhD

The immune system protects the body against countless pathogenic agents, from viruses to cancer. T cells are an essential component of the adaptive immune system, whose function is necessary to specifically defend against these agents. However, T cells have recently been described to enter a dysfunctional state of exhaustion, characterized by altered transcriptional profiles and upregulation of inhibitory receptors on the cell surface. As a consequence, T cell exhaustion severely diminishes the T cell response to both pathogenic invasion and tumorigenesis. Reversal of T cell exhaustion at the molecular level has therefore become a major focus in the development of novel infections and oncologic therapies.

Carcinoembryonic antigen cell adhesion molecule 1 (CEACAM1) and T cell immunoglobulin domain and mucin domain-3 (TIM-3) are cell surface receptors that are known to be upregulated on the surface of exhausted T cells during chronic viral infections and in specific cancers. CEACAM1 functions as a ligand for TIM-3 and mediates enhanced T cell inhibition. Recent studies have shed light on the binding mechanism of human CEACAM1 (hCEACAM1) and TIM-3 (hTIM-3), revealing the specific domains involved. However, the fine molecular details of the interaction remain unknown.

This project investigates the interaction of CEA-

CAM1 and TIM-3, specifically aiming to determine three things: the oligomeric state of the binding reaction (that is, the ratios with which the proteins bind); the binding kinetics; and the sequence determinants (specific amino acid residues that are involved in binding). Previous studies have been complicated by the propensity of hCEACAM1 to homodimerize (bind to itself), and I am therefore taking an alternative approach by using the murine variants of CEACAM1 (mCEACAM1) and TIM-3 (mTIM-3). Co-immunoprecipitation studies have provided evidence for mCEACAM1–mTIM-3 binding, and co-blockade of mCEACAM1 and mTIM-3 enhances anti-tumor immune responses in mouse colorectal cancer models. The aim of this project is to provide direct biophysical evidence for the mCEACAM1–mTIM-3 interaction. To achieve this, I have cloned a naturally-occurring splice variant of the mCEACAM1 (mCEACAM1 D1,D4) into a suitable prokaryotic expression vector, expressed the protein in *E. coli*, and isolated it through the development of a novel purification strategy. The purified protein is being studied to determine the oligomeric state of the mCEACAM1 D1,D4 construct through a series of biochemical assays, including size exclusion chromatography, native gel electrophoresis, and analytical ultracentrifugation.

Studies of mCEACAM1 D1,D4 will inform subsequent binding studies with both the murine and human forms of TIM-3 to determine both species-specific and cross-species binding characteristics. Similar techniques will be employed to determine the behavior of the mCEACAM1–hTIM-3 complex as were used for the mCEACAM1-mCEACAM1 complex. If mCEACAM1 D1,D4 also demonstrates high-affinity homodimerization, as seen with hCEACAM1, alternative binding assays will be employed, such as surface plasmon resonance (SPR) and enzyme-linked immunosorbent assay (ELISA), where CEACAM1 homodimerization can be monitored and controlled. These assays will provide valuable information regarding the interaction mechanism of TIM-3 to mCEACAM1. Ultimately, study of the murine proteins will provide insight into the human system and aid development of specific and efficient therapies for reversal of T cell exhaustion.

Rational Design and Synthesis of Nicotinamide *N*-Methyltransferase (NNMT) Bisubstrate Inhibitors

Brandon Wright
Chemistry, 2018

Harvard University

Advisor: Matthew D. Shair, PhD

Mentor: Rocco Policarpo

Nicotinamide *N*-methyltransferase (NNMT) is an enzyme that catalyzes a simple methyl-transfer from cofactor *S*-adenosyl-L-methionine (SAM) to a nicotinamide substrate. Nicotinamide is an important precursor in the formation of NAD(H) and NADP(H), two key cofactors in metabolic regulation. NNMT's expression is inversely related to production of the GLUT4 sugar transporter that protects against diabetes. As such, NNMT knockdown has been shown to prevent obesity in mice on high-fat diets. Additionally, NNMT is known to be overexpressed in several types of cancer, including breast cancer, bladder cancer, and renal cell carcinoma. Although NNMT plays an important role in various diseases pathologies, no small-molecule inhibitors currently exist for the purpose of studying NNMT's biological function. The fact that NNMT has been shown to regulate metabolic processes in adipocytes and is overexpressed in several forms of cancer makes it a prime target for understanding and treating these diseases.

Rational design of novel bisubstrate NNMT inhibitors began using rigorous *in silico* screening to determine candidates that were both biologically potent and synthetically viable. These docking studies were performed with Schrödinger Maestro computational chemistry suite and assigned a "Glide score" to assess predicted compound efficacy. Bisubstrate inhibitors generally consist of two fragments covalently attached, each fragment targeting a different binding site of a bisubstrate enzyme, therefore achieving a tighter binding by sampling both enzyme binding pockets at once. The initial proof-of-concept compound NS1 was made on a small scale (10 mg) using an asymmetric synthesis consisting of 18 steps from D-ribose. This compound was subsequently evaluated for its ability to inhibit NNMT in an HPLC-based NNMT inhibition assay developed in-house. Lead compound NS1 was determined to inhibit NNMT with an observed IC50 value of 600 nM. A scale-up of the linear synthesis is currently ongoing to generate sufficient material for NS1 analogue synthesis and ultimately for optimized *in vitro* biochemical and cell-based NNMT assays.

Understanding the biological efficacy of NS1 is

essential for evaluating the success of our rational approach to designing NNMT inhibitors using *in silico* screening. Once a <100 nM compound is discovered according to optimized assays, it will be subjected to methyltransferase counter-screening to evaluate NNMT selectivity. The compound will then be evaluated for its ability to reduce 1MNA levels in cells known to express high levels of NNMT. From these efforts we hope to arrive at a potent, selective, patentable drug-like NNMT inhibitor for preclinical development.

Directed Evolution of Cas9

Christina Zeina
Chemistry, 2018

Harvard University

Advisor: David R. Liu, PhD

Mentor: Johnny Hu

The ability of Cas9, a nuclease, to target and edit DNA sequences has made it an invaluable tool for studying biology and curing human diseases. The most commonly used Cas9 is found in *S. pyogenes* as a defense mechanism against viruses. Using a single-stranded RNA guide sequence (the gRNA), Cas9 can be programmed to recognize any DNA sequence with the only restriction being a protospacer-adjacent motif (PAM) sequence that is required to be present.

This summer, my project involved evolving Cas9 to recognize a broader set of target sequences. In order to do this, we utilized phage-assisted continuous evolution (PACE) as a means to evolve Cas9. Developed in Professor David Liu's lab, PACE allows for the evolution of proteins tied to the production of a necessary phage protein, pIII. Its advantage over other evolution methods lies in its ability to complete many more rounds of evolution without researcher intervention, allowing for the quick and precise evolution of macromolecules. We cloned the catalytically-dead dCas9 along with a transcription activator into the phage genome, moving the pIII gene into a separate plasmid. Upon dCas9 activity, pIII production is activated. The system was optimized and was shown to be able to activate pIII production 100-fold over background levels.

After evolution, we submitted phage from PACE for sequencing to analyze the frequency of mutations. As Cas9 activity is tied to the production of new phage, the most prevalent phage mutations should also contain the most active Cas9 mutants. A number of assays were then conducted to characterize the Cas9 and profile the target sites that can be targeted by both the wild-type Cas9 and the evolved Cas9. A library of target sites was encoded on a plasmid,

which was then targeted by catalytically active Cas9. By sequencing the remaining plasmid pool after cutting with Cas9, we can gain an understanding of the target sequences that the Cas9 can target. In addition, both plasmid and genomic target sites were chosen in mammalian cells to test. Cell culture assays were conducted and the results were processed either through flow cytometry or high-throughput sequencing.

Ultimately, in engineering a Cas9 capable of cutting a broader set of target sites, we hope to take steps in the advancement of genetic engineering. Hopefully, this will eventually lead us to evolve and harness a Cas9 capable of targeting and cleaving any DNA. This Cas9 variant may potentially be very important and beneficial for the future of both science and medicine as a gene-editing tool.

Discovering the Biosynthetic Origin of N-N Bonds

Christine Zheng
Chemistry, 2018

Harvard University

Advisor: Emily Balskus, PhD

Mentor: Tailun Ng

Although they are rare, natural products that contain nitrogen-nitrogen (N-N) bonds have long been of interest to medicinal and synthetic chemists alike, due to their unusual structures and potent biological activities, particularly in the treatment of various cancers. Many FDA-approved drugs with indications for oncology contain N-N bonds, as do many of the top 100 most prescribed drugs in the United States. However, despite their usefulness, N-N bonds are difficult to synthesize in a laboratory setting because their precursors tend to be unstable, and very little is known about the biosynthetic origin of these compounds.

Members of the Balskus group have identified various natural products that contain N-N bonds as subjects of further study to determine how these bonds are made biosynthetically. I am working on discovering the biosynthetic pathway for the molecule alanosine, made by the bacterium Streptomyces alanosinicus, which has been used in Phase I and Phase II clinical trials for the treatment of methylthioadenosine phosphorylase (MTAP)-deficient cancerous tumors. To study how this molecule is made, we must first find the genes that code for the enzymes that are responsible for its synthesis. Previous work in the Balskus group has already identified a strong candidate gene cluster, named the *ala* gene cluster, in the *S. alanosinicus* genome. We have since created and screened a genomic library of *E. coli* colonies contain-

ing different segments of the *S. alanosinicus* genome, and have successfully isolated this cluster of interest. Our current work is focusing on definitively linking the *ala* gene cluster to the production of alanosine through heterologous gene expression—transferring the *ala* cluster into host cells that do not naturally produce alanosine in order to see if alanosine production then occurs—and gene knockout experiments—deleting various genes in the *ala* cluster in alanosine-producing strains to see if alanosine production is halted.

Notably, the *ala* gene cluster contains two enzymes that are homologous to enzymes found in the biosynthetic pathways of other N-N bond-containing compounds, and therefore we believe that these two enzymes could be ultimately responsible for the formation of the N-N bond itself. We are also currently working on purifying these enzymes in order to better understand their functions.

It is our hope that studying the biosynthetic mechanisms for N-N bond formation in alanosine and other molecules like it will provide insight into this unique class of natural products and aid scientists in developing new methods of synthesizing them in the future.

Oxygen-Dependent Phosphorescent Porphyrins for the Measurement of Blood Oxygenation Levels

George Qiao
Chemistry, 2018

Massachusetts General Hospital

Advisor: Conor L. Evans, PhD

Mentor: Emmanouil Rousakis, PhD

Oxygen, despite its vital role in the proper function of the human body, is notoriously difficult to measure and quantify in the bloodstream. While electronic methods exist for measuring atmospheric oxygen content, such methods are frustrating and tedious to translate into medical use. Those crossover methods that do exist are unreliable and moderately invasive; moreover, they provide only one-time measurements, and are incapable of monitoring changes in blood oxygen concentration over time. An ideal oxygen sensor would be minimally invasive, finely calibrated, devoid of toxicity, and time-sensitive.

Recent work in the Evans lab has focused on the production of such a sensor utilizing a liquid bandage consisting of oxygen-sensitive porphyrins (an organic compound) suspended in oil. These porphyrins, consisting of four cyclic aromatic structures

bound together and coordinated to a metal atom, are capable of exhibiting oxygen-dependent fluorescence. Put simply, the porphyrin glows strongly in an oxygen-deficient environment, while its luminescence is quenched when placed in an oxygen-rich environment. As luminescence is far simpler to measure than oxygen concentration, the porphyrin thus simplifies the process of elucidating oxygen concentration. Moreover, the porphyrin is stable and long-lasting, allowing for continual oxygen measurement over time.

My project involves further improving the oxygen-sensing properties of these porphyrins. In order to cause the porphyrin to glow, the electrons within its structure must first be excited with a laser emitting photons of the proper wavelength. The current porphyrin's excitation wavelength does not match the wavelengths of commercially available lasers, making it less than ideal for widespread use. Excitation wavelength is related to the size of the aromatic structures that make up the porphyrin's backbone; my goal is to enlarge these aromatic structures in order to decrease the porphyrin's excitation wavelength, pushing it into commercially viable ranges.

Previous attempts at synthesizing these larger porphyrins involved first synthesizing larger aromatic structures and subsequently fusing them together; such methods result in low yields due to the unfavorable entropic loss associated with the fusion step. My approach is to take the already formed precursor porphyrin and grow their aromatic structures. This is a two-step process, involving first the oxidation of alcohol groups present on the ends of each of the aromatic structures into ketones, and second the attachment of small carbon chains to the newly formed ketones. The entropic loss associated with these steps is small; the end goal is to produce the improved porphyrin in larger yields. Attempts at oxidation have already yielded results, with the Swern method in particular showing particular promise. The next steps are the verification of successful oxidation and the search for a carbon attachment method.
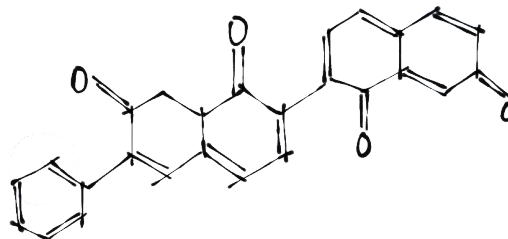
Computer Science

Automated Placement of Components in Printed Circuit Boards

Artidoro Pagnoni
Computer Science & Physics, 2018

Cadence Design Systems

Advisor: Taylor Logan

Mentor: Gu-Yeon Wei, PhD

The design of electronic chips has advanced significantly over the past thirty years, with the number of transistors on a chip nearly doubling every year. Integrated Circuits (IC) techniques can now fit over seven billion transistors on a single piece. Resources were fruitfully placed in the development of IC circuitry, leaving out the optimization of the connections among components at the higher level of printed circuit boards.

This question is becoming of greater importance today in the design of large electronic systems. The placement of components is still done manually for the most part, even when the number of components is in the tens of thousands.

High complexity has increased the price of designing new chips. Specific purpose devices are progressively being replaced with larger groups of standardized chips. The number of components in complex designs is therefore increasing and their high level of optimization makes other factors come into play.

My work has focused on finding algorithmic techniques to optimize the placement of components on printed circuit boards. Our project goes beyond the partial solution proposed in the late 80s by Professors Sangiovanni and Newton at Berkeley University for equal-sized components. We intend to propose a truly practical solution that takes into account connectivity constraints but also magnetic fields, heat dissipation, and multilayer design constraints.

To include these numerous constraints in our optimization, we adopted the well-known genetic algorithms, which generate heuristic solutions simulating the biological process of evolution. The optimization of parameters in our evolutionary framework and our representation of individuals (solutions to the placement problem) will be done by drawing techniques from machine learning. In addition, much effort is being put in the clustering of strongly connected components. We cannot rely on traditional clustering algorithms because connectivity does not increase enough within the groups.

We are hopeful that the overall program will automate the process of component placement, reducing the cost and time of designing new systems and improving their overall performance.

Improving Sequence Handling in Deep Learning

Carl Denton
Computer Science & Physics, 2019

Harvard John A. Paulson School of Engineering and Applied Sciences

Advisor: Alexander Rush, PhD

Recent years have seen an explosion of interest in the field of deep learning, a branch of machine learning that uses today's immense computational resources and massive datasets to train models at a scale that would have been unimaginable just a few years ago. Deep learning is particularly appealing because it requires far less human intervention. Traditional approaches to machine learning rely heavily on domain-specific knowledge, inserted by hand as a set of "features" in the data; in contrast, deep learning simply uses raw data and allows the models to recognize patterns for themselves.

Deep models have seen considerable success, finding widespread application in fields including computer vision, speech recognition, and natural language processing (the focus of this work). Nevertheless, there are a number of tasks in which traditional models still outperform other methods. Conditional random fields (CRFs), a type of probabilistic graphical model, are widely used for sequence prediction (i.e. predicting sequences rather than individual tags) and set the state of the art in tasks including part-of-speech tagging and named entity recognition. Whereas deep models often rely on a huge number of parameters for their modeling power, CRFs exploit the sequential structure of the data they process.

The goal of this work is to integrate CRFs and deep models. A number of current techniques in deep learning make use of sequence predictions, often with far less sophisticated modeling techniques than those used by CRFs. In machine translation, for example, the model is given a source sentence in a source language and is expected to produce an equivalent target sentence in a target language. Many of the most

successful approaches to the task from deep learning process the source sentence by using some notion of "attention"—that is, they weight some parts of the source sentence more highly than others in producing the translation. However, most models ignore the sequential structure of the input in producing these distributions. By applying a CRF, we hope to obtain attention distributions that more accurately reflect the structure inherent in the input.

More broadly, we hope that by integrating CRFs with deep models, we can capture both the sequential modeling power of CRFs, and the flexibility of deep models to create models which can more effectively learn from sequential data.

Concurrency in Transactional Data Types

Lillian Tsai
Computer Science,
2017

Harvard John A.
Paulson School of
Engineering and
Applied Sciences

Advisor: Eddie
Kohler, PhD

Parallelism, where a program runs simultaneously on multiple CPUs, is increasingly critical for performance in computer software systems. This is because CPUs are becoming more numerous per machine, rather than more performant individually. However, parallel programming is also inherently difficult because the programmer must coordinate the execution of multiple CPUs. This has driven a search for better programming tools. The most promising is software transactional memory (STM), that allows programmers to write concurrent code using sequential programming paradigms. Unfortunately, STM is often very slow in practice and rarely considered practical. Recently, a major performance improvement has been discovered at Harvard with the development of a novel type of STM, STO (software transactional objects). The key insight was to implement the STO system at a higher level—data structure operations—than most previously developed systems, which work at the level of individual memory words. With STO transactional data structures, programmers can now easily write correct concurrent code and avoid the performance overhead of a word-STM.

The STO framework relies on libraries of data types built for concurrent operation. These data structures are currently the limiting factor of STO's performance; although they ensure correctness and transactional properties, they do not come close to maximizing scalability or performance. However, there has been significant research done by others on maximizing scalability and performance of these data structures without the concern for transactional correctness. We look to combine these research ideas to produce data structures that both perform well and simplify parallel programming by providing transactional guarantees.

We first experimented with new, *hybrid* concurrency control designs for high-contention data structures. Contention is a performance problem that occurs when many threads have to wait for one another. Most data structure libraries use either pessimistic locking, which behaves well under contention, or optimistic concurrency control (OCC), which is faster in all other cases. STO generally uses OCC, but we introduce hybrid designs that use pessimistic locking for naturally-contentious operations, such as "pop" on queues and priority queues. With this simple change, our naive transactional queues gain performance nearly equal or even better than the cutting-edge research on concurrent data structures. This is a great result, because naive algorithms are much more likely to be implemented correctly. We therefore achieve close to optimal performance with a greater likelihood of correctness.

We did find one strategy, Flat Combining, that performs better than any of our STO implementations. However, when we place the Flat Combining queue into the STO transactional framework even without ensuring transactional correctness, we discover the inherent overhead of the STO system. The STO-Flat Combining queue performs worse than even the original STO queue, and we had not altered the Flat Combining algorithm in the slightest.

This result leads us to think that reasoning about concurrent data structures is *fundamentally different* than reasoning about transactional data structures. We have already discovered the benefits of pessimistic aborts in transactions, in which a transaction will assume that conflicts will occur and therefore either lock the data structure for the entirety of the transaction, or abort other transactions immediately. Because high-concurrency data structures have no notion of transactions, this idea would not be found in high-concurrency research. Perhaps the high-concurrency algorithms have inherent properties that make them unsuitable for conversion to transactional data structures, and the specification for transactional data structures must be weakened in order to make them performant. As we investigate these questions further and come closer to maximizing the performance of the STO data structures, we develop a tool unmatchable in its ability to both simplify parallel programming and speed up program execution.

Effectiveness of Multi-Task Peer Prediction on *Google Local Guides*

Matthew Leifer
Applied
Mathematics, 2019

Harvard John A.
Paulson School of
Engineering and
Applied Sciences

Advisor: David
Parkes, PhD

Mentors: Debmalya
Mandal, MEng;
Victor Shnayder,
PhD

We often hear about the "wisdom of the crowd," but can we trust it? Moreover, if we try to exploit this "wisdom" by paying people for their opinions, can we ensure that they will invest effort and not make up uninformed opinions? In essence, can information be crowd-sourced and reliable? Consider, for example, as *Google Local Guides* now does, asking participants about the qualities of a particular restaurant such as whether it is "noisy," "family-friendly," "romantic," or "wheelchair accessible." While questions may have correct answers, determining those answers ourselves could be near impossible, especially if we wish to learn about restaurants around the world. To solve this problem of eliciting truthful information, a number of protocols for evaluating people's responses, known as peer prediction mechanisms, have been developed. These mechanisms rely on comparing responses from two or more participants on a particular task and perhaps their responses to other tasks, and pay a participant if his report is predictive of other participants' reports. Ultimately, the protocols are designed such that investing effort and telling the truth is an optimal strategy.

The focus of this research project is to apply retrospectively a number of these mechanisms (*Robust Peer Truth Serum*, a mechanism designed by Kamble et al., and a generalized form of *Correlated Agreement*) to the response data gathered by *Google Local Guides*. We will compare the size of the payouts of these mechanisms to the current system, which pays users according to the number of questions they answered. We are going to determine how much these mechanisms would have paid, had they been employed, as well as quantify how large of an incentive they would create for people to provide truthful information. We also plan to calculate what fraction of the participants would have to be truthful in order for these mechanisms to be effective. We also hope to learn if these mechanisms are any more effective on factual questions, such as those that ask about wheelchair accessibility, than questions that seek to elicit an opinion, such as whether a place is "noisy."

Characterizing Predictive Neural Networks

Mingu Kim
Applied
Mathematics, 2018

Harvard John A.
Paulson School of
Engineering and
Applied Sciences

Advisor: David Cox,
PhD

Mentor: William
Lotter, MS

Deep learning is a branch of machine learning that has been gaining attention in the past couple years. It has especially been shown to be effective in the field of computer vision. Using a biologically-inspired neural network architecture, our lab has created a "PredNet" that is designed to accept video sequences of objects and continually predict future frames in a novel manner. The PredNet is able to accurately predict sequences of events that are both synthetically and naturally created, and has a better understanding of the physical nature of object representation than other network architectures. There are two main questions we wish to address. First, we would like to understand how biologically related the network is to our own brains. Second, we want to investigate the generalization capabilities of the network, and whether it has truly learned physical transformation properties. For example, if the network is only trained on video sequences of human faces rotating and animals moving, how well can it predict a bottle spinning? These questions can help show whether the PredNet is actually a functional hybrid that possesses desirable characteristics of deep learning, and the natural methods by which people learn.

To answer the first question, we analyzed the firing rates of each neuron in each convolutional layer to see whether we can find any evidence for positional invariance. This would correspond well with known neuroscience literature. We also wanted to see whether our network can be fooled like our brains can, by specifically using an optical illusion known as the flash-lag effect. We can do this by measuring the firing rates of the neurons for static sequences and comparing them to dynamic sequences at various angles. Either a positive or negative result would be useful, as we could see the degree to which our network has similar characteristics that can be measured quantitatively.

For the second question, we can train the network on video sequences for various different categories of objects and test this network on objects that it has never seen before. We can quantitatively measure performance by using metrics such as the structural similarity (SSIM) index, and per-

forming such tests allows us to analyze what features the PredNet might be using. We can also see if the PredNet can generalize different speeds of movement, by training on a given range of rotation speeds, and then determine if it can accurately predict an object rotating at a speed greater than it has been trained on. We hope to see that the network can demonstrate a deep understanding of general physical properties of object representation.

Automatically Scalable Computation

Peter Kraft
Computer Science, 2017

Harvard John A. Paulson School of Engineering and Applied Sciences

Advisor: Margo Seltzer, PhD

Mentor: Amos Waterland

I am working on the Automatically Scalable Computation (ASC) project, a collaboration between Harvard and Boston University, the goal of which is to speed up single-threaded computation by exploiting additional cores. Almost all modern computers contain several execution engines, called cores. In theory, most computationally expensive programs can be run on multiple cores simultaneously, and will run linearly faster as they use more cores. However, in reality, writing programs that use multiple cores efficiently is hard—much more difficult than writing a similar program that uses only a single core. Most programs that people write are single-threaded, meaning they can only use a single core. This is extremely inefficient as it means the auxiliary cores are simply wasted, their power completely unused. Given the prevalence of multicore computers, it is crucial to develop ways to transform programs written for a single core into ones that take full advantage of multiple cores.

ASC addresses this problem in an entirely new way. It relies on the fact that most programs are deterministic. Deterministic programs are always in one of some number of distinct states, and will always do the same thing after they enter a particular state. The state of a program is the individual 1s and 0s it manipulates in memory. ASC works by monitoring running programs and extracting their states. It then uses machine learning to guess potential future states of the program. Next, ASC has auxiliary cores speculatively execute from the predicted future states. What this means is that additional cores pretend the program has entered the predicted state and figure out what it would do from that state. ASC then creates a cache that maps between predicted states and speculatively executed states. A program can look up a state in the cache and see what it would end up doing if it ever entered that state. Therefore, if the main thread of the program ever realizes that its current state is identical to one of the predicted states in the cache, it can instantly update its current state to the speculatively executed state corresponding to that predicted state, "fast-forwarding" itself into the future and taking advantage of computation done on an auxiliary core. By continuously repeating this cycle of prediction, speculative execution, caching, and fast-forwarding across multiple cores, ASC can allow a single-threaded program to take full advantage of multiple cores.

The goal of my project is to enhance ASC in two related ways: improve the quality of its predictions and expand the class of programs for which ASC can provide significant speedup. The computations that ASC has to perform during operation, like extracting state, making prediction, and doing cache lookup, are not free. Achieving ASC's goal of speeding programs up requires that these be made as fast and efficient as possible, otherwise overhead from them will outweigh the benefits of automatic parallelization. As ASC is made smarter and faster, the set of programs it can speed up will hopefully grow as well, making it a more general and useful tool.

Swarm Locomotion: Biologically Inspired Rules for Self-Organized Bridge Building in an Army Ant-Inspired Soft Robot

Suproteem Sarkar
Computer Science & Mathematics, 2019

Harvard John A. Paulson School of Engineering and Applied Sciences, Wyss Institute for Biologically Inspired Engineering

Advisor: Radhika Nagpal, PhD

Natural systems exhibit complex behavior while employing individual agents with limited information. Understanding emergent behavior in nature can help inform the engineering of artificial systems that operate at scale. *Eciton burchellii* (commonly known as the army ant) creates complex static and dynamic structures, including bridges on foraging trials and living nests, which consist of the organisms' own bodies. Bridges form spontaneously without any apparent central planning, and though their assembly hinders the movement of specific ants that enter the structure, the shorter paths created by these structures increase the efficiency of raids and prey collection for the colony as a whole. Prior work has found these bridges to be robust under changing conditions, sensitive to traffic flow, and favorable to paths that in-

crease colony movement efficiency. The implications of *E. burchellii*'s transport methods were studied in relation to a soft robot that is capable of locomoting horizontally, vertically, and upside-down. A model of the robot in a swarm was examined in physics simulation, given varying sets of rules for locomotion and attachment to other robots in the system. Conditions were varied across randomly-generated terrain types, leading to the formation of robot bridges and ramps. Bridge formation and dissolution was studied in the presence of terrain with valleys, given variable conditions for traffic and valley dimensions. The simulation provided insight into which terrain types and traffic levels would lead to self-organization of structures and which rules would lead to efficient collective locomotion across rugged terrain. Future work will further examine the formation and dissolution of these structures and account for differing rules and initial conditions. By drawing inspiration from nature and computationally simulating self-organization this work aims to help understand collective movement in natural and artificial systems.

MAB Algorithms and Content Recommendation

Val Leifer
Applied Mathematics, 2019

Harvard John A. Paulson School of Engineering and Applied Sciences

Advisor: David Parkes, PhD

Mentors: Nicole Immorlica, PhD; Greg Stoddard

Websites recommend content to their users, from news articles to videos. In doing so, they face a trade-off between recommending old links, which they know a lot about, and new links, about which they want to learn more. The multi-armed bandit (MAB) algorithms are a class of algorithms that are useful for such content recommendation systems.

In machine learning, there is a trade-off between exploration and exploitation—between choosing an action that yields a current high reward (exploitation), and one that sacrifices maximizing current reward for the chance of obtaining better rewards in the future (exploration). The multi-armed bandit problem is a scenario in which a player at a row of slot machines, also known as "one-armed bandits," must choose which machine (or arm) to play. If each arm returns a $0 reward or a $1 reward with a certain probability, then after infinitely many plays, that arm's average payoff equals the arm's probability of returning $1. At each trial, the player faces a trade-off between exploration and exploitation; at each trial, she learns more about the arm that she has pulled, but sacrifices information about all other arms.

If we consider the $0 and $1 reward in the context of content recommendation, $1 becomes a "click" on a link, and $0 becomes "no click." Thus, MAB algorithms learn more about a system's content as users click, or do not click, on the content. Thus, the average "payoff" becomes the average "click through rate" (CTR) of a link.

Today, websites have more information about a video than its CTR. Indeed, videos have tags and users have accounts, which means that we don't start with a blank slate when recommending content. For example, a video might be tagged as "funny" and its title might contain the phrase "Donald Trump"; a user might have an account, where she has stated that she is 24 and from Chicago. If female users over 21 from urban settings have historically enjoyed funny videos about Donald Trump, perhaps so will this user.

The goal of my research is to improve content recommendation and increase videos' CTRs by developing an algorithm that uses these features—video tags and user information—in conjunction with the MAB framework.

In practice, however, observing features is costly—assigning tags to a user or video takes time and effort. Further, in reducing a user or video to a set of tags, certain features are not observed. Maybe when our 24-year-old female user watches a video about Donald Trump, it's not because she likes funny videos about Trump, but because she likes John Oliver, who made the video. When the set of features for a given user-video combination is incomplete, the algorithm can make false connections between known features.

Prior work on content recommendation algorithms assumes full access to relevant features. We are exploring the efficacy of existing algorithms when faced with missing information, and working to develop an improved algorithm that performs well even in the absence of relevant features.

Engineering and Applied Sciences

Test Bed for Minerva, a Deep Neural Network (DNN) Integrated Circuit, and RoboBee Power Electronics

Bahlakoana Mabetha
Electrical Engineering, 2018

Harvard John A. Paulson School of Engineering and Applied Sciences

Advisors: Paul Whatmough, PhD; Mario Lok

Test Bed:

Machine learning has become more popular because of its ability to solve problems that seem challenging (or impossible) to solve using current traditional computing algorithms. Machine learning implements pattern recognition techniques to solve problems such as poor facial recognition and poor handwriting analysis. The Brooks and Wei Lab developed an application-specific integrated circuit (ASIC) that computes deep neural networks (DNN), which, like most machine learning approaches, uses a large set of input and output examples to learn about complex function approximations. The input/output dataset is used as training for data fitting, thereby enabling relatively accurate predictions. My project focuses on designing a test bed for the ASIC designed using Minerva, a co-design framework for optimal design of deep neural network hardware. The test bed needs to flexibly power the ASIC from multiple sources, namely, USB, bench-top power supply, and AC mains. Other specifications of the test bed include portability and compatibility with USB to enable interfacing with laptops to demonstrate the ASIC. The end goal is to make the test bed as compact as possible, ideally the size of a small memory stick.

RoboBee:

Aerial vehicles and microaerial vehicles, commonly known as drones, have recently seen increased applications. Several research groups at Harvard SEAS and beyond joined forces to develop RoboBees. RoboBees are biologically-inspired microaerial vehicles that can fly using flapping wings driven by bimorph piezoelectric actuators. RoboBees have the potential to expand the application of aerial vehicles to crop pollination, search and rescue missions, and environment monitoring. To this end, RoboBees are meant to be autonomous and capable of self-contained, self-directed flight. Currently the RoboBees are powered off of external sources through thin tethered wires, while researchers explore sufficiently light-weight alternatives. The challenge with using tethered wire is its high resistance (ranges from 20 Ohms to 30 Ohms), which when coupled with the 30mA-100mA variation in current drawn by the robot, results in up to a 3V voltage drop. My project involves developing a power supply system that will power chips on the RoboBee with constant output voltage independent of the variations in the current drawn by the robot BrainSoC. This is achieved by implementing a tethered wire feedback loop that adjusts the output voltage. The analysis of the regulator output voltage will enable characterizing different responses of power supply feedback loop.

Harvard iGEM 2016

Daniel Um
Integrative Biology, 2019

Wyss Institute for Biologically Inspired Engineering

Advisor: Neel Joshi, PhD

Mentors: David Lips; Kevin Hoff; Bom Pichet Praveschotinunt; Anita Chandrahas

iGEM is a premier international competition to advance synthetic biology and develop collaboration in this field. This year's Harvard team consists of eight undergraduates, four graduate school mentors, and one professor. After three months of brainstorming and researching, we narrowed our search to four ideas to present to the team. Subsequently, we decided on our project, "Breaking PET," that would break down polyethylene terephthalate (PET)—the plastic molecule that makes up one-sixth of the world's plastic products—with enzymes produced by a recently-discovered microbe. The additional goal of the project is to use the degradation products to generate electricity in a microbial fuel cell.

Project "Breaking PET" is derived from the Oda Group's research conducted at the Kyoto Institute of Technology. Their paper, "A bacteria that degrades and assimilates PET," outlines the research conducted on a bacteria (*I. sakaiensis*) found at a bottle recycling plant. We plan to use two enzymes, PETase and MHETase, to break down this substance into its constituent parts. PETase breaks down PET into the compound MHET, or mono (2-hydroxylethyl) terephthalic acid. MHETase further breaks down MHET into terephthalic acid and ethylene glycol. These two enzymes degrade PET

120 times faster than other PET-degrading enzymes. From the Oda et al. paper, we found the DNA sequences for the enzymes PETase and METase, which we plan to insert into *E. coli*, a commonly used bacteria with a faster reproduction rate than the original *I. sakaiensis*. Using this process, we hope to optimize production of PETase and MHETase in *E. coli* by characterizing the necessary conditions to degrade PET most efficiently. We will create Bio-Bricks (DNA constructs that can easily be swapped around) for PETase and MHETase and characterize these parts against enzymes used in previous PET-degrading iGEM projects. Additionally, we hope to engineer a secretion system for these enzymes.

Following the breakdown of PET, we plan to use one of the degradation products, terephthalic acid, in a microbial fuel cell to generate electricity. Although microbial fuel cell technology is still in its nascent stage, and thus cannot generate a large amount of electricity, we are using our project as a proof of concept to show that this technology could potentially be used to recycle plastic bottles into something useful such as energy.

Photocatalytic Sand for Removing Multiple Classes of Toxins from Water

Deepika Kurup
Neurobiology, 2019

Harvard University

Advisor: Roy Gordon, PhD

Mentor: Marlene Wolfe, PhD Candidate

Water is a substance that is essential for life, yet one-ninth of the world's population lacks access to safe drinking water. According to the World Health Organization, waterborne diseases such as cholera, dysentery, and diarrhea are responsible for more than 2.2 million deaths a year. These diseases are by caused harmful pathogens such as *Escherichia coli* and organics such as benzene, toluene, ethylbenzene, and xylene. Even in developed countries, natural disasters such as earthquakes, hurricanes, and tsunamis can cause flooding and sewage overflow, resulting in water pollution.

Currently in developing countries, solar disinfection (SODIS) that uses ultraviolet radiation from the sun (UVA, $\lambda = 315$–400 nm) is used to purify water; however, SODIS is very slow and can take up to two days to purify water. In recent years, photocatalyst such as titanium dioxide (TiO_2) have been used to accelerate the SODIS process through the creation of reactive oxygen species. While dispersing photocatalytic nanoparticles in water is an effective way to remove bacterial and organic contaminants, it is extremely difficult to recover the nanoparticles from the purified water. Thus, people end up drinking the photocatalyst, and the TiO_2 cannot be reused.

The overall objective of this research is to develop and evaluate a viable method of using photocatalysis to purify water. It is hypothesized that a novel photocatalytic sand could be synthesized, and would be more effective in removing bacteria and degrading organics than a SODIS control. A simple method of synthesizing photocatalytic sand will be developed, in which TiO_2 nanoparticles are coated on the surface of uniformly graded sand. The TiO_2 nanoparticles will not wash off into the water, and the sand will be easy to recover once the water is purified.

For preliminary studies, several ratios of TiO_2 to sand (0.5%, 1%, 2%, and 5% TiO_2 in sand) will be evaluated. The color change of methylene blue, an organic indicator dye, will be quantitatively determined using UV-Vis spectroscopy in order to find the minimum ratio of TiO_2 to sand that is needed for maximum organic degradation efficiency. This optimal ratio will be selected to evaluate the removal of *E. coli* that has been spiked in water to mimic drinking water contamination. Further tests will be conducted to determine the largest volume of water a set amount of the photocatalytic sand can purify, as well as determining what proportion of the TiO_2 (if any) washes off into the water.

It is expected that the photocatalytic sand will be able to degrade organics and remove bacteria at a much faster rate than traditional SODIS methods. The photocatalytic sand is envisioned to have potential applications in point-of-use water purification systems for individual users, water tanks for large communities, and even in wastewater treatment plants in developed countries. This project opens numerous possibilities for green, sustainable, and economically viable water purification.

Printed Sensors for Medical Devices

Nicolas Ontiveros
Electrical Engineering, 2018

Harvard John A. Paulson School of Engineering and Applied Sciences, Wyss Institute for Biologically Inspired Engineering

Advisor: Robert Wood, PhD

Mentors: Michael J. Smith; Ruth Peña Velasco

Medical devices equipped with sensors can greatly improve medical care quality. Specifically, medical practitioners are interested in measuring forces exerted by devices and tracking a device's location and history. Exerted forces are important because certain human tissues must be dealt with delicately. Furthermore, device tracking can ensure instruments are sterilized before a surgical operation and no instruments are left in the body after surgery. For many medical instruments, one can measure forces by equipping the device with a strain gauge, a sensor that measures mechanical strain. Strain gauges are not widely used in the medical industry because they require adhesives, which cannot survive the sterilization process and are at risk of failing during surgery. Adhesives are also problematic for device tracking. Hospitals have begun using radio-frequency identification (RFID) tags to track instruments, but some instruments would require significant modifications and adhesives to be equipped with RFID tags.

When compared to traditional sensors, printed sensor technology offers several benefits: low-cost, low profile, and good adhesion to the device's surface. The goal of this project is to develop printed strain gauges and RFID tags for medical devices. Medical devices under consideration include surgical staplers, surgical robot end effectors, forceps, and hemostats. The printing process involves advanced microfabrication techniques, including chemical and physical vapor deposition. The surface of the device is roughened with either sanding or abrasive blasting and cleaned with acetone. A thin layer of Parylene-C, a biocompatible and biostable polymer, is applied to create a conformal coating on the surface of the device. The strain gauge or RFID antenna is then deposited; this metallization process involves sputter deposition of the desired metals, which include Constantan and copper. Electrical traces are also deposited and attached to the sensor. For the RFID tags, a small chip is attached to the antenna; this chip transmits information to an RFID interrogator when the antenna receives electromagnetic radiation from the interrogator. A final layer of Parylene-C is applied, encapsulating the sensor for electrical insulation. Additional circuits and signal processing techniques are developed to condition and accurately read signals.

Programmable Motion in Colloidal Systems

Will Fried
Mechanical Engineering, 2019

Harvard John A. Paulson School of Engineering and Applied Sciences

Advisor: Vinothan N. Manoharan, PhD

Mentor: Emily Gehrels

Traditionally, colloids, or suspensions in which microscopic particles are dispersed within liquid media, are chaotic and unpredictable. The system I study, where one-micron plastic spheres are suspended in aqueous solution, is no exception to this rule of thumb. In this setup, particles are driven to move randomly through the liquid as they continuously collide with passing water molecules. My project seeks to apply chemical and thermodynamic principles to harness these randomizing processes to program the colloid to behave in a controllable manner.

One method of achieving this controlled behavior involves creating a track of particles on which another particle (called a dancer) is constrained to move, and then manipulating the steps that the dancer takes by altering the environment of the system. To control the interactions between the track and dancer particles, we attach various sequences of single-stranded DNA to the surfaces of different particles such that only particles coated with complementary DNA bind together. We then control the motion of the dancer along the track by embedding different colors of dye into the track particles and sequentially exposing the system to different wavelengths of light. Each dyed track particle should absorb light of a different color and dissipate the energy as heat. Provided that at a given moment the dancer is bound to a dyed particle capable of absorbing the incident wavelength, the subsequent heat release should break the nucleotide bonds between the two particles and drive the dancer towards a neighboring particle to which it is attracted. This process is then repeated to trigger more substantial movement.

However, before we are able to perform experiments using the dancer system, we must conduct several preliminary experiments. First, we must verify that illuminating dyed particles at proper wavelengths can in fact locally heat the particles' surroundings. To confirm this hypothesis, we saturate a solution with dyed particles and add free-floating complementary strands of DNA to the medium. These DNA

strands are modified so that they fluoresce when excited by light of a certain wavelength, but only when they are unbound from their complementary strand. Consequently, the fluorescent signal is weak at low temperatures, where the DNA is predominantly bound, and strong at higher temperatures, where the increase in thermal energy causes the DNA molecules to detach from their complements. Therefore, when we illuminate our sample with two beams of light, one that excites the fluorescence and one that is absorbed by the dyed particles, we expect the fluorescent output to grow as heat is generated by the dyed particles.

The other preparatory experiment involves characterizing the melting curves of DNA strands attached to the surfaces of dyed polystyrene particles. Using a microscope to examine these particles, we can identify the wavelength, intensity, and duration of light exposure required to cause particles coated in complementary strands of DNA to detach from each other for a given dye and at a given temperature. We can then use this data to precisely and efficiently carry out our motion experiment.

A Novel Design for Compact and Parallelizable Current Stimulation and Recording

Youbin Kim
Physics, 2018

Harvard John A. Paulson School of Engineering and Applied Sciences

Advisor: Donhee Ham, PhD

Mentor: Jeffrey Abbott

Accurate micron-scale variable current sources are extremely useful in the field of biotechnology and in improving the interface between electronics and living cells. Especially when electrically exciting cells, current-based stimulation is preferred in comparison to voltage-based stimulation, as the latter could compromise cell viability due to the nonlinear characteristics of the electrode-cell interface. Furthermore, complex communication between circuitry and biology requires multiple stimulation and recording sites, creating a desire for small-scale and parallelizable current sources. This project utilizes a switched capacitor circuit (SCC) as a micron-scale CMOS current source. Although the SCC-based design addresses the problems of size and parallelizability, it also induces a ripple voltage whenever the output switch is closed. The resulting voltage shift, although periodic and based on the frequency of the SCC, can interfere with simultaneous stimulation and recording from the target cells. This project addresses this issue by designing a novel auxiliary circuit to synchronize the SCC frequency to a constant sampling clock while still allowing for the immediate control of the SCC frequency and thus the output current. As a result, all voltage measurements can be taken at the same phase of the SCC input, nullifying the influence of the ripple voltage on the measured signal. The SCC-based CMOS current injection circuit in combination with the newly developed auxiliary circuit allows for the fabrication of tools with large arrays of stimulation and recording sites, which have immediate application in many biological research fields including neurobiology and neurotechnology.

Genetics
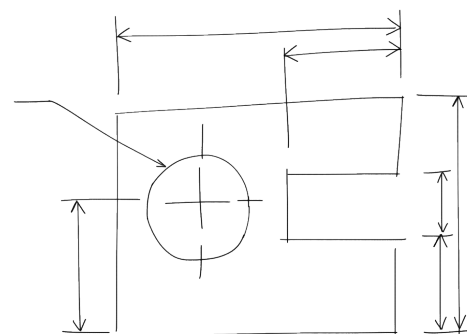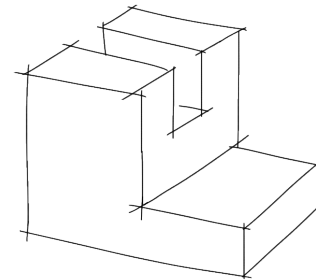
X-Linked Dystonia-Parkinsonism

Anne Cheng
Neurobiology, 2017

Massachusetts
General Hospital

Advisor: Cristopher
Bragg, PhD

Mentors: Christine
Vaine, PhD; William
Hendriks, PhD;
Bjorn Brandl; Lilian
Cruz

Dystonia is a movement disorder characterized by sustained muscle contractions causing abnormal, often repetitive, movements and postures. These movements are typically patterned, twisting, and may be tremulous, which is often initiated or worsened by voluntary action and is associated with overflow of muscle activation.

X-linked dystonia-parkinsonism (XDP), the only known form of dystonia that is inherited as an X-linked trait, is a neurodegenerative movement disorder endemic to the Philippine Island of Panay. Patients with XDP not only experience dystonic symptoms but also exhibit parkinsonism in their later years. XDP is associated with neuronal loss and astrocytosis in the caudate and lateral putamen of patient brains. Because of the high rate of XDP in Panay and the genetically non-lethal nature of the disorder, researchers believe that the disorder must have arisen from a genetic founder effect.

The disease locus, DYT3, affects the gene encoding the TATA-box binding protein-associated factor 1 (TAF1). Within this region is a set of mutations common to all XDP patients that is believed to significantly decrease TAF1 expression levels: four single nucleotide polymorphisms, five disease-specific sequence changes, a 48-base pair deletion, and an SVA retrotransposon insertion. Although these disease specific mutations in the patient genome have been identified, their roles in the development of XDP are not yet well understood.

My goal in the XDP project is to further our understanding of the formation of the disease by studying the disease-specific SVA retrotransposon insertion located in the 32nd intron of TAF1. Retrotransposons comprise about 0.13% of the human genome and have the capability of influencing local transcription. SVA retrotransposons are composed of a CCCTCT hexamer repeat, an Alu-like sequence, a GC-rich variable number tandem repeat (VNTR), a short interspersed element (SINE), and a poly-A tail. The SVA retrotransposon insertion could possibly modulate TAF1 transcription through a variety of different ways such as methylation of protein aggregates.

To further study the SVA retrotransposon insertion, I plan to study its downstream effects in the absence of the other XDP-specific mutations. To do this, I will isolate the SVA retrotransposon from a patient line and clone it into a luciferase vector to allow assessment of the SVA transcriptional activity. Understanding the SVA retrotransposon insertion and its effects on local genes is crucial for both XDP and many other diseases, including some types of cancers that incorporate retrotransposon insertions as a mutation.

Manipulating Innate Immunity Genes to Measure Their Regulation of the Microbiome

Bonirath Chhay
Engineering Sciences, 2018

Rowland Institute at Harvard

Advisor: Robert Brucker, PhD

Mentors: Jessica Dittmer, PhD; Wenchao Gu, PhD; Chris Stokes; Winfield Hill

Host-microbiome interactions are recognized as a fundamental part of an organism's biology, closing gaps in understanding phenotypes left by genomics alone. However, there are still many unknown mechanisms that establish the relationship between the genome of the host and that of the microbiome. In particular, how does the host regulate the microbiome and how does the microbiome influence the host's fitness? It is widely accepted that the host's innate immunity can rapidly evolve and is likely a source of microbiome regulation. Specifically, antimicrobial peptides (AMPs) have been shown to have a role in determining host species-specific microbiota communities. One question, then, is that if the microbiome is so vital to the characteristics of a species, how would the microbiome of an organism be structured if native AMP genes were knocked out or even replaced with AMP genes from another species? Thus, we experimentally test how host species-specific immunity genes affect the composition of species-specific microbiomes in the model animal *Nasonia*, a genus of parasitic wasps. By utilizing the CRISPR-Cas9 genome editing system to manipulate the AMP genes of the *Nasonia* host, we measured phenotypic responses in the animal and its microbiome.

This experiment consisted of two steps. First, we developed a CRISPR-Cas9 genome editing system for *Nasonia* embryos. The CRISPR-Cas9 complex is usually delivered via microinjection, but this method is technically difficult and time-consuming. We developed an electroporation device that is both simpler to use and has higher transformation efficiency than microinjection. Our device consists of two aluminum electrodes held on top of a microscope glass slide by silicone polymer. The device deviates from the traditional electroporation cuvette to make it more suitable for handling embryos.

Second, we manipulated the immune genes to see their effects on the microbiome. We started by knocking down AMP genes of interest with apparent induced expression and then observed the composition of the resulting microbiome via 16S rRNA sequencing, quantitative PCR, and fluorescent *in situ* hybridization. We then knocked in AMP genes specific to other *Nasonia* species to directly study the aforementioned question of how changes in species-specific immunity genes affect the host-microbiota symbiosis.

The results of this research help us better understand vital host-microbiome interactions, which can yield not only a fuller explanation of the current phenomena in an organism, but also the mechanisms that have driven its evolution.

Genome-Wide Analysis of Body Proportion

Kristin Tsuo
Organismic and Evolutionary Biology & Integrative Biology, 2017

Boston Children's Hospital, Broad Institute of MIT and Harvard

Advisor: Joel Hirschhorn, MD, PhD

Mentor: Rany Salem, PhD

Hundreds of common genetic variants have been discovered to be associated with human height, but the genetics underlying body proportion have not been as extensively studied. Genetic analysis of sitting height ratio (SHR), a measure of body proportion, has yielded insights into the biology of skeletal growth, but only a few SHR-associated variants have been identified.

SHR is calculated by dividing the sitting height, the length from a person's head to the surface on which they are seated, by total (standing) height. To expand our understanding of skeletal growth, we conducted the largest (to date) genome-wide association study of SHR, performed on approximately 140,000 individuals of European ancestry. Broadly, a genome-wide association study aims to identify parts of the genome that potentially affect the phenotype being studied, by testing for associations between individuals' genotypes and their trait values. The first genome-wide association study of SHR found six regions of the genome significantly associated with SHR (Chan et al. 2015). In our study, we have identified more than one hundred novel genetic loci associated with SHR and will conduct further analyses to determine the biological pathways and tissues most highly associated with the loci. Together these loci can offer new insights into the genetic and biological basis of body proportion and skeletal growth.

Application of Open Source Machine Learning Methods to Enhance Genetic Variant Classification Accuracy

Matthew Aguirre
Applied Mathematics, 2017

Broad Institute of MIT and Harvard, Massachusetts General Hospital

Advisor: Daniel MacArthur, PhD

Mentor: Laura Gauthier, PhD

The emergence of next-generation sequencing (NGS) technologies has enabled the cost-effective expansion in both number and scale of data-generating projects that empower medical and population genetics. Although the error rate of NGS is very low at the level of individual bases, errors accumulate over large data sets, resulting in false variant discovery. Existing tools, such as the Broad Institute's Genome Analysis Toolkit (GATK), have been designed to generate a set of high-quality genetic variants from short NGS reads aligned to a human reference genome. In the GATK pipeline, variant classification is accomplished via a statistical model built on sequence- and alignment-based scoring metrics. However, these models have been fine-tuned using truth data from older projects consisting primarily of single nucleotide polymorphisms (SNPs) in well-characterized regions of the genome. They therefore have significant room for improvement in classifying other types of genetic variation as true data or technical artifact. For example, the accuracy of indel classification (i.e. insertions or deletions of sequence in the sample genome with respect to the reference) has historically been subpar relative to that for SNPs. In this project, ensemble methods (e.g. boosting and random forests), which make predictions based on a weighted average of the votes of weaker classifiers, were used along with other supervised learning techniques to develop new models for variant classification. A thorough comparative analysis of these models against those currently in GATK production was also conducted. Preliminary results show comparable classification accuracy (~91%) for SNPs between current GATK production and ensemble methods, and significant improvement in the accuracy of indel classification (~90% vs. ~75%) under the new model.

Leveraging Large-Scale Databases of Genetic Variation to Study the Effects of LRRK2 Loss-of-Function Mutations on Health

Sahar Ashrafzadeh
Molecular and Cellular Biology, 2017

Broad Institute of MIT and Harvard, Massachusetts General Hospital

Advisor: Daniel MacArthur, PhD

Mentor: Jamie Marshall, PhD

Parkinson's disease (PD) is a complex neurodegenerative disorder characterized by progressive loss of motor control and cognitive impairment. Approximately 10% of patients have known, disease-causing genetic mutations. Despite being the second most common neurodegenerative disorder, there are currently no preventative therapies available to reduce the progression of the disease.

Studies indicate that rare mutations in leucine-rich repeat kinase 2 (LRRK2) greatly increase susceptibility to PD. Although the pathogenic mechanism of LRRK2 is not fully understood, amino acid changes due to missense mutations likely result in enhanced kinase activity of LRRK2 on α-synuclein. In turn, phosphorylated α-synuclein aggregates in regions of the brain, contributing to the progression of PD. Thus, reducing LRRK2 activity in people with gain-of-function mutations appears to be a promising therapeutic approach against PD. Before seeking a LRRK2 inhibitor, however, we must first understand whether decreased presence of the protein has detrimental consequences on health. Fortunately, the large collection of human DNA sequencing data from relatively healthy individuals in the Exome Aggregation Consortium (ExAC) provides an opportunity to gain insight into the function of LRRK2 by studying the effects of loss-of-function variants on human health.

The goal of our project is to determine whether a LRRK2 inhibitor may be a safe and effective approach to reducing risk and/or progression of PD in individuals who have gain-of-function LRRK2 mutations. For this objective, we will first use ExAC to identify people with potential loss-of-function mutations in LRRK2. Our analysis is restricted to heterozygous individuals since no individuals in ExAC are homozygous for loss-of-function LRRK2 variants—an indication that complete loss of LRRK2 is likely detrimental to survival. Second, we will collect available patient cell lines and use CRISPR-Cas9 genome editing technology to develop cell lines with the mutations of individuals whose cells are

not available. Using these cell lines, we will study protein expression by performing immunoblotting analyses in order to verify that the candidate mutations reduce the amount of functional LRRK2 protein present in human cells. Finally, we will collect phenotypic data from patients and characterize the effects on health, if any, of loss-of-function mutations in LRRK2. Ultimately, demonstrating that people living with reduced levels of functional LRRK2 have no adverse health effects will indicate that LRRK2 inhibitors may be a safe therapeutic approach against Parkinson's disease.

Investigating the Effect of Regulatory Genomic Sequences on Susceptibility to Osteoarthritis

Siddharth Yarlagadda
Human Evolutionary Biology, 2018

Peabody Museum

Advisor: Terence Capellini, PhD

Mentors: Evelyn Jagoda, M.Phil.; Jiaxue Cao

More than three million individuals in the U.S. are diagnosed with osteoarthritis annually. Osteoarthritis is a disease in which the joint cartilage degrades, leading to pain and immobility. While environmental factors influence osteoarthritis, genetic factors are particularly important. Growth differentiation factor 5 (*GDF5*) is the gene region most reproducibly associated with osteoarthritis, with mutation doubling risk. Specifically, risk variants show reduced transcriptional activity in bone and cartilage tissues when compared to the activity of alternative alleles that have no association with risk or disease progression. Importantly, GDF5 is expressed in key joint areas, such as the articular surfaces of bones and ligaments that help stabilize joints. *GDF5* mutations in mice and humans reveal that the gene is required for proper joint, long bone, and tendon development.

Given the importance of *GDF5*, the Capellini lab has elucidated the regulatory control of the gene during embryogenesis: numerous on/off switches (enhancers) around *GDF5* control the gene's expression in specific joints. Within these regions, there are sequences where proteins can bind in order to facilitate transcription of the gene. Each switch potentially influences skeletal development in different areas. One specific switch, the "R37 enhancer region," has been shown to modulate activity in the knee joint. Currently, we are developing a knockout version of the R37 enhancer in mice. This entails removing the region from the genome of the mouse and observing how its loss leads to changes in GDF5 expression and function. The phenotype of the mouse is expected to show abnormal joint development and possibly osteoarthritis at later lifestages.

This phenotypic change is important when assessing osteoarthritic risk, because subtle changes in the development of the joints could also increase the likelihood of osteoarthritis since maintaining homeostasis of the skeleton becomes more difficult. Specific human mutations residing within the R37 enhancer were identified that may alter its activity, i.e., drive significantly lower transcriptional activity than the non-risk allele. In order to test the transcriptional modulation of the variant, we use luciferase-based readout in mouse cell lines. First, risk and non-risk versions of the enhancer from human DNA were inserted and amplified using bacterial vectors. Each vector has a built-in promoter region, which initiates transcription of a nearby luciferase reporter gene. An active enhancer drives the transcription of the luciferase reporter gene, which leads to a higher bioluminescent readout. Thus, if we test both enhancer variants we should expect to see decreased luminescence with the risk allele, corresponding to lower *GDF5* activity. Therefore, the remaining methods involve inserting the vectors into relevant mouse cell lines to see if the risk variant significantly affects transcriptional activity. These experiments will help reveal novel human mutations that play a role in the etiology of human osteoarthritis, with the goal of elucidating the mechanisms that lead to disease onset and progression.

CRISPR/dCas9-VP64-SAM Screen of Cis-Regulatory Elements of the Cdx2 Gene

Yong (Daniel) Shen
Human Developmental and Regenerative Biology, 2019

Brigham and Women's Hospital

Advisor: Richard I. Sherwood, PhD

Mentor: Budhaditya Banerjee, PhD

We know to a great extent how information stored in the genome as DNA produces proteins, but we barely understand how this information flow is controlled by regulatory elements and the transcription factors that bind to these elements. How do we find which regions of the genome play an important role in gene regulation? How do those regions control gene expression? My research endeavors to answer these questions.

We are investigating Cdx2, a homeobox transcription factor that plays an important role in intestinal tissue

development and intestinal cancer development. Recent studies suggest that the overexpression of Cdx2 transcription factor is required for the proliferation of colorectal cancer cells. Cdx2 has also been identified as a promising prognostic marker for intestinal cancer. Though Cdx2 appears to be a vital regulator of intestinal cell development and cancer development, the mechanism by which Cdx2 is regulated remains largely unknown.

In order to discover the regions important in the regulation of Cdx2 (e.g., cis-regulatory elements), we are employing a modified version of the CRISPR-Cas9 genome editing system. In this genome editing system, the Cas9 protein binds to a guide RNA (gRNA) targeting a region of the genome, and excises DNA at the region specific to the gRNA. Scientists have used CRISPR-Cas9 to introduce double strand breaks at desired locations to create mutations and allow the cell's repair machinery to incorporate exogenous genes into locations of interest. We have modified the CRISPR-Cas9 system so that transcription activator proteins could be brought to regions near Cdx2 to induce its expression. To do so, we knocked into mouse embryonic stem cells (mESCs) a catalytically inactive version of Cas9 protein (dCas9) that can no longer cut, fused with ten copies of the transcriptional activation domain VP64—a protein complex composed of four copies of a viral protein that activates transcription. We also designed the SAM activation module—gRNAs that bind strongly to MS2-P65-HSF1 activation helper proteins. Together, the VP64 proteins and SAM module (dCas9-VP64-SAM system) have the potential to induce gene expression. Once bound to a gRNA from a library of thousands of gRNAs that cover the regions flanking the Cdx2 gene, the dCas9-VP64-SAM transcriptional activation domain fusion protein will target a region of about twenty base pairs near the Cdx2 locus and attempt to activate the gene.

We first attempted to create a Cdx2:GFP mESC cell line for fluorescence-activated cell sorting (FACS). After sorting through cells that give high GFP fluorescence, we would use next-generation DNA sequencing to sequence the gRNAs in those cells. From an analysis of the gRNA sequences, we can infer regions important for the activation of Cdx2 gene. However, we were unable to create a Cdx2:GFP mESC cell line. We are now using real-time qPCR to identify cells with high Cdx2 expression and subsequently using next generation DNA sequencing to identify gRNAs and regions important in Cdx2 regulation. We will also explore whether fusing dCas9 to a different activator protein module—p300—can identify the same or different set of important regulatory elements. Our project will show regula-

tory regions of the Cdx2 gene and offer insights into its regulation. We hope that our knowledge of the regulation of Cdx2 will contribute to our understanding of intestinal development and cancer.
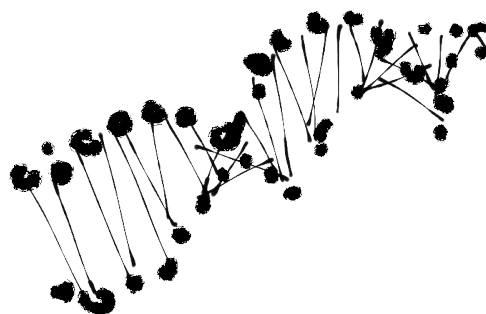
Developing a System for Testing Synchronous Tree-Adjoining Grammar Analyses of Linguistic Phenomena

Jennifer Hu
Mathematics & Linguistics, 2018

Harvard John A. Paulson School of Engineering and Applied Sciences

Advisor: Stuart Shieber, PhD

Mentor: Cristina Aggazzotti

Understanding the interface between syntax (sentence structure) and semantics (sentence meaning) is one of the most fundamental questions in modern linguistics, and our goal is to unify the syntactic and semantic structures of natural language despite their surface-level inconsistencies. The sentence "John hates every boy," for example, has "John" as a syntactic subject but puts semantic emphasis on "every boy." Two sentences can even have the same semantics but different syntactic structures, such as "John likes Mary" and "It is Mary that John likes." With incompatibilities like these in mind, linguists have struggled to model a parallel derivation for syntax and semantics without requiring extremely complicated rules. Our work offers a potential solution to this problem, suggesting that these surface structures could be derived from the same underlying structure through a formalism called a synchronous multicomponent tree-adjoining grammar (MCTAG).

Tree-adjoining grammars (TAGs) produce fully-formed sentences by combining trees that represent natural language expressions. To derive the sentence "Bill saw Mary," for example, we would start with trees that represent the phrases "Bill," "saw," and "Mary," and combine them using the appropriate operations specified by the grammar. A synchronous TAG simultaneously applies these composition operations across multiple trees, giving us a way to characterize relationships between languages. In fact, we can represent syntax and semantics as two languages defined by TAGs, which allows us to derive syntactic and semantic surface trees in parallel through a synchronous process.

To assess the power and limitations of the model, I have been developing a compiler that converts a high-level description of a grammar into a program that can perform the corresponding MCTAG derivations. The user provides a text-based representation of phrases and grammar rules, and my system produces the final syntactic and semantic structures of the sentences by combining the phrases according to the rules. The system currently handles many simple sentences as well as quantifiers and reflexive pronouns, and I hope to extend it to features such as case, person, and number.

Automating these derivations will allow us to refine the MCTAG formalism to deal with complex linguistic phenomena and even extend across different languages. By supporting a common derivation of syntactic and semantic structures, MCTAGs could provide an elegant model of natural language that will not only help us understand how humans communicate, but also lead to exciting applications in machine translation, speech recognition, and other tasks in computational linguistics.
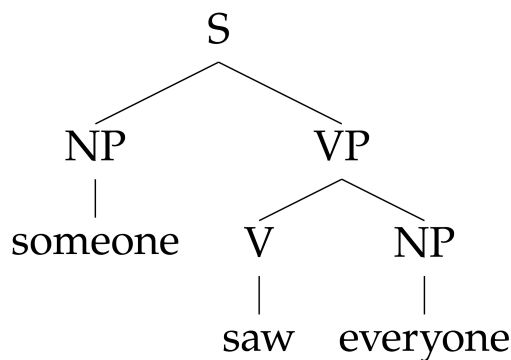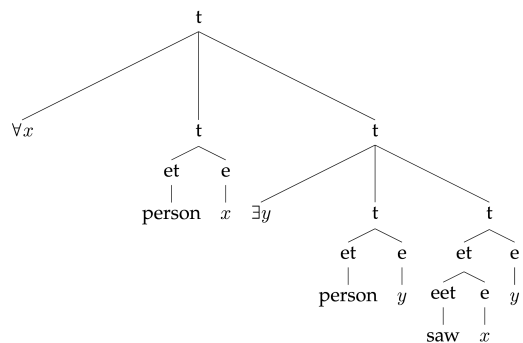

Figure 1: A syntax tree.

Figure 2: A semantics tree.

Mathematics

Elliptic Curves

Alan Bidart
Computer Science, 2018

Harvard University

Advisor: Joseph D. Harris, PhD

In 2000, the Clay Mathematics Institute focused its attention on seven unsolved questions, which leading mathematicians considered of key importance to modern mathematics. These problems became known as "The Millennium Problems." Of these seven problems, six remain unsolved, and my project for the summer is based around developing a better understanding of the premises behind one of them, the "Birch and Swinnerton-Dyer Conjecture," through the study of elliptic curves.

An elliptic curve is a special type of order 3 polynomial on two variables which usually has the form $y^2 = x^3 + ax + b$, where a and b are constants. For every different pair of a, b there will be different set of solutions x, y to the equation. Moreover, for a particular pair, the set of solutions for the elliptic curve could have no elements, a non-zero finite number of elements, or an infinite number of elements. The BSD Conjecture establishes a relationship between the structure of the set of solutions of a given elliptic curve with a, b belonging to the rational numbers and the behavior of some special type of functions called L-functions.

The study of these sets of solutions is of particular importance not only to mathematicians, but also to computer scientists, because these sets have a group structure. In simple terms, a group is a set of elements that can be composed with each other, under a given operation, with the certainty that the result of the composition is still an element in the set. For example, the integers form a group: if we add two integers, the result will still be an integer. Progress towards the proof of the BSD conjecture would help predict the behavior of elliptic curves over the rational numbers and lead to developments in cryptography and computational algorithms.

Investigating the Symplectic Ellipsoid Embedding Function

Handong Park
Mathematics, 2018

Harvard University

Advisor: Daniel Cristofaro-Gardiner, PhD

A symplectic manifold is a smooth (differentiable) manifold M equipped with a non-degenerate skew-symmetric bilinear form ω, called a symplectic form. Understanding these manifolds has important physical applications, as symplectic geometry comprises the fundamental mathematics underlying Hamiltonian mechanics.

To understand the properties of these manifolds, mathematicians often ask whether one manifold can be embedded inside another. We say that a symplectic manifold (M_1, ω_1) can be embedded into another symplectic manifold (M_2, ω_2) if there exists an injective map $\phi: M_1 \to M_2$ such that its inverse (ϕ^{-1}) is smooth and its pullback (ϕ^*) satisfies $\phi^*(\omega_2) = \omega_1$.

Mathematicians haven't uncovered the exact conditions under which a general symplectic manifold embeds into another; however, we can tackle the problem for specific manifolds. One such specific manifold is the four-dimensional symplectic ellipsoid in $\mathbb{R}^4$ with axes of lengths a and b, denoted $E(a, b)$. To better understand the general case, we can study the conditions that govern embeddings of these specific ellipsoids. In this research project, we attempt to study such ellipsoid embeddings by investigating the crucial ellipsoid embedding function $C(x, y)$. Given x and y, this function gives the smallest constant λ for which the ellipsoid $E(1, x)$ embeds into the ellipsoid $E(\lambda, \lambda y)$.

Previously, values for the ellipsoid embedding function were only known when y is 1. It was known that for x less than the critical point $(17/6)^2$, $C(x, 1)$ is given by a piecewise linear function with infinitely many horizontal pieces. In addition, researchers also knew that for x greater than or equal to $(17/6)^2$, the function is given precisely by the square root of x.

This project will determine such critical points for other y-values as well. By using computer-assisted computational methods to apply number-theoretic approaches to the problem, we can estimate $C(x, y)$ for a range of x- and y-values. Using these estimates, we can attempt to find critical x-values when $y \neq 1$, and identify patterns among the shifting points.

Throughout this project, we will attempt to prove such patterns, which provide new knowledge on the nature of the ellipsoid embedding function $C(x, y)$ and could fuel new advances in solving symplectic embedding problems. Along with these patterns, we will also attempt to develop sharp estimates for several different points of the ellipsoid embedding function to better understand its overall behavior. In the future, these developments could eventually help us not only describe, but also determine an explicit formula for the ellipsoid embedding function, which would improve our understanding of symplectic geometry and its various physical applications.

Lines on Hypersurfaces

Hannah Larson
Mathematics, 2017

Harvard University

Advisor: Joseph D. Harris, PhD

Algebraic geometry is the study of geometric objects, called varieties, that arise in the following algebraic way: an (affine algebraic) variety is defined as a collection of points that are simultaneous solutions to a set of polynomials. When the variety is defined by a single polynomial, it is called a hypersurface. The degree of a hypersurface is the degree of the polynomial defining it, where the degree of a polynomial in many variables is the largest total degree of all terms. An example of a degree 2 hypersurface we've all seen before is the solutions of the polynomial $x^2 + y^2 - 1$: the collection of points (x, y) in the plane that satisfy $x^2 + y^2 - 1 = 0$ is just the unit circle. A variety defined by multiple polynomials is the intersection of the hypersurfaces corresponding to those polynomials.

Given a variety, we might want to understand what varieties of a certain type are contained in it. The simplest such question is: when does a given hypersurface contain a line? Trivial examples of hypersurfaces containing lines are lines themselves, planes, and more generally all degree 1 hypersurfaces. As a more interesting example, consider the degree 3 hypersurface defined by $x^3 + y^3 + z^3 - 1$, that is triples of complex numbers (x, y, z) such that $x^3 + y^3 + z^3 = 1$. It turns out that this hypersurface contains 27 lines: if $\omega = e^{2\pi i/3}$ is a complex number whose cube is 1, then you can check that there are nine lines of the form $x = \omega^a, y = -\omega^b z$ where $a, b \in \{0, 1, 2\}$; nine of the form $y = \omega^a, x = -\omega^b z$; and nine of the form $z = \omega^a, x = -\omega^b y$. In fact, this example is representative of a more general phenomenon: every smooth hypersurface of degree 3 contains exactly 27 lines. (Here smooth means that the partial derivatives of the defining polynomial do not all simultaneously vanish at some point on the hypersurface.)

We will discuss the question of when we can expect a hypersurface of a given degree d in n variables to contain lines. We will find that when $2n - d - 3 \geq 0$, a general hypersurface of degree d contains lines, but when $2n - d - 3 < 0$, a general hypersurface of degree d contains no lines.

Enumerative Problems in Algebraic Geometry

James Hotchkiss
Mathematics, 2018

Harvard University

Advisor: Joseph D. Harris, PhD

Algebraic geometry seeks to study, from a geometric point of view, solutions to systems of multivariate polynomials. Central to the field is *preservation of number*—the idea that even if one varies the coefficients of polynomials in a system, the number of solutions to the system generally remains the same. The desire to exploit this idea has spurred numerous developments in algebraic geometry and continues to propel the subject.

Preservation of number gives rise to the possibility of solving an enumerative problem, a problem that asks for the number of geometric objects satisfying some conditions. For example, given three circles in a plane, one might ask how many circles one can find that are tangent to all three (where two circles are said to be tangent if they meet at a single point). Such problems are often classical; a variant of the above problem was first posed by Apollonius of Perga (262 BC–190 BC). Preservation of number suggests that the answer to Apollonius' problem does not generally rely on the original position of the three given circles.

One branch of my research is focused on solving enumerative problems, which proceed in five steps. First, we try to find an appropriate parameter space, a geometric space with nice properties, the points of which correspond to the objects of interest. Second, we calculate an algebraic invariant of the parameter space known as the Chow ring, whose points loosely represent the objects satisfying additional conditions. For example, in the Chow ring of the parameter space of circles, there is a point representing the circles tangent to a single given circle. Third, we identify each point in the Chow ring for each condition that one has imposed. Fourth, we calculate the point in the Chow ring representing objects satisfying all of the conditions. At this step, the point in the Chow ring becomes an enumerative formula, which, if the objects in question behave nicely, yields the correct answer to the enumerative problem. Fifth, we attempt to prove that the objects

in question behave nicely.

A second branch of my research focuses on answering subtler questions about the solutions to a given enumerative problem. For example, in Apollonius' problem, can we solve for the equations defining the tangent circles directly from the equations of the given circles? This line of inquiry leads to a notion of the Galois group of an enumerative problem, which encodes the geometric structure of the solutions to an enumerative problem, and we can solve for the solutions to our problem if and only if the Galois group satisfies an algebraic condition known as solvability. Finding a Galois group requires a delicate analysis of the problem's geometry and is more difficult than the enumerative problem. Indeed, there are many long-solved enumerative problems for which the Galois group is unknown.
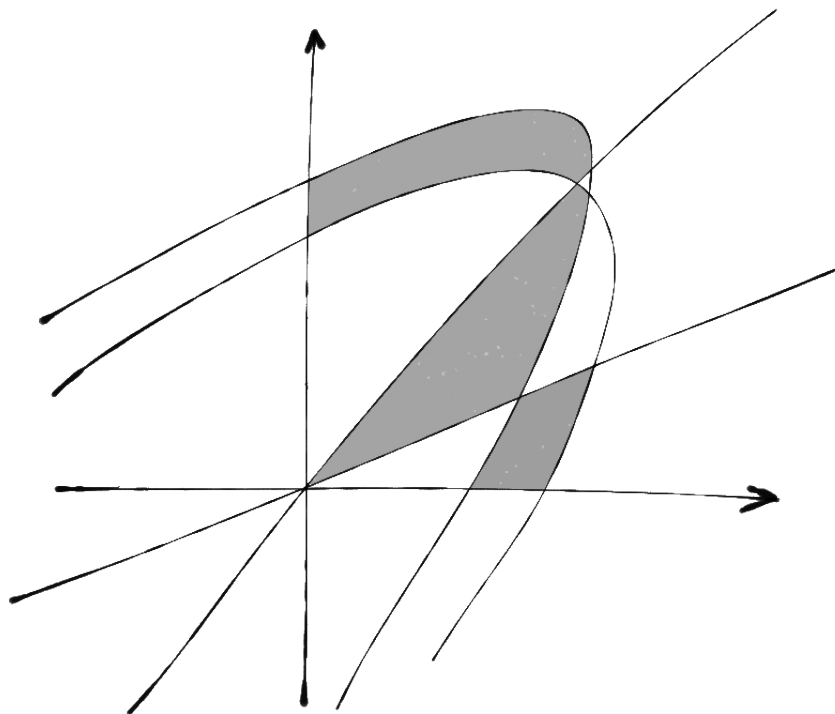
Generating a Comprehensive Alveolar Rhabdomyosarcoma Model in *Danio rerio*

Alexander H. Jin

Molecular and
Cellular Biology,
2018

Massachusetts
General Hospital

Advisor: David
Langenau, PhD

Mentor: Madeline
Hayes, PhD

Rhabdomyosarcoma is the most common pediatric soft core tissue cancer, afflicting approximately 4.6 children out of every million in the United States (Gurney et al., 1996). The malignancy is divided into two subtypes, alveolar (ARMS) and embryonal (ERMS), the former of which is more aggressive and difficult to treat. Molecularly, the two subtypes are distinguished by a chromosomal translocation in ARMS that increases tumor aggressiveness in comparison to ERMS (Barr et al., 1992). As a survivor of Stage 4 alveolar rhabdomyosarcoma, this translocation and its consequences are of particular and poignant interest to me.

The Langenau Lab specializes in and is at the forefront of developing novel disease models in *Danio rerio*, or zebrafish. With low maintenance costs, high rates of reproduction, and ease of genetic manipulation, zebrafish are ideal organisms for generating new models of genetically complex malignancies such as ARMS. With the continued development of the incredibly powerful CRISPR-Cas9 genome editing technology, inducing desired mutations is also easier than ever.

Additionally, the current field of research is lacking in alveolar rhabdomyosarcoma animal models. Of the rhabdomyosarcoma models in both zebrafish and mice, only one is not of the embryonal subtype. Therefore, this summer's project is a continuation of work begun last year to generate a faithful and comprehensive model of alveolar rhabdomyosarcoma in zebrafish by using CRISPR genome editing technology. If successful, this will not only be the first zebrafish and best animal model of ARMS, but will also present a groundbreaking proof-of-concept regarding fusion gene models, which have not been attempted in zebrafish and of which ARMS is only one of many diseases that can be modeled using the techniques we are pioneering here this summer.

The primary approach is to insert a construct at the splice site that can be activated when desired, cleaving the genome at a later time in development and allowing the chromosomes to reattach improperly, mimicking natural ARMS. The fish are injected as single-cell eggs with CRISPRs designed to execute this task and allowed to grow to adulthood. Upon maturation, the injected fish are then screened for the insertion in two ways, the first being a simple evaluation as to whether or not any of the cells in the organism have taken the construct at all. For the individuals that pass this first screen, the next is a deeper sequencing to measure the frequency of such insertions in the fish.

The current project has progressed up to this point. While a few potential mosaic founders (fish that have mutations in some of their cells, but not all) have been identified, no genetically pure offspring have been identified and isolated, indicating that none of the insertions have taken hold in the germline cells of the injected fish thus far. Newer and hopefully better variants of the second method are currently being tested as well, in the hopes of generating the first pairs of mutant fish for breeding, testing, and, with luck, the first artificially-induced alveolar rhabdomyosarcoma tumors in zebrafish ever, by the end of this summer.

Modulation of Intestinal Epithelial Homeostasis and Microbiota with Exercise in a Murine Model

Alexander Munoz

Applied
Mathematics, 2018

Massachusetts
General Hospital

Advisor: Richard
Hodin, MD

Mentor: Sulaiman
Hamarneh, MD

Inflammatory bowel disease (IBD) is a classification of diseases characterized by inflammation of the colon and small intestines. Common symptoms of IBD include abdominal pains, rectal bleeding, vomiting, and diarrhea. These symptoms often present in "flares," where the symptoms are suddenly activated for a time period. In fact, IBD patients' flare management can have a significant psychological and social impact on their lives, as frequent flares can decrease self-esteem and isolate a patient from social spheres. Thus, finding a methodology to increase the time between IBD flares would not only be a somatic therapy for patients, but also a psychological and a social one. But while nearly 400,000 Americans are newly diagnosed

with IBD each year, both the cause of the disease as well as preventative measures are still not entirely clear.

Current research has shown that IBD patients have both decreased microbiotal diversity as well as increased intestinal permeability. In addition, observational research has shown that exercise has a preventative correlation with intestinal inflammatory disease; however, formal research has not yet shown that relationship. As a result, we chose to investigate the effect of exercise on intestinal inflammation in a murine model. Mice were separated into two groups: one that underwent two weeks of daily, rigorous exercise and one that was sedentary for the two-week period. Within those groups mice were divided into two subgroups. One of the subgroups was treated with one week of 3% dextran sodium sulfate (DSS) in drinking water, while the other retained normal autoclaved drinking water. Because DSS causes chemically-induced inflammation of the colon, it has been established as a model for colitis. Following the exercise and DSS colitis phases, mice were sacrificed. Intestinal permeability and gut microbiotal diversity were quantified. In addition, the colon was formalin-fixed, paraffin-embedded, sectioned, and scored for the degree of colitis via histopathology. Using this methodology of exercise and colitis induction with DSS, we compared the exercised and DSS-treated mice to the controls. We thus quantified the effect of exercise as a preventative therapy for the development of increased intestinal permeability, decreased microbiotal diversity, and increased inflammation, which are all associated with colitis. Conclusive differences between the subgroups of mice (for example, one where exercise induces increased innate immune regulation, thus preventing harsh acute colitis with DSS treatment) could yield the potential for exercise as a human colitis preventative measure—and could have therapeutic value to patients attempting to increase time between IBD flares.

Following confirmation of the protective role of exercise in colitis, we chose to investigate the cause of this relationship. To investigate this, we noted a significant exercise-induced upregulation of a brush-border gut enzyme—intestinal alkaline phosphatase (ALPI). We hypothesized that the upregulation of this enzyme, which is involved in regulating intestinal permeability, had a causative role in exercise's colitis prevention. We thus bred mice lacking ALPI and compared the phenotypes with wild-type mice via a repetition of the exercise/DSS experiment. Significant differences between the two groups (such as exercise serving a preventative role for colitis in the wild-type group but not in the group lacking ALPI) could demonstrate a necessary role for ALPI

in the exercise-induced colitis prevention pathway.

Towards the Structure of an Outward-Facing Nramp Transition Metal Transporter

Brandon Lee
Statistics & Molecular and Cellular Biology, 2019

Harvard University

Advisor: Rachelle Gaudet, PhD

Mentor: Aaron Bozzi

Transition metal ions, such as iron and manganese, are essential to life, from enabling oxygen transport to the catalysis of redox reactions. These ions are maintained in a crucial balance within organisms by various mechanisms, including specialized membrane transport proteins such as Nramp family proteins. Nramps (natural resistance-associated macrophage proteins) likely use an alternating access transport mechanism, with the protein alternating between at least two states: one with the substrate-binding site accessible to extracellular space, and another with the binding site accessible to the cytoplasm. The structure of an outward-facing conformation remains unknown, in addition to the mechanism that alternates between the different facing states. I engineered Nramp proteins locked in an outward-facing state from which I now aim to obtain an atomic resolution structure by crystallography.

I introduced individual mutations to tryptophan at eleven positions within the *Deinococcus radiodurans* Nramp (DraNramp), which I predicted would be exposed in an outward-facing conformation. I reasoned that tryptophan, as the largest amino acid, may force DraNramp into a locked outward conformation. This strategy was previously successfully used in another transporter, lactose permease. This summer I conducted preliminary screens of the mutants' conformational preference to determine whether each tryptophan mutation is outwardly-locking the protein.

First, I used *in vivo* cobalt-uptake assays to quantify the relative cobalt transport by my DraNramp variants. I expected that the tryptophan mutants would exhibit varying degrees of cobalt transport, with low activity supporting the hypothesis that the mutant is indeed locked. Consequently, six of my eleven mutants demonstrated the lowered transport activity.

I also used assays that take advantage of the unique reactivity of cysteine to form covalent bonds with its sulfhydryl group. A cysteine must be solvent-accessible to chemical probes, such as the small

membrane-permeable N-ethylmaleimide (NEM), to be chemically modified. By collecting data from various single-cysteine DraNramp mutants, I can infer the likely conformational state imposed by the tryptophan mutation (see Figure 1). After using both inward- and outward-facing reporter cysteines, I have identified six mutants which are highly likely to be locked in an outward-facing state; all six mutants also displayed low transport activity.

I plan to use additional inward- and outward-facing reporters in cysteine accessibility assays and measure binding data for my mutants with various metal ligands in comparison with the WT DraNramp. Following these studies, I will move forward to obtain X-ray diffraction data of my mutants. I will employ membrane-protein specific techniques such as lipidic cubic phase (LCP), which has been successfully employed in the crystallization and structure determination of two conformations of an unrelated metal transport protein. As initial crystal hits often diffract to low resolution, I will also attempt to optimize crystallizable constructs by truncating and/or mutating the loop regions that appeared disordered in our inward-facing structure.
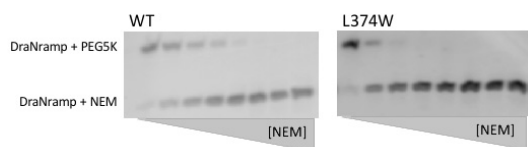
Figure 1: **L374W shows evidence of an outward-facing solvent-accessible vestibule**. I introduced each tryptophan mutation in an A61C background, serving as a reporter for that the outward-open conformation. I then applied different NEM concentrations to the resulting DraNramp variants, denatured the protein, and added a PEG5K-maleimide to mark any unmodified cysteines. Finally, I compared the amount of A61C modified by NEM vs. PEG5K-maleimide using a Western Blot. NEM will modify A61C at lower concentrations in an outward-locked mutant, protecting it from PEG5K-maleimide modification. The WT protein, sampling both the inward- and outward-facing conformations, requires a larger NEM concentration to protect A61C from PEG5K-maleimide modification. In contrast, L374W, an outward-locked candidate, is protected at a smaller NEM concentration.

Manipulating Cellular Morphology through Light Activation of Membrane Curvature-Inducing Proteins

Catherine Li
Chemical and
Physical Biology,
2019

Harvard University

Advisor: Adam E.
Cohen, PhD

Mentor: Hongkang
Zhang, PhD

Neurons have complex cellular morphology, with highly branched dendrites and axons extending from the cell body to receive or send electrochemical signals. Understanding the growth of these projections could allow for the manipulation of new functional synapse formation and shed light on its role in neurological disorders.

I sought to demonstrate a light-inducible system for initiating growth of filopodial and neurite projections in HEK cells and neurons, respectively. We harness the activity of membrane curvature-inducing proteins belonging to the Bin/Amphiphysin/Rvs (BAR) domain protein superfamily. Each class of BAR domain proteins has a uniquely curved three-dimensional shape, and positive residues on one surface of the protein allow it to bind directionally to a cell's negatively charged plasma membrane. Once bound, the curved BAR domain acts as a "scaffold" to which the membrane conforms, resulting in local membrane curvature that can trigger cellular outgrowths. Depending on the class of the BAR domain, the induced membrane curvature can be either inwardly (F-BAR, N-BAR domains) or outwardly (I-BAR domain) oriented. Our work focuses on MIM (an I-BAR protein) and FBP17 (an F-BAR protein).

We coupled the membrane-bending activity of each BAR domain protein to light signals through co-expression with the plant photoreceptor Cryptochrome 2 (CRY2) and its photodimerization partner, calcium- and integrin-binding protein (CIB), in cultured cells. The BAR domain protein is fused directly to CRY2, while CIB is targeted to the plasma membrane by a genetically encoded CAAX motif. Blue laser or LED stimulation excites CRY2, causing it to stick to membrane-bound CIB, thereby also bringing the fused BAR domain to the membrane. We hypothesize that the BAR domain will bind, induce curvature, and promote filopodia or dendritic spine growth. By localizing the initial blue light stimulus to a certain population of cells, or to a specific site on a single cell's membrane, we predict that this light-inducible system will allow us to command the formation of membrane projections with spatial precision at designated locations and times. These structures may in turn develop into functional spines,

which interface and communicate with axons of other cells at a newly-formed synaptic junction. This ability to induce outgrowths and construct new synapses between neurons using simple light stimulation would open doors to many experimental pathways in the study of neural networks and neurological disorders.

Investigating Interactions Between RNaseY and the Y Proteins

Eliza Llewellyn
Molecular and
Cellular Biology,
2018

Harvard University

Advisor: Richard Losick, PhD

Mentor: Aaron DeLoughery, PhD

RNase Y, an endoribonuclease in *Bacillus subtilis*, is responsible for cleaving hundreds of transcripts, playing an important regulatory role in the cell. Strains without the gene for RNase Y have defects in biofilm formation, competence, and sporulation, strategies important for *B. subtilis* in stressful conditions. Bacterial two-hybrid experiments have shown that YlbF, YmcA, and YaaT (the Y proteins), required for biofilm formation, interact with RNase Y and may be necessary for RNA cleavage. My work this summer will investigate the exact region where RNase Y and YlbF interact, using the concept of binary search. This information will help us learn more about the protein, and it will serve as a basis for further experimentation. Using information about where RNase Y binds to accessory proteins, we might ask whether or not it is able to function at all without these proteins.

The foundation for this experiment is a bacterial two-hybrid experiment, which identifies interaction between proteins. One protein is fused to RNA polymerase, while another is fused to a DNA binding domain. The interaction will be tested in both directions by building plasmids that contain RNase Y and YlbF fused to both RNA polymerase and the binding domain. These are expressed in a reporter strain in which the DNA binding domain is upstream of reporter gene *lacZ*. If there is interaction between the proteins, they will bind to each other, which in turn will localize RNA polymerase to the binding domain just preceding the reporter gene *lacZ. lacZ* will then be transcribed, and its activity can be measured and ultimately used to quantify the degree of interaction. This project will involve searching for residues or domains of RNase Y to find the exact region of interaction with YlbF, beginning by testing the different halves of the protein, then searching within the half of the protein that shows a positive interaction. Steps will include creating PCR products of the first and second halves of RNase Y and

a middle fragment of RNase Y (to account for the possibility that binding might occur at residues in the middle of the protein that span the first and second halves), then introducing these to plasmids and transforming bacteria with these plasmids. Ultimately, I hope to make successful constructs of both RNA polymerase and the regulatory region fused to the first and second halves of RNase Y. These will be tested with constructs containing YlbF, as well as different controls.

If successful, this experiment can lead to further understanding of this important ribonuclease. Repetitions of the search—by splitting whichever part of RNase Y interacts with YlbF again in half—can more precisely identify the region of interaction between RNase Y and the Y complex. Then, we can explore the effect of amino acids substitutions at this region to see if RNase Y can cleave transcripts independent of the Y complex. We will also use co-evolutionary analysis to find candidate regions to test. This work will help us learn more about this important ribonuclease and will elucidate details about its recently-discovered interaction with the Y proteins.

The NLGN4X Gene as a Potential Target for Clinical Therapy in Triple Negative Breast Cancer (TNBC)

Felipe Flores
Human
Developmental and
Regenerative Biology,
2019

Dana-Farber Cancer
Institute

Advisor: Kornelia Polyak, MD, PhD

Mentor: Bojana Jovanovic, PhD

Breast cancer is not a single disease, but rather a family of diseases with distinct molecular subtypes. Among these is triple negative breast cancer (TNBC) lacking the expression of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2), hence the name "triple negative" (Foulkes, Smith, & Reis-Filho, 2010). Although TNBC represents only 15–25% of all breast cancer cases in the U.S. (Foulkes et al., 2010), it nevertheless contributes to a significant fraction of breast cancer-related deaths, and in general has a poor prognosis mainly due to limited treatment options beyond chemotherapy (Hudis & Gianni, 2011). Hence, one important focus for TNBC research is the search for alternative TNBC targets in order to find more effective treatments.

The focus of this project is to evaluate the role of *NLGN4X*, encoding for a protein implicated in neural survival, in TNBCs. Specifically, we hypothesize

that *NLGN4X* is required for the survival and proliferation of TNBC cells and may also play a role in maintaining them in a poorly differentiated stem cell-like state. Polymorphisms in *NLGN4X* have been associated with autism and Asperger syndrome (Laumonnier et al., 2004), but it has not been studied in cancer. Based on our analysis of single nucleotide polymorphisms (SNPs) of somatic cell fusions derived from luminal and TNBC cell lines (Su et al., Cell Reports 2015), a SNP linked to *NLGN4X* was the only SNP that was associated with the inheritance of basal features (resembling another subtype of aggressive breast cancer).

Based on this interesting finding we decided to evaluate whether *NLGN4X* has any functional properties in TNBC cells. Thus, we silenced the expression of this gene in SUM159 TNBC cell line using doxycycline-induced shRNA system. The cells were manipulated with lentivirus to introduce engineered fragments of RNA into the cells, and then treated with doxycycline or kept untreated as a control, providing us great power over *NLGN4X*'s expression. A proliferation assay showed that silencing of *NLGN4X* significantly decreased cell proliferation. We will replicate the experiment in a wider spectrum of cell lines, representative of the different TNBC subtypes. If the results are similar to those obtained in SUM159 cells, we will perform xenograft assays in mice, where cancer cell lines will be transplanted into mammary gland of mice to evaluate the effect of *NLGN4X* knockdown on tumor growth. We will also characterize molecular and phenotypic changes that may occur as a consequence downregulation of *NLGN4X* in TNBCs including characterizing gene expression changes by RNA-seq and assessing cell migration, invasion, and metastatic behavior. The successful completion of the proposed project may link a gene implicated in neural progenitors to TNBCs, highlighting the epigenetic plasticity of this breast cancer subtype and identifying a potential novel therapeutic target in a subset of TNBCs.

Transuclear iNKT Mice Show that iNKT Functional Subsets are not Determined by TCR Specificity, but Correlate with Tissue of Residence

Gui Zhen Chen
Molecular and
Cellular Biology,
2019

Dana-Farber Cancer
Institute

Advisor: Stephanie
Dougan, PhD

Mentor: Eleanor
Clancy-Thompson,
PhD

Natural killer T (NKT) cells are a subset of T cells that play a crucial role in immunity, influencing the immune response towards diseases such as diabetes, inflammatory bowel disease, and cancer. These cells recognize lipid antigens presented through the MHC class I-like molecule CD1d, a set of surface proteins found on antigen-presenting cells. There are two main types of NKT cells; here we focus on type I, invariant NKT (iNKT) cells, which express a fixed TCRα (T cell receptor) chain that is complemented with a limited repertoire of TCRβ chains. Once activated, these cells rapidly produce key cytokines, which are chemical signals that can skew the subsequent T cell immune response. iNKT cells are classified into functional subsets that are distinguished by varying production of cytokines and signature expression of transcription factors during thymic development. Among these subsets it has been shown that a combination of the the transcription factors Tbet, PLZF, and RORγt distinguishes NKT1, NKT2, and NKT17 subsets that secrete the cytokines IFNγ, IL-4, and IL-17, respectively.

The question remains as to what factors affect the lineage choice of iNKT cells. Therefore, our research sought to determine 1) whether the specificity of the TCRβ chain impacts lineage choice and 2) whether the tissue of origin impacts lineage choice. To address this, we used transnuclear mice generated through somatic cell nuclear transfer, using the nuclei of iNKT cells specific for different TCRβ chains to generate four lines of mice. One line contains polyclonal mice that expresses iNKT cells with diverse beta chains, and thus, different TCR specificities. Each of the remaining three lines are of monoclonal specificity, meaning each line has one unique TCRβ chain, causing their TCRs to have the same specificity. These mice have increased numbers of iNKT cells in their organs compared to wild-type mice, which allows for the study of rare populations of iNKT cells.

To determine whether the specificity of the TCRβ chain impacts lineage choice, we harvested thymus

and spleen from the transnuclear mice and analyzed iNKT cells for expression of transcription factors and cytokines through flow cytometry, a method that detects cells successfully stained with fluorescent antibodies that bind to the targets of interest. First, intracellular staining for IFNγ, IL-4, and IL-17 assesses cytokine production by these iNKT cells. To determine lineage choice, we then stained intracellularly for the transcription factors PLZF, Tbet, and RORγt. Together, both stainings will allow us to determine if skewing to a certain subset (NKT1, NKT2, NKT17) occurs in our monoclonal mice. If the TCRβ chain plays a determinate role in iNKT cell lineage choice, we would expect to see skewing of monoclonal populations towards one subset.

To determine whether tissue of origin impacts lineage choice, we repeated the same analysis on iNKT cells from multiple organs in the transnuclear mice. If the tissue of origin plays a role, we would expect to see skewing of all iNKT cell subsets in a particular organ.

The central innovation of this project lies in the use of mice cloned by somatic cell nuclear transfer as they provide an abundant source of iNKT cells to study. Expansion of patients' iNKT cells is an approach currently in clinical trials for cancer, and the results have been modest, but promising. As iNKT cells can rapidly produce cytokine in response to antigen, the knowledge of how to induce particular subsets would enhance the efficacy of iNKT cell therapy. Therefore, we aim to determine which factors influence iNKT cell lineage choice to better harness them for future anti-cancer studies.

The Effects of Kin Discrimination on *P. mirabilis* Swarming in Type VI Secretion Mutants

Ian Maynor
Linguistics, 2017

Harvard University

Advisor: Karine Gibbs, PhD

Mentor: Murray Tipping, PhD

When one thinks of social organisms, humans, other primates, or pack animals such as dogs come to mind. Yet even microbes display fascinating social interactions, such as quorum sensing, altruism, and kin discrimination. Kin discrimination is the ability of individual organisms to differentiate other individuals in a population as genetically related or not and modify their behavior accordingly.

The Gibbs lab studies kin discrimination in the bacterium *Proteus mirabilis*, an opportunistic urinary tract pathogen. *P. mirabilis* colonies exhibit kin recognition behavior while swarming on solid surfaces: two genetically identical swarms will merge, while swarms of two different strains form a macroscopic boundary. Furthermore, when two different strains form a mixed co-swarm, one will territorially exclude the other from the swarm front. The Gibbs lab uses *P. mirabilis*'s behavior as a model for bacterial kin discrimination to study the social behavior of bacteria and other microbes.

The Gibbs lab has identified a key gene cluster involved in *P. mirabilis* kin discrimination: Ids. Previous experiments indicate Ids works in a type VI secretion system (T6SS)-dependent fashion to introduce proteins into neighboring cells which mediate kin discrimination behavior. The Ids cluster includes six genes, two of which encode proteins that mediate kin identity: IdsD, an exported protein that signals kin identity, and IdsE, which binds cognate IdsD and is thought to mediate the kin recognition response. The Gibbs lab has produced an *idsD* deletion strain, an *idsE* deletion strain, and an exchange mutant in which the *idsE* is exchanged with that from another strain to study the mechanics of kin discrimination. The Gibbs lab has also created a strain deficient in type VI export, which continues to receive the IdsD identity signal from other *P. mirabilis* cells yet cannot export the signals itself.

This summer, I used these *idsD* deletion, *idsE* deletion, and *idsE* exchange mutants to study how mutations of *idsD* and *idsE* affect kin recognition and understand how these genes mediate kin discrimination, specifically in cases where liquid cultures of two strains are allowed to swarm together. I studied these mutations in a T6SS mutant background, thereby simplifying the two-way interaction of kin discrimination and better isolate how *idsD* and *idsE* mediate kin discrimination effects on non-related bacteria. To study this behavior, I characterized the behavior of T6SS mutants with *idsD* and *idsE* mutations when co-swarmed with the wild-type parent strain. In addition, previous experiments have found that the starting ratio of the two strains inoculated together in a co-swarm can affect Ids-mediated kin recognition behavior. I therefore also varied the starting ratios of each strain to investigate ratio-dependent changes to kin recognition in these co-swarms.

Through these experiments, we hope to further our understanding of the exact mechanism by which *idsD* and *idsE* allow *P. mirabilis*'s kin discrimination system to affect swarming behavior. This knowledge will expand our model of kin discrimination in *P. mirabilis* and allow us to apply these insights to social behavior in other microbes.

Hsf1 Activity and Cell Fitness in Yeast Aging

Jason Li
Molecular and
Cellular Biology,
2018

Harvard University

Advisor: Vlad Denic,
PhD

In their folded states, proteins are the workhorse of cellular machinery. When exposed to proteotoxic stresses, proteins misfold and aggregate, disrupting cell function. One of the hallmarks of aging is the accumulation of protein aggregates in old cells. By contrast, young cells resist the accumulation of protein aggregates, a phenomenon known as protein homeostasis (proteostasis). Exposure of young cells to heat shock revealed a striking example of proteostasis: heat transiently induces the formation of protein aggregates that become resolved over time due to transcriptional up-regulation of protein folding and disaggregation factors.

Studies in the budding yeast have led to the discovery of many longevity factors conserved across the eukaryote lineage, making it a good model system to study aging. In yeast, the heat shock response is mediated by activation of transcription factors Heat shock factor 1 (Hsf1) and Msn2/4. By contrast, there are numerous cellular processes that, if genetically disrupted, only activate Hsf1 but not Msn2/4. Hsf1 is presumed to be essential for cell viability due to its basal transcriptional program, the disruption of which is lethal to yeast. The Denic lab identified the basal core transcriptional program of Hsf1 and termed these genes Hsf1-dependent genes (HDGs), showed that restoring basal expression of just two HDGs encoding the folding chaperones Hsp70 and Hsp90 enables young cells to live without Hsf1, and suggested that Hsf1 is critical for thermotolerance because it up-regulates expression of its basal gene targets.

Preliminary work with fluorescent transcriptional reporters shows that in old yeast cells Msn2/4 is constitutively active (i.e., in the absence of stress) while Hsf1 activity becomes uninducible by heat shock. Moreover, genetic experiments have suggested that Hsf1 inducibility may be repressed by Msn2/4 activity. My goal is to better define the phenomenon of deregulation of Hsf1 and Msn2/4 during yeast replicative cell aging and test if it has consequences on cell aging. My working hypothesis is that Msn2/4 activation in old cells is a double-edged sword. On the one hand, it is most likely helping cells respond to some type of internal stress. On the other, it creates a negative cross-feedback to Hsf1 activity that prevents expression, either basal and/or stress-inducible, of HDGs that are not targets of Msn2/4, which I will call strict HDGs.

Since the basal and heat-induced expression of HDGs depends on Hsf1 in young cells, I will examine the expression of HDGs during aging. I will fluorescently tag endogenous HDG proteins, and track their expression levels by fluorescence in both basal and heat-stress conditions during various ages. Since Msn2/4 represses Hsf1 heat-inducibility in old cells (measured by a transcriptional reporter), I expect to see reduced HDG expression in old cells. Complicating matters, however, is the fact that some HDGs are also regulated by Msn2/4, which becomes constitutively active in old cells. My hypothesis is that basal expression of strict HDGs will be reduced in aging cells. I will also examine the effect of uninducible Hsf1 in old cells on tolerance to a variety of genetic perturbations that normally induce Hsf1. Genome-wide screening has identified gene deletions that induce Hsf1 activation but not Msn2/4. Some of these mutations disrupt protein targeting and misfolded protein degradation, suggesting that Hsf1 activation is adaptive for these types of internal stress. I will call these strict Hsf1 stressors. My hypothesis is that Msn2/4 activation by way of cross-feedback inhibition of Hsf1 makes old cells uniquely sensitive to strict Hsf1 stressors. To investigate this, I will induce degradation of strict Hsf1 stressor in both young and old cells, and fluorescently track if Hsf1 is no longer activated in old cells following strict Hsf1 stress and how this affects old cell morbidity.

Subclonal Cooperation in Triple Negative Breast Cancer—A Study in Intratumoral Heterogeneity

Jett Crowdis
Molecular and
Cellular Biology,
2019

Harvard Medical
School

Advisor: Joan
Brugge, PhD

Mentor: Sabin
Dhakal

Cancer has recently been established as a heterogeneous disease composed of phenotypically and genetically diverse cell subpopulations. This phenomenon, known as intratumoral heterogeneity, is poorly understood, but it has important implications for drug resistance, metastasis, and tumor development. The population dynamics of the tumor ecosystem have long been thought to be governed by competition for scarce resources, but recent studies have indicated that some subpopulations can communicate via secreted proteins, leading to co-

operation within the tumor. The exact mechanism through which subclones crosstalk and the extent to which this impacts tumor progression is largely unknown.

My lab has isolated thirty-one different populations derived from single cells (SCPs) from the triple negative breast cancer cell line MDA-MB-468. Preliminary studies have demonstrated that these SCPs display remarkably different phenotypic properties, including resistance to cell death in suspension. In addition, some SCPs grow better when cultured with the original cell line, indicating positive signaling interactions.

Our research aim is to investigate clonal cooperation within this triple negative breast cancer cell line. Demonstrating a mechanistic basis for clonal cooperation is an important step in understanding how intratumoral heterogeneity and cooperation between the heterogeneous subpopulations can support tumorigenesis.

To investigate clonal interactions, we are utilizing a coculture assay, where subclones are cultured with other subclones and their growth is assessed after two weeks. In previous experiments, we have found that three clones display enhanced survival and growth when cocultured with the original cell line; these are our "dependent clones." Other clones do well in isolation and are resistant to cell death (strong clones). Fluorescently tagged versions of the dependent clones are being grown with the strong clones in a 1:1 ratio, and their growth will be examined with a fluorescent microscope after two weeks. We are examining thirty different pairwise combinations by mixing three dependent clones with ten strong clones. If one strong clone proves able to stimulate growth of all three dependent clones, this indicates positive cooperation.

To investigate the mechanism of clonal cooperation, we are using a number of different approaches. Proteomic analysis of the MDA-MB-468 cell line shows that these cells express higher levels of the TGF-beta family of cytokines, protein factors known to be involved in cell-cell signaling. To investigate if these cytokines are responsible for mediating clonal interactions, we are currently performing knockdown studies with short hairpin RNAs, genetic overexpression studies, and synthetic cytokine/cytokine inhibitor introduction with the coculture assays.

We are currently waiting for data from the coculture analysis, but our biased approach to the cytokine knockdown has been promising. Using Noggin, an inhibitor of the BMP signaling cytokine,

we have seen decreased growth of some clones in the coculture assay. Identifying which clones can support other clones will be important, as we can later examine their genetic and proteomic profiles to identify any active pathways that can explain the clonal interactions. Establishing positive interactions within these SCPs additionally offers the scientific community a resource to study a poorly understood and clinically significant topic.

An Age-Dependent Study of Innervation Patterns of Lymph Nodes

John Austin
Molecular and Cellular Biology, 2017

Harvard Medical School

Advisor: Ulrich von Andrian, MD, PhD

Mentor: Siyi Huang, PhD

The nervous and immune systems work together to help defend the body against harmful pathogens. The peripheral nervous system (PNS) in particular is responsible for sensing and responding to stimuli. Thus, understanding the PNS and, in particular, the sensory nervous system is crucial to better understanding the process of host defense. Recent studies have shown that the innervation of various organs contributes to their function. Studying the innervation of lymph nodes could lead to a better understanding of the role that the nervous system plays in immunity.

Aging in mice produces significant anatomical and biochemical changes in lymphatic vessels. The lymphatic vessels' close association with the surrounding capillary system means that the blood vasculature as well displays some aging effects such as vascular wall remodeling and changes in endothelial barrier integrity. Because lymph node innervating-sympathetic fibers and nociceptors physically associate closely with the vessels, we believe that they will both exhibit age-dependent changes. We are specifically looking at the changes in density of lymph node-innervating neurons with regard to age. Any correlation could suggest an age-dependent relationship between PNS neurons and lymph nodes, in turn suggesting possible functional relationships between the nervous and immune systems.

To research this question, we will use mice of varying ages to look at the innervation density of the lymph nodes over time. Popliteal lymph nodes will be harvested from the mice, and the nociceptors, sympathetic neurons, and lymphatic vasculature will be immunolabeled. Images generated via confocal microscopy will help characterize the innervation pat-

tern and quantify the total amount of neuronal fibers.

We expect that the data will suggest that nociceptors and sympathetic neurons retract with age, leading to an overall drop in neuronal density. It is entirely possible, however, that no relation with age will be seen. However, understanding how age affects the innervation of the lymph node can help us understand the interaction between the nervous and immune systems.

Mechanisms of Virulence in *Klebsiella pneumoniae* ST-258

Julian Braxton
Chemical and
Physical Biology,
2018

Broad Institute of
MIT and Harvard,
Massachusetts
General Hospital

Advisor: Deborah
Hung, PhD

Mentor: Christoph
Ernst, PhD

Antibiotic resistance is a global threat of rapidly growing significance, as more than two million people in the United States alone are infected with antibiotic-resistant bacteria every year. Carbapenem-resistant *Klebsiella pneumoniae* (CRKP) is one such species, and is a major cause of hospital-acquired infections and infections associated with long-term care in the United States. While infections with CRKP are still relatively rare, they have dramatically increased in frequency in the last 10 years and have mortality rates of ~30–70%, making these bacteria a severe public health threat. Among CRKP, ST-258 is the most frequently isolated sequence type, and can further be divided into two genetic clades, denoted 1 and 2. In a recent surveillance study at several Boston area hospitals, clade 2 appeared to be associated with more invasive infections than did clade 1. These clades differ mainly in the capsule polysaccharide synthesis (cps) region, which encodes the necessary proteins for the production of bacterial capsule, a layer of densely packed polysaccharides on the outside of Gram-negative bacteria.

The objective of this project is to compare the impact of the capsule on the virulence of *Klebsiella pneumoniae* ST-258 clades 1 and 2. My hypothesis is that the clade 2 capsule contributes to increased virulence relative to the clade 1 capsule. To test this, I plan to delete the entire cps region in a clade 2 strain, and then insert the corresponding region from a clade 1 strain, generating a hybrid strain of clade 2 genetic background but with a clade 1 capsule. In assays of virulence, this hybrid strain should be less virulent than the corresponding wild type clade 2 strain, suggesting that the capsule at least par-

tially contributes to the differing virulence between these two clades of *K. pneumoniae* ST-258. Such a finding would have the potential to validate the capsule/capsule biosynthesis genes as legitimate drug targets, focusing future drug development efforts on this particular virulence factor of *K. pneumoniae.*

SID-1 Independent RNA Uptake in *C. elegans*

Julio Fierro
Molecular and
Cellular Biology,
2017

Harvard University

Advisor: Craig P.
Hunter, PhD

Mentor: Eddie Wang

SID-1 is a transmembrane protein that facilitates import of extracellular double stranded RNA (dsRNA) in *C. elegans*, which initiates RNA interference (RNAi), a mechanism through which gene silencing occurs. However, there is reason to think embryos can also take up extracellular dsRNA independently of SID-1. The SID-1-independent RNA uptake requires the yolk protein (YP) endocytosis receptor RME-2. While SID-1 is known to facilitate the import of extracellular dsRNA, it is unknown if endogenous RNAs in adults depend on the YP-endocytosis pathway to gain access to embryos. Yolk protein is a vital instrument that serves as a food source for animal embryos and is considered important for successful replication in egg-laying animals, including *C. elegans.* Prior research has led to the hypothesis that RNAs in *C. elegans* could hitch-hike onto the yolk protein and transmit from adults to embryos.

To test the hypothesis that RNA-YP interactions are necessary for transmission of RNAs from adults to embryos, we will pursue three aims. The first seeks to confirm via RT-PCR and qPCR that a mutant lacking *ceh-60*, a YP transcription factor, does not express the vitellogenin precursors of yolk. If the results are successfully replicated, the mutant will fail to express YP and be considered viable for experiments requiring YP knockdown such as creating a *ceh-60*; *sid-1* double mutant. This mutant can then be used to test for RNAi silencing signal uptake defects as well as embryo uptake of endogenous extracellular RNA. Essentially, if we have *C. elegans* who no longer have the means to generate yolk protein, we can then examine if the lack of yolk protein inhibits transmission of RNA from adults to the embryo.

The second and third steps will investigate if endogenous RNAs utilize the YP-endocytosis pathway to

gain access to the embryo. First, wild-type worms will be transformed with tissue-specific mCherry-tagged reporter constructs, DNA plasmids that allow confirmation of expression of the tissue specific genes via fluorescence imaging. Using only the progeny who express mCherry, we will monitor their lineage for loss of expression of reporter constructs in somatic cells, observing which specific somatic cells lose expression. Those worms will then be selected and their embryos taken to analyze what, if any, RNA is found. If the RNA corresponding to the reporter constructs is detected in the embryos of such animals, it suggest these endogenous RNAs use RNA-YP interactions to gain access to the embryo. If the hypothesized results are confirmed, the procedure will then be replicated in a double mutant that knocks out both SID-1 and RME-2 with the expectation that the RNA should not be able to be transmitted to embryos because both SID-1 and the YP-endocytosis pathway will be rendered inaccessible.

The final step will continue investigating the pathway endogenous RNAs take, this time using single molecule fluorescent in situ hybridization (FISH) of tag-257, a protein kinase found in neurons. Single molecule FISH uses fluorescent probes that bind to a specific DNA or RNA sequence, in this case the target sequences, to observe and localize the presence and abundance of the target sequence. Prior research identified tag-257 as a candidate to be an endogenous RNA that uses RNA-YP interactions to go from the soma to embryos. Single molecule FISH will allow us to visualize where in the body tag-257 RNA localizes to and evaluate if it does indeed use these RNA-YP interactions.

Understanding Regulation of Splicing in Response to Infection

Kimberley Yu
Chemical and Physical Biology, 2017

Massachusetts Institute of Technology

Advisor: Chris Burge, PhD

Mentor: Athma Pai, PhD

Eukaryotic organisms have evolved remarkable innate immune systems that allow cells to fight infection by foreign pathogens. These defense mechanisms are dependent on a cell's ability to carefully regulate gene expression and protein levels. One mechanism by which cells can regulate gene expression is through mRNA splicing, which involves excising introns from pre-mRNA molecules and joining exons to produce mature mRNA. Deci-sions about which exons are included in the final transcript—called alternative splicing—enables fine-tuning of mRNA expression levels and the production of a diverse array of proteins.

Previously, my lab characterized the role of alternative splicing in the dynamic cellular response to infectious agents. Using RNA-seq to quantify gene expression levels and isoform patterns in primary macrophages, they found significant global shifts in splicing patterns in response to infection with *Listeria monocytogenes* or *Salmonella typhimurium* bacteria. However, the specific regulatory mechanisms that regulate these global shifts are still unknown, and it is still unclear whether shifts in splicing are part of the cell's immune defense or a deleterious result of the infectious pathogen.

Previous studies have shown that variation in gene expression amongst different conditions and individuals can be informative of the regulatory networks that the cell utilizes to modulate expression. My goals are therefore to understand variation in splicing levels between individuals before and after infection in order to elucidate regulatory mechanisms of splicing that may play a role in innate immune responses in the cell. I propose that an increase in the variance of splicing levels between individuals after infection indicates that the normally tight regulation of splicing has been disrupted, by either individual-specific changes or stochastic variance after infection. A decrease in variance would indicate that infection targets splicing towards a single response regardless of interindividual differences. Little to no change in variance would indicate a tightly controlled event not affected severely by infection. Preliminary results indicate that there are many genes that belong to all three of these categories, and each set of genes differs in their associated functions and attributes. For instance, genes with an increase in splicing variance after infection are likely to be associated with inflammation, cell death, and phosphorylation-related processes, while genes with little change in variance are often involved in essential housekeeping processes.

Currently, I am analyzing a measure of how efficiently cells from each individual were able to kill the foreign bacteria (bacterial clearance counts), to investigate whether ability to fight infection can be predicted by patterns of splicing variance amongst individuals after infection. Ultimately I hope to determine, at a population-level, which splicing events and associated genes, as well as per-individual features, may be important in determining an individual's susceptibility to infection and the efficacy of their innate immune responses.

Enabling Natural Killer Cell Tumor Immunotherapy Using CRISPR-Cas9 Deletion of Inhibitory Receptor Genes

Kruti Vora
Molecular and Cellular Biology, 2017

Harvard Stem Cell Institute

Advisor: Chad Cowan, PhD

Mentor: Torsten Meissner, PhD

Cancers, highly proliferative and prone to mutation, evade the body's immune system by escaping both T and natural killer (NK) cells. In 2016, it is estimated that over 1.6 million new cases of cancer will emerge while the disease will claim around 600,000 lives. It is thus of pressing urgency and importance to find new methods to help patients' immune systems recognize and eradicate malignant cells to prevent unchecked growth of tumors.

Normally, circulating T cells act as a primary defense mechanism against malignant cells, scanning peptides presented on major histocompatibility complex class I (MHC-I) molecules on the surface of cells as potential markers for destruction. However, cancer cells escape T cell surveillance by down-regulating MHC-I.

As a backup system, NK cells recognize aberrant cells without the aid of MHC-I molecules, using the release of cytotoxic granules to eliminate cancerous and virus-infected cells. But certain cancers fight back, expressing various inhibitory ligands that bind to corresponding receptors on NK cells to prevent their activation.

NK cells play an especially large role in targeting cancers with low mutation rates, such as pancreatic cancer, as T cells cannot recognize these tumor cells. However, prior research has shown that pancreatic cancer can inhibit NK functioning, and it currently remains the fourth most lethal cancer.

My research goal is to use the CRISPR-Cas9 system to target the KLRC1 inhibitory receptor, its accessory chain (KLRD1), and inhibitory KIR family genes in NK cells to test the resistance of these modified cells to the effects of various inhibitory ligands.

Deletion of these inhibitory receptors should enable NK cells to recognize and kill malignant pancreatic tumor cells, even in the presence of inhibitory ligands. We use 721.221 cells as our target cell line since these MHC-I negative cells are able to bypass the T-cell immune response and act as an optimal target for NK cells.

In order to create these genetically modified NK cells, I first designed and cloned CRISPR guide RNAs (gRNAs) into lentiviral expression vectors to knock out our target genes in both immortalized and primary NK cell lines. Next, I characterized the normal inhibitory NK cell receptor expression using fluorescence-activated cell sorting (FACS). Once expression is established, I can compare the degranulation and killing activity of the modified NK cell lines against native 721.221 cells or 721.221 cells that have been reconstituted with the corresponding inhibitory ligands. I will also include a negative control cell line in the killing assays that should not be susceptible to modified NK cell attack.

I hypothesize that my genome-edited NK cells will be resistant to the inhibitory effect of these ligands, showing greater degranulation and killing activity. If my hypothesis is correct and we are able to enable NK cells to recognize and kill cancer cells, we could create new treatment options to fight cancers using NK cell tumor immunotherapy.

Artificial Antigen-Presenting Cells for CAR T Cell Activation

Leah Marsh
Molecular and Cellular Biology, 2019

Massachusetts General Hospital

Advisor: Marcela V. Maus, MD, PhD

Mentor: Felipe Bedoya, PhD

Harnessing the body's own immune system to fight diseases such as cancer is an idea that has been around for many years, but in the past decade the field of immunotherapy has seen tremendous advances. One recent immunotherapy innovation is the chimeric antigen receptor (CAR)-modified T cell. In CAR T cell cancer therapy, the chimeric antigen receptor targets an antigen that is expressed on cancer cells, and then causes the CAR-modified T cells to traffic to the cancer cells and destroy them. To create a CAR T cell, T cells are genetically modified to express the CAR, which is made of parts from the T cell's own antigen receptor as well as parts from antibodies (proteins the immune system uses to recognize antigens). The CAR recognizes a specific antigen based on the type of antibody used to create it, and CAR T cells will traffic to cells expressing that antigen.

In order for T cells to proliferate in the body, they must be activated by the T cell's receptor recognizing and binding to a complex on the surface of an

antigen-presenting cell. In the laboratory, one can engineer artificial antigen-presenting cells (aAPCs) to activate T cells *in vitro*. These can be used for normal T cells, or CAR T cells. The goal of my project was to create and evaluate aAPCs that activate CAR T cells recognizing the antigens CD19 and BCMA. These are antigens expressed on the cells of multiple myeloma and other blood cancers.

First, we created lentiviruses carrying the genes for CD19 and BCMA that insert this antigen DNA into the genome of "blank" cells, causing the blank cells to become aAPCs expressing CD19 or BCMA on their surface. To make the virus, we used a process called transfection to introduce the viral DNA, which coded for the virus's physical structure, and the CD19 or BCMA antigen DNA into a packaging cell line. After this, the viruses assembled themselves in the cytoplasm. We then collected the viruses and used them to transduce the blank cells as described above, in which the viral vector delivers the antigen DNA to the genome. After transduction, the blank cells became aAPCs, expressing either CD19 or BCMA on their surface (they do not express any viral proteins).

We also used this process to create the CAR T cells, first creating lentivirus that carried DNA for the CAR, and then transducing a T cell line with the virus so that those T cells expressed the CARs recognizing CD19 or BCMA.

By culturing the aAPCs and CAR T cells together, we can observe whether the CAR T cells become activated and begin proliferating. We can also test different ratios of aAPCs to CAR T cells to see which ratios work best; we hypothesize that there is a specific ratio of aAPCs to CAR T cells that will induce high levels of activation and proliferation, and we hope to find this ratio. Using these aAPCs will facilitate the study of CAR T cells as treatment for multiple myeloma and other blood cancers.

Studying Epithelial Polarity in the Nematode *C. elegans*

Leah Rosen
Applied Mathematics, 2019

Harvard University

Advisor: Susan Mango, PhD

Mentor: Stephen Von Stetina, PhD

Epithelial cells have a crucial role as barriers and transporters in a diverse range of complex organisms, from humans to *C. elegans* (roundworms). In the intestine, for example, epithelial cells keep toxins out while transporting nutrients in. Epithelial cells need to be able to differentiate between their different sides, and for this reason they are polarized. They polarize by differentiating between the proteins contained in their membrane on different sides. The junctional domain separates the two main sides of the cell (the basal from the apical membrane), and connects the epithelial cells together so that they form a continuous barrier. Many diseases arise when epithelial polarity is disturbed, and so it is crucial that we understand the pathways and molecules that give epithelia their polarity.

The *C. elegans* pharynx is an ideal system in which to study epithelia due to how simple, well-studied, and well-conserved across organisms it is. The pharynx, also known as foregut, is at the front of the digestive tract, just behind the mouth. *C. elegans* are both simple and transparent, so that many things can be observed *in vivo*, and we have a lot of context while studying them. Despite the simplicity of *C. elegans*, the molecular nature of epithelial cells is very similar in *C. elegans* and in mammals. Specifically, PHA-4 (FoxA in mammals) is a transcription factor that has been found to be essential for pharynx development in all organisms tested.

The Mango lab has found a pathway of conserved molecules, culminating in the expression of the junctional protein DLG-1. The Mango lab has shown that in embryos lacking pha-4, DLG-1 protein is reduced in the foregut. Surprisingly, the RNA for *dlg-1* is still abundant, which is unexpected as PHA-4 is a transcription factor. The first aim of my project is focused on discovering why there is still abundant *dlg-1* RNA in *pha-4* mutants. We are looking at two possibilities. Firstly, we are studying the gene *lin-26*, which is known to regulate *dlg-1* expression in the skin, and is found in the pharynx in mutants that lack *pha-4*. The hypothesis we are testing is that *lin-26* is normally inhibited by PHA-4. It follows that in the absence of PHA-4, *lin-26* would be expressed and responsible for the activation of *dlg-1* RNA. In the second possibility, we are investigating

whether a PHA-4 interacting protein and its homolog are involved in *dlg-1* regulation. The lab previously observed a phenotype only when both the protein and its homolog are absent.

In the second part of the project we are studying another part of the pathway. A protein called ZEN-4 regulates DLG-1 protein production based on the *dlg-1* RNA that was regulated by PHA-4. The molecular process of how ZEN-4 controls this transition is completely unknown. The Mango lab performed an immunoprecipitation followed by mass spectrometry to identify ZEN-4 interacting proteins. We are looking at these and seeing if they have an impact on DLG-1.

We are studying all of these hypotheses by inactivating the relevant genes and analyzing their effects in worms of different backgrounds, either by using mutant worms, or by performing RNAi, in which the worms are fed bacteria that produce a molecule that interrupts the gene expression pathway. We then analyze the effects in embryos, when the mutants are still viable, using antibody staining to see protein expression and single molecule fluorescent *in situ* hybridization (smFISH) to see RNA expression.

Elucidating the Role of FXR1d in MicroRNA Activity and Cancer Development

Madeleine Granovetter
Molecular and Cellular Biology, 2019

Massachusetts General Hospital

Advisor: Shobha Vasudevan, PhD

Our lab studies gene expression in quiescent cancer cells, a subpopulation of cells in cancers that resist clinical therapy and give rise to cancer recurrences. FXR1 is an RNA binding protein with multiple forms, and is associated with aggressive cancers, such as leukemia, lung cancer, and breast cancer. The lab previously uncovered that one isoform, FXR1a, is overexpressed in quiescent cancer cells and promotes translation by microRNAs. MicroRNAs are small, non-coding RNA regulators of gene expression that play critical roles in cancer, either as tumor suppressors or as oncogenes. In the cancers with which it is associated, FXR1 is overexpressed and promotes translation of important tumor invasion genes.

Research in the lab has found that knockdown of endogenous FXR1 (all isoforms) alters levels of microRNAs and their regulatory proteins, and reduces clinical resistance in cancers; however, FXR1a does not interact with microRNA regulators. This suggests that other FXR1 isoforms may regulate microRNA levels. FXR1d is the other isoform present in the cancer cells studied in our lab. Previous research has found that unlike the a isoform, FXR1d does not promote translation; however, the role of FXR1d in microRNA levels is not known. The goal of my research is to investigate the roles of FXR1d in regulating microRNA levels and activity. I am investigating the effects of the overexpression of FXR1d on microRNA levels and their regulators. My research will therefore provide insights into the role of FXR1 in regulating these important RNAs in cancer and clinical resistance.

A typical experimental procedure I use is to transfect different types of cancer cells with GFP as a control (to show that the transfection was successful), as well as with FXR1a, FXR1d, and forms of the FXR1a gene with mutations that prevent its ability to interact with RNA. The proteins expressed have a small tag called Flag that allows me to detect them separately from endogenous FXR1 protein. The cells I use are MCF7 (breast cancer cells) and H520 (lung cancer cells), which express FXR1. I provide serum to half of the samples, and serum-starve the other half so that I can make a comparison between proliferating cancer cells and cancer cells that are quiescent—or in the G0 state—due to serum starvation. I perform western blots with Flag antibody to detect the FXR1 I transfected, as well as with various antibodies to see whether overexpression of these FXR1 forms affects the levels of microRNA regulators. Ultimately, I will use qRT-PCR to measure and compare the roles of FXR1a and FXR1d in microRNA regulation. I can consequently see which isoforms of FXR1, or which mutations of FXR1a, are involved in microRNA regulation.

Given that my experiments run successfully, I hypothesize that overexpression of FXR1d causes an alteration in microRNA levels by either promoting the activity of oncogenic microRNAs or reducing the activity of tumor suppressor microRNAs. Ultimately, I strive to understand how FXR1d is contributive to cancer development in both proliferating and clinically resistant G0 cells.

Role of Hexamer Repeats in Pathogenesis of XDP

Mai-Linh Ton
Molecular and
Cellular Biology &
Visual and
Environmental
Studies, 2019

Massachusetts
General Hospital

Advisor: Cristopher
Bragg, PhD

X-linked dystonia-parkinsonism, referred to as XDP, is a rare neurodegenerative movement disorder. It combines symptoms of dystonia, such as sustained and repetitive muscle contractions and twisted postures, and parkinsonism, which includes the slowing of movements. Those affected can have their maternal ancestry traced back to Panay in the Philippines, with individuals passing down the XDP-specific mutations in an X-linked recessive pattern.

XDP involves the gradual loss of medium spiny neurons; however, the mechanism behind this loss is not fully understood. Additionally, it is not known why the age of onset can vary so greatly for those with XDP. While some are affected in early adulthood, others may only show symptoms late in life.

The XDP-specific variants are believed to affect the gene encoding for TATA-box binding protein associated factor 1 (TAF-1). These variants include five disease-specific nucleotide changes, DSCs 1, 2, 3, 10, and 12; a 48 bp deletion; and an SVA retrotransposon insertion in the 32nd intron of TAF-1.

Genotyping individuals has revealed that all affected individuals contain all markers for the disease. However, the length of the SVA retrotransposon can vary among individuals. One region within the retrotransposon that varied the most among XDP patients was the region containing the CCCTCT hexamer repeats. My project will concern the SVA retrotransposon, with added focus on these hexamer repeats. The length of these repeats could help elucidate why the age of onset for XDP can vary so greatly.

Validating Mouse Models of Transient Lubricin Deficiency

Mara Coyan
Molecular and
Cellular Biology,
2017

Boston Children's
Hospital

Advisor: Matthew
Warman, MD

Mentor: Yajun Cui,
PhD

Increased participation in competitive sports has led to the rising incidence of traumatic joint injuries. These injuries can have lasting consequences on joint health; the mean arthritis-free interval after an anterior cruciate ligament tear is only about fifteen years. Thus, young athletes with this injury are at high risk for being burdened with disabling joint disease at a relatively young age. Effective clinical interventions that prevent joint injuries from progressing to joint failure are needed. Lubricin (encoded by *Prg4*), the principal boundary lubricant in articular cartilage, plays a key role in preventing cartilage wear. Genetic lack of lubricin causes precocious joint failure, and acquired transient deficiencies of lubricin occur in patients following acute joint injury. It is not known if transient lubricin deficiency in a previously normal joint causes joint damage and, if so, whether this damage is irreversible and progressive.

The objective of the present study is to determine whether and how cartilage becomes damaged in the context of acquired lubricin deficiency. To address these questions, I am using mice with a conditional lubricin knockout allele ($Prg4^{Frt}$); the first two exons of *Prg4* are flanked by flippase recognition (Frt) sites, enabling the allele to behave like a wild-type allele until flippase activity renders it nonfunctional. Another mouse allele ($Rosa26^{FlpER}$) permits the temporal induction of flippase activity; the protein is not active until it is bound to tamoxifen. I expect mice with the genotype $Prg4^{Frt/-}$; $Rosa26^{FlpER/+}$ to have adequate lubricin expression until they receive tamoxifen, after which lubricin expression will cease. I intend to administer a ten-day course of tamoxifen to these mice when they become skeletally mature (i.e., two months old) and then monitor their development and progression of cartilage damage.

The Warman lab has already confirmed that the $Prg4^{Frt}$ allele functions as a wild-type allele in the absence of flippase and as a knockout allele after flippase-mediated recombination. Critical to my experiments is that the $Rosa26^{FlpER}$ allele is not enzymatically active in the absence of tamoxifen, and that it is highly efficient at recombining the $Prg4^{Frt}$ allele when tamoxifen is present. I am as-

sessing the "leakiness" of the $Rosa26^{FlpER}$ allele by taking 2-month old $Prg4^{Frt/+}$ $Rosa26^{FlpER/+}$ mice that have not received any tamoxifen, and performing droplet-digital PCR on DNA extracted from their joints to measure recombination at the $Prg4$ locus. I will repeat this process for mice that have received a 10 day intraperitoneal injection program of tamoxifen to assess the the efficiency of flippase activity. If I demonstrate that the $Rosa26^{FlpER}$ allele is not "leaky" in the absence of tamoxifen and highly efficient when it is present, then I will be in an excellent position to temporally inactivate lubricin expression in juvenile and adult mice, and then determine the consequences on joint health.

Metabolomic Profiling of Chronic Kidney Disease and Type 2 Diabetes in the Jackson Heart Study

Mary Wan
East Asian Studies, 2019

Beth Israel Deaconess Medical Center

Advisor: Robert E. Gerszten, MD

Mentor: Jordan Morningstar

Chronic kidney disease (CKD) is caused by various factors including type 2 diabetes (T2D) and high blood pressure, which also contribute to cardiovascular disease. Although studies have shown that African Americans exhibit higher risk of developing CKD than caucasians, the biological basis of CKD development has yet to be fully understood. Occurrence of CKD and T2D depends on a combination of numerous genetic and lifestyle variations such as body mass index (BMI) and blood glucose levels; the large number of possible factors complicates research in the field. However, we hypothesize that metabolomics, which is the study of small molecules, or metabolites in human cells, tissues and biofluids, may help to highlight novel mechanistic underpinnings of CKD and its connection to T2D as well as to find novel biomarkers of both diseases. In order to do so, we will profile metabolites in human plasma from participants in the Framingham Heart Study (FHS) as well as the Jackson Heart Study (JHS) and, because the samples were obtained in 2002, we will identify metabolites that presage the clinical diagnosis by over a decade. Additionally, we hope that a metabolomics approach will provide insights on the high rates of CKD and T2D in African Americans.

The Jackson Heart Study comprises 5,302 African American participants from the Jackson, Mississippi area. Participants underwent a series of clinical measurements including cholesterol, glucose, insulin, and creatinine levels, and completed an extensive lifestyle questionnaire. To initiate a large pilot analysis, we selected 400 of these participants at random for metabolomic profiling. We used a liquid chromatography-tandem mass spectrometry (LC-MS/MS) platform that measures approximately 80 organic acids and phosphorylated metabolites in the negative ion mode. Data was reviewed and normalized to correct for instrument drift over the course of the run. After correction, metabolite levels were log transformed and compared to clinical phenotypes using a Pearson correlation test. Using metabolomic data previously gathered from the Framingham Heart Study, we performed a similar correlation of metabolites to clinical phenotypes to what was done in JHS. My primary contributions over this summer thus far have been to prepare the samples, extract the metabolites, perform LC-MS/MS, generate the data on the mass spectrometer, and integrate the metabolite peaks to translate them into quantitative information.

We hope to find associations of metabolites to phenotypes in JHS that are consistent with associations previously made in the FHS cohort. As an example, we have previously shown that uric acid is highly associated with a variety of metabolic risk factors including BMI, glucose, and HOMA-IR. Additionally, we hope to find novel metabolite correlations with the estimated glomerular filtration rate (EGFR), a measurement of kidney function performed in JHS, that may be inconsistent with FHS, indicating an African American-specific metabolic risk factor.

Directed Evolution of Ferritin for the Biomineralization of Heavy Metals

Michael Giles
Applied Mathematics, 2019

Wyss Institute for Biologically Inspired Engineering

Advisor: Pamela Silver, PhD

Mentor: Leon Liu

Ferritin is a cage-shaped protein that controls the concentration of free iron ions in the cell by oxidizing them into iron oxide crystals in its center. Ferritin's structure is relatively simple and well-characterized, and the ferritin family of proteins is well-conserved throughout an extensive lineage of both eukaryotes and prokaryotes. These features have made ferritin a promising platform for bioengineering and bionanotechnology, and many *in vitro* applications of the protein have been explored.

Less has been done with ferritin *in vivo*, and this summer we attempted to engineer *E. coli* by modifying the ferritin that the cells expressed. Previous experiments performed in the Silver lab have created *E. coli* possessing ferritin that crystallizes a larger number of iron ions than wild-type ferritin. Interestingly, these mutant bacteria have been shown to better survive poisoning by the toxic metals arsenic, cobalt, cadmium, and nickel. These metals react well with iron oxide in an inorganic setting, and we surmised that something similar could be occurring within the cell—that the modified ferritin proteins were able to biomineralize these metals, co-crystallizing with the iron oxide inside of the modified ferritin more than in the wild-type ferritin.

We attempted to take advantage of this observed effect by engineering ferritin to extract ions of these metals from the cells' environment. We approached this problem using a technique called directed evolution: organisms are mutated and placed in a selective environment that favors the survival of a particular trait. Mutants possessing favorable traits reproduce more than less fit individuals. Over one or more rounds of selection, the most successful traits are favored, and the mutations that led to this increased survival can then be characterized.

We applied directed evolution to ferritin by using a technique known as mutagenic PCR to create many copies of ferritin gene, each with an average of one DNA base changed. We forced many *E. coli* cells to express a copy of these mutant ferritins, and plated them on agar supplemented with one of the four metals of interest at a concentration lethal to bacteria expressing unmodified ferritin. We picked the healthiest colonies of bacteria from these plates and sequenced their ferritin genes to find the mutations of the most successful survivors. We verified that our method produced the desired outcome using mass spectrometry, as well as four novel biosensors, which were constructed using promoters that have been previously shown to be responsive to our four metals of interest. We designed plasmids including these promoters so that they drove the expression of green fluorescent protein (GFP), allowing us to quantify the intracellular concentration of these metals by measuring fluorescence.

Successful mutants will be ones that sequester significantly more metals than the wild type. With future development, these could be useful in remediating water that has been contaminated by these metals, as well as in mining. The biosensors we developed, provided they are reasonably sensitive and specific, could also be useful as a cheap and resilient way to test for heavy metal contamination.

Synergy Between *Myc* and PI3K Inhibitors

Muhammed Ors
Molecular and Cellular Biology, 2017

David H. Koch Institute for Integrative Cancer Research

Advisor: Angela Koehler, PhD

Mentor: Francisco Caballero, PhD

My thesis research is being done in the Koch Institute of MIT in the Koehler lab. The lab studies dysregulated transcription in cancer such as *c-Myc* overexpression, one of the most common aberrations in human cancers, having been linked to upwards of 70% of cancers. Being able to inhibit aberrant transcription factors would be instrumental, as overexpression of even one transcription factor, especially a master regulator like *c-Myc*, leads in turn to overexpression of any number of the proteins whose codes are downstream to the transcription factor. Therefore, they make a smart therapeutic target in cancer research. Importantly, mouse models have shown that *c-Myc* inactivation stereotypically results in the proliferative arrest, differentiation, and/or apoptosis of tumor cells. Thus, I believe any step forward in selectively modulating *c-Myc* expression is synonymous with a step forward in tackling cancer. Unfortunately, for the most part, transcription factors lack typical binding pockets and have been considered "undruggable." Of course, there is hope: the Koehler lab has been finding and optimizing small molecule inhibitors of *c-Myc*, allowing me the incredibly unique opportunity to experiment with something completely novel.

My research is focusing on the synergistic effect in breast and brain cancer cell lines between the inhibitors of *c-Myc* and PI3K, a family of kinases (phosphoinositide 3-kinase) which are referred to as PI3K in shorthand. Previous research using RNA interference (RNAi) of *c-Myc* has shown that resistance to PI3K inhibitors is dependent on *c-Myc* induction. Being unable to inhibit *c-Myc* directly, another group decided to do so indirectly by inhibiting BRD4, a protein with a role in the same pathway as PI3K and *c-Myc*, using a BET inhibitor called JQ1. The synergy of BET and PI3K inhibitors produced a sustained inhibition of PI3K and tumor cell death in twenty cell lines.

My methods for research can be partitioned into three phases. The first aim is to use co-immunoprecipitations and western blotting to establish and verify that the *c-Myc* inhibitors were inhibiting the Myc/Max/DNA interaction. The second aim is to do cell viability assays in multiple cell lines using the PI3K, BET, and the Myc inhibitors individually in serial dilutions from 40 μM to 0 μM/DMSO in order to generate dose-dependent curves, and additionally to do western blotting and confirm a reduction in protein expression due to increasing concentration of inhibitors. Finally, the third aim will be to use yet-to-be-developed synergistic cell viability assays to contrast between the conditions of PI3K/BET inhibitors and the PI3K/Myc inhibitors.

My hypothesis is that the synergistic effects of the PI3K and *c-Myc* inhibitors are better than those of the PI3K and BRD4 inhibitors. Namely, this would be a smaller IC50, the concentration of the drug required for there to be 50% inhibition *in vitro*.

Odd-Skipped Related 1 in Female Reproductive Tract Development

Naomi Asimow
Chemical and Physical Biology, 2018

Brigham and Women's Hospital

Advisor: Ursula Kaiser, MD

Mentor: Adriana Lofrano Porto, MD, PhD

Our project began with three Brazilian sisters, ages 33, 27, and 26, born of a consanguineous marriage, each presenting the same novel phenotype. All three displayed primary amenorrhea, but otherwise normal secondary sex characteristics. MRI revealed thin endometrium and uterine hypoplasia. The sisters had normal hormonal levels, but endometrium that proved unresponsive to sex steroids. Intriguingly, two of the sisters had been treated for ectopic tubal pregnancies. Whole exome sequencing (WES) analysis followed by array-CGH and homozygosity mapping revealed two candidate genes in which the affected sisters carried novel homozygous mutations: *OSR1* and *HS1BP3*, of which Odd-Skipped Related 1 (*OSR1*) is the more likely candidate based on in silico analysis. *OSR1* encodes a zinc finger transcription factor which has previously been studied for its role in the development of the heart and kidney. It is expressed in the intermediate mesoderm during embryogenesis. Our project focuses on the role of OSR1 protein in the Müllerian duct, the embryonic structure which differentiates from intermediate mesoderm into the female reproductive tract (FRT), forming the fallopian tubes, uterus cervix, and the upper vagina. Our hypothesis is that OSR1 plays an important role in differentiation of the intermediate mesoderm into the Müllerian duct, such that mutations to OSR1 could explain the FRT anomaly seen in our patients. We are using both *in vivo* and *in vitro* studies to investigate the role of OSR1 in intermediate mesoderm differentiation.

Global *Osr1* knockout mice have previously been shown to die early during embryogenesis. Therefore, to investigate OSR1 protein in the development of the FRT, our mouse model had to be a conditional *Osr1* knockout, where *Osr1* is only knocked-out in the Müllerian duct and is expressed normally elsewhere. To generate our conditional mouse, we used the *Cre-Lox* recombination system. *LoxP*-flanked *Osr1* genes generated a homozygous *Osr1-Floxed* female mouse. In males, the *Cre* gene was inserted into the promoter of a Müllerian-expressed *Wnt7a* allele. OSR1 knockout was thus limited to the Müllerian duct in F2 generation individuals homozygous in *Osr1-Floxed* and hemizygous in *Cre* allele.

We are also looking at the interactions of the OSR1 protein with other proteins, such as HOXA10. We transfect human embryonic kidney (HEK 293) cells with plasmids containing the sequences for human *OSR1* and *HOXA10* respectively. Using western blot and co-immunoprecipitation experiments, we analyze interactions of OSR1 with HOXA10 and with other proteins.

Finally, we aim to examine expression patterns of *Osr1* in female mouse embryos using immunohistochemistry. We use 13.5-day-old mouse embryos with a TY1 tag on OSR1. We bind the anti-TY1 antibody to OSR1 protein and then a fluorescent secondary antibody to anti-TY1 so that we have fluorescence anywhere OSR1 is expressed. We are looking for fluorescence in the Müllerian duct and/or surrounding mesenchyme.

Female reproductive tract abnormalities are one of the primary causes of infertility. Understanding the role of OSR1 and its interactions with other proteins in FRT development may provide insight into many types of FRT abnormalities, such as Mayer-Rokitansky-Küster-Hauser syndrome.

Mechanisms of Tumor Resistance to TIL Therapy and PD-1 Inhibition in Melanoma

Ramtin Talebi
Molecular and
Cellular Biology,
2018

Dana-Farber Cancer
Institute

Advisor: Levi
Garraway, MD, PhD

Mentor: Shruti
Malu, PhD

In the past few years, immunotherapies such as tumor infiltrating lymphocytes (TILs) and checkpoint receptor blockade (such as anti-PD1) have demonstrated that durable responses can be obtained in patients with advanced cancers. However, the majority of patients still fail to respond or progress after an initial response, creating the need for combination therapies. Moreover, understanding the resistance mechanisms that the tumors impose can help us define better biomarkers of response to immunotherapy. This will be beneficial in predicting which patients might respond to immunotherapy as well as guiding them into the effective combination therapies for their own tumor. Thus, we have recently performed whole exome sequencing and RNA sequencing of tumors from melanoma patients that underwent TIL based immunotherapy and are using this data to understand novel resistance mechanisms to T cell based therapies.

In TIL therapy, autologous T cells that have infiltrated a patient's tumor are removed from the harsh tumor microenvironment, cultured and grown in large numbers *in vitro* in presence of interleukin-2 (IL-2), a cytokine essential for specific growth of T cells. These cells are then re-administered to the patient, leading to an anti-tumor immune response. The first part of this project will focus on studying the genes that appear significantly enriched in tumors of patients who do not respond to TIL therapy as compared to the tumors of patients who respond well. In order to effectively identify the genes potentially responsible for tumor resistance to this adoptive cell therapy, we have performed whole exome sequencing and RNA sequencing the patient tumor samples. The hypothesis is that overexpression of these genes will lead to resistance to T cell-mediated killing and knockout of these genes will make the cell lines more sensitive to T cell-mediated death. To test this, we will overexpress these genes in human melanoma cell lines and perform T cell cytotoxicity assays. Using CRISPR technology, we will knockout these genes from the melanoma cell lines and perform similar assays. Finally, we will study the mechanism of action of genes that are shown to be mediators of resistance to T cell killing by utilizing techniques of molecular biology and protein analyses.

The second part of the project will analyze another highly promising immunotherapy, anti-PD1 (an immune checkpoint receptor), which has been recently approved for both melanoma and lung cancer. Despite its promise, many patients still fail to respond to anti-PD1 therapy, and some research suggests the cause to be an up-regulation of alternative immune checkpoints like TIM-3 or LAG-3 in CD8$^+$ T cells. However, many mechanisms of tumor resistance to anti-PD-1 therapy remain unknown. Thus, using anti-PD-1 antibody *in vitro*, we will improve the cytotoxicity of TILs on the melanoma cell lines and develop this system for large scale genomic screens to identify novel resistance mechanisms for the joint therapy. In the long term, these findings may lead to the discovery of combination therapies designed to anticipate tumor resistance and create a more durable and positive patient response for those with advanced melanoma.

Characterization of Intron Elements that Control Splicing in Unc-16 Gene

Santiago Vargas
Chemistry and
Physics, 2019

Harvard University

Advisor: John
Calarco, PhD

Mentors: Adam
Norris, PhD; Xico
Gracida, PhD

In eukaryotes, the process of precursor mRNA (pre-mRNA) splicing involves the removal of non-protein coding segments called introns and the stitching back together of coding segments called exons. These segments are cut and stitched by a piece of cell machinery known as the spliceosome. Additionally, exons themselves can be selectively included or skipped in a process known as alternative splicing, which allows for a greater degree of protein diversity from a more limited repertoire of genes. A gene can code for various different proteins depending on splicing patterns and this allows for more diversity without necessarily increasing the genetic complexity of an organism's DNA. Splicing complexity has been tied to the complexity of an organism itself. Humans display a large amount of alternative splicing and this allows for complicated and differentiated cell function within our bodies without the need for exceedingly

large sets of genes. Understanding splicing events is important. For example, splicing characterizes how certain human mental disorders are caused on a genetic level.

We know that certain DNA motifs upstream or downstream of alternatively spliced exons can serve as up-regulators or down-regulators of splicing events. These elements are known as cis elements and they are believed to be recognized by trans factors within pre-mRNA and this will either increase or decrease the likelihood of splicing. Trans factors are proteins that are coded for elsewhere in the DNA that are made specifically to identify these splicing control sequences. The model organism we are studying is *C. elegans*, as it has a simple, well-characterized nervous system, has known splicing, and has DNA that is relatively easy to manipulate with exogenous fragments.

Two interesting cis elements had been previously found in the *unc-16* gene that stimulate exon inclusion in neuronal cells. To search for more sequences that can regulate splicing, my work this summer started with analysis of high-throughput sequencing data of cDNA collected from animals expressing a library of splicing reporter transgenes containing random 6-base pair sequences (testable candidate cis elements). I developed efficient computational tools in order to sift through this data and to quantify splicing patterns and test if splicing ratios changed from wild-type and mutant benchmarks depending on what hexamer sequence was inserted into a given reporter. Positive results would include finding hexamer (6-base pair segments) sequences that greatly affect the ratio of splicing in the animal. Other possible results could be the characterization of other, previously uncharacterized splicing regulatory elements. I then used this data to create oligomer sequences (DNA segments of interest) that I later inserted into *C. elegans* to experimentally confirm our computational data. This data could then be used to create an evolutionary comparison to detect identical regulatory elements in other species of nematodes and possibly other clades of organisms. My work this summer also included the adaptation of a fluorescent reporter into a format that could be easily used in *C. elegans.* This reporter glows a different color depending on the splicing pattern occurring in a certain tissue. I used this reporter in the experimental trial of our hexamer sequences of interest.

Identification of C3Orf59/*MB21D2* as a Novel Cancer Gene

Scott Xiao
Chemistry, 2019

Broad Institute of MIT and Harvard, Dana-Farber Cancer Institute, Harvard Medical School

Advisor: Matthew Meyerson, MD, PhD

Mentor: Douglas Wheeler, MD, PhD

Recent genomic sequencing efforts by the Meyerson lab of lung adeno and squamous cell carcinomas have identified several novel oncogenes and tumor suppressors. Part of the efforts determined common neoepitopes, or recognizable antigen segments, of lung cancer cells, which included well-documented cancer genes. This analysis found that C3orf59 (the 59th open reading frame of Chromosome 3), also known as *MB21D2*, is frequently mutated at residue 311 from a glutamine to a glutamate (also known as p.Q311E). Interestingly, we find that this mutation frequently and significantly co-occurs with activating mutations in the oncogene *PIK3CA* in lung cancers and other tumors. Our goal is to explore the potential of *MB21D2* as a novel cancer gene and to identify the function of this gene in tumorigenesis.

MB21D2 is an unstudied gene which is tightly conserved from zebrafish to humans. The *MB21D2* protein contains a Mab-21 domain, which is also found in a number of other proteins involved in developmental processes but whose function is poorly understood. Although the function of *MB21D2* is unclear, phosphoproteomic studies exploring the importance of *MB21D2* phosphorylation have identified it as a substrate of the AGC kinase family, a group which includes protein kinase A, C, and AKT. These data suggest that this protein may be phosphorylated downstream of cellular signal transduction pathways.

To study this uncharacterized gene, we first identified cell lines with high and no expression of *MB21D2* . By immunoprecipitating MB21D2, we found that the cell lines HeyA8, BT-474, and NCI-H2009 expressed high levels of MB21D2 while the cell line NU-DUL-1 did not express MB21D2 at all. We have used these cell lines to validate antibodies and cDNA reagents towards MB21D2. Immunoprecipitation of MB21D2 from high expression and no expression cell lines has revealed several interacting proteins, which indicates that MB21D2 likely exists in one or more protein complexes. Interestingly, two cancer cell lines, BT-474 and NCI-H2009, contain the p.Q311E mutation.

Because MB21D2 is a phosphoprotein, we asked whether it was downstream of signal transduction

pathways. We treated high expression cells with a variety of kinase inhibitors, and found that inhibition of protein kinase A (PKA) reduced the phosphorylation of MB21D2, which indicates that this protein lies downstream of PKA signaling. We also used the clustered regularly interspaced short palindromic repeats (CRISPR-Cas9) system to disable the *MB21D2* gene in several cell lines and asked whether loss of MB21D2 function altered cellular signaling in any way. We found that functional inactivation of MB21D2 diminishes phosphorylation of certain PKA substrates such as glycogen synthase kinase 3 (GSK-3) and cAMP response element-binding protein (CREB). These data further suggest that MB21D2 may have some regulatory function within the PKA signal transduction pathway.

Finally, to define the role of *MB21D2* in oncogenesis, we plan to manipulate this gene *in vivo* and *in vitro*. We are in the process of constructing vectors to conduct these experiments. In the meantime, we are also trying to better understand the connection between PKA and MB21D2, and whether such a link would have implications for *MB21D2* as a novel cancer gene.

A Computational Tool for Classifying Nervous System Cell Types Using mRNA Expression Profiles

Victoria Lin
Molecular and Cellular Biology & Statistics, 2017

Harvard Medical School, Broad Institute of MIT and Harvard

Advisor: Steve McCarroll, PhD

Mentor: Evan Macosko, PhD

In healthy organisms, genes are expressed at different levels in different cell populations; it is this property that allows for the functional specialization of various cell types to form tissues and organs. When an organism is afflicted with a genetic disease, another kind of differential gene expression may appear, as disease-risk alleles can result in abnormal expression only in certain subsets of cells. As such, a comprehensive characterization of cell types in the nervous system would not only provide insight into the function of a complex network, but also contribute to a greater understanding of the pathology of neurological disease. Despite the ongoing categorization of such cells by location, morphology, target specificity, electrophysiological properties, and other molecular markers, however, the complexity of the brain has thus far limited attempts to fully identify the diversity of cells that it comprises.

One metric by which cells can be simply grouped is through their mRNA expression profiles, since these transcripts in large part dictate the cells' function. Large-scale classification following this method has been limited in the past by the fact that RNA-seq at the single-cell level can only be achieved for small batches of cells, as transcripts must be sequenced individually for each cell. Recently, however, Macosko et al. introduced the Drop-seq method of massive single-cell mRNA expression profiling, in which individual cells are encapsulated in nanoliter droplets containing a bead bearing a distinct molecular barcode, allowing all transcripts to be sequenced in parallel and subsequently matched to their cell of origin. With transcriptional profiles now available for tens or even hundreds of thousands of neural cells, it is possible to use computational methods to group cells into distinct classes based on biologically significant factors.

We are currently developing a tool that will use Drop-seq data from brain tissue to help discover previously unknown subtypes of known cell classes in the nervous system; subsequent identification of enriched genes for each of these subtypes both can serve as markers and can provide an idea of their functional diversity. We use principal component analysis (PCA) on mRNA expression data from a single cell type to explain variation in the data through a small number of statistically significant variables, then map data points in two dimensions using the t-distributed stochastic neighbor embedding (t-SNE) method of nonlinear dimensionality reduction. Following dimensionality reduction, we will cluster and classify the cells into subpopulations using the k-nearest neighbors (k-NN) algorithm, in which the class of a cell is determined by a majority vote of its k nearest neighbors; then extract, through differential expression analysis, enriched genes from each group, both to serve as markers and to provide an idea of the function of the subtype. These results from such a systematic examination will contribute to a more thorough understanding of the diversity of cell populations in the nervous system, lend insight into how cell types specialize by tissue, and allow for the pinpointing of both healthy and disease-related phenotypes to more specific classes of cells.

The Role of PDK4 in Regulating Th17 Cell Pathogenicity

Zhuang Miao
Chemical and
Physical Biology,
2019

Brigham and
Women's Hospital

Advisor: Vijay
Kuchroo, PhD

Mentor: Chao Wang,
PhD

T helper 17 (Th17) cells are $CD4^+$ T cells that produce the cytokine interleukin-17 (IL-17) and can exhibit distinct phenotypes. Th17 cells are found in the intestine of healthy mice and man, maintaining mucosal barriers and defending against extracellular pathogens. However, Th17 cells are also found at sites of inflammation, playing a pathogenic role in inducing autoimmune diseases such as multiple sclerosis, inflammatory bowel disease, and rheumatoid arthritis. These findings raised an interesting hypothesis that not all Th17 cells are created equal. Research on the mechanisms controlling the fate of Th17 cells will allow precision targeting of pathogenic Th17 cells and block their differentiation in autoimmune diseases, while sparing the non-pathogenic Th17 cells that are beneficial for health.

A recent study in my lab shows that Th17 cell pathogenicity is correlated with changes in cellular metabolism. Analysis of the Th17 cell transcriptome revealed that genes involved in central carbon metabolism show differential expression in pathogenic and non-pathogenic Th17 cells. Furthermore, our metabolome study shows that non-pathogenic Th17 cells have a higher ratio of polyunsaturated fatty acids (PUFA) to saturated fatty acids (SFA), suggesting altered lipid metabolism. These observations lead to some questions: what is the key regulator of these metabolic changes, and can it drive Th17 pathogenicity? We hypothesize that metabolic genes at the junction of glycolysis and fatty acid metabolism could contribute to Th17 cell function. I chose to focus on a mitochondrial enzyme, pyruvate dehydrogenase kinase 4 (PDK4), which inhibits the conversion of pyruvate to acetyl-CoA, promotes glycolysis, and inhibits fatty acid synthesis. Moreover, PDK4 is a nutrient sensor: its expression is turned on in response to alterations in lipid metabolism, which might then drive Th17 cell pathogenicity.

To study the role of PDK4, I used mice with targeted deletion of $PDK4$ ($PDK4^{-/-}$) and determine the effect of the loss of PDK4 on Th17 cell pathogenicity with several approaches. I cultured naive T cells isolated from wild-type and $PDK4^{-/-}$ mouse spleens with different cytokines to differentiate them into pathogenic or non-pathogenic Th17 cells. Furthermore, I chose to culture some of the cells with PUFA to investigate how PDK4 knockout affects Th17 cells with lipid environmental cues. Specifically, I measured the changes in the expression of genes that were previously identified to play a critical role in regulating Th17 cell pathogenicity by using the following techniques: 1) flow cytometry to measure protein expression at the single cell level, 2) ELISA (enzyme-linked immunosorbent assay) to measure cytokines secreted by Th17 cells, and 3) qPCR and nanostring to quantify to determine gene expression at the transcriptome level.

Because the function of PDK4 to upregulate glycolysis is consistent with increased glycolysis in pathogenic Th17 cells, I initially hypothesized $PDK4^{-/-}$ cells would be less pathogenic. However, my earlier experiments showed higher expression of some pro-pathogenicity proteins in the $PDK4^{-/-}$ cells, suggesting PDK4 plays a more nuanced regulatory role in Th17 cells. I find these unexpected results especially interested and will further investigate the function of PDK4 in Th17 pathogenicity.

Elucidating the Resistance Mechanism of Medulloblastoma to BET Bromodomain Inhibition

Bluyé DeMessie
Molecular and
Cellular Biology,
2019

Dana-Farber Cancer
Institute, Broad
Institute of MIT and
Harvard

Advisor: Rameen
Beroukhim, MD,
PhD

Mentors: Pratiti
Bandopadhayay,
MBBS, PhD; Ryan
O'Rourke

Medulloblastoma is a high-grade, fast-growing tumor, and is the most prevalent malignant brain tumor in children. Surgical removal, chemotherapy, and radiation are ineffective and damage the adjacent tissue. A pharmacological approach to target overexpressed genes is being explored by our lab. The gene MYC is overexpressed in patients with a poor prognosis and the transcription of MYC is regulated by the bromodomain and extra-terminal domain (BET) family of proteins (e.g. BRD4). The BET bromodomain regulates gene transcription by binding to acetylated histone tails.

JQ1, a BET bromodomain inhibitor, has demonstrated high activity towards $BRD4$ by displacing it from acetylated histones. Medulloblastoma cells treated with JQ1 have a sharp reduction in viability and proliferation, but eventually develop resistance

to JQ1 and other BET bromodomain inhibitors. The aim of this work is to understand the mechanism by which cells become resistant to JQ1, and to determine the role of *BRD4* as an essential gene in resistant models.

To determine the JQ1 resistance mechanism, it is necessary to identify the proteins and overexpressed genes in resistant cells which allow the cells to proliferate. We performed a genome-wide open reading frame (ORF) screen to identify genes that confer resistance when overexpressed. Medulloblastoma lines that are resistant to BET bromodomain inhibitors were generated by chronically passaging cell lines in drug until they exhibited growth. The *in vivo* behavior of resistant cells was studied using cell line based xenograft models. Western immunoblots were performed to study expression of candidate resistance proteins in resistant lines. Protein-protein interaction was studied by using immunoprecipitation to extract proteins bound to BRD4. These proteins were then identified using mass spectroscopy.

We generated resistant cell lines by chronically passaging sensitive cells in BET bromodomain inhibitors. We next confirmed that the resistant lines also exhibited resistance *in vivo*. Following flank injection of both sensitive and resistant cells, we observed the resistant xenografts to proliferate in both vehicle and JQ1 treated mice. Mice harboring xenografts from the sensitive lines exhibited an initial response, however we observed the eventual development of resistance.

Affinity purification mass spectroscopy of the immunoprecipitated BRD4, and its related proteins in resistant cells, revealed specific proteins binding to BRD4. These results demonstrate that BRD4 may contribute to resistance through this domain.

This study has identified a multitude of overexpressed genes involved in developing resistance to BET bromodomain inhibitors. We have determined that *BRD4* is an essential gene in resistance to BET bromodomain inhibitors. We have identified which proteins BRD4 binds to in resistant cells. Future work involves identifying targetable proteins in order to create a drug treatment for BET bromodomain inhibitor resistant medulloblastoma.
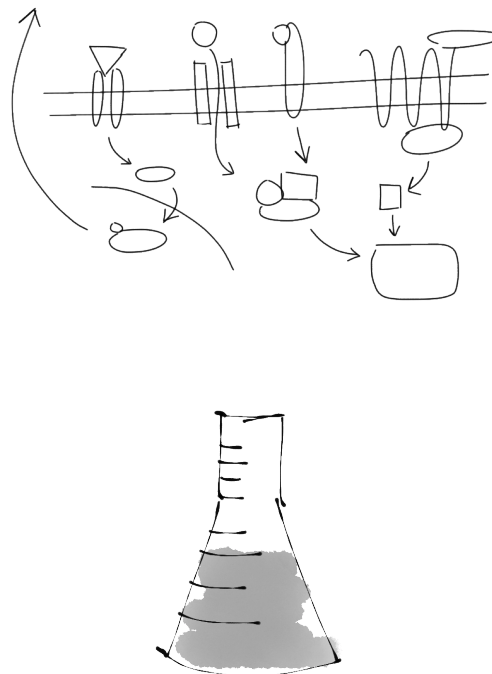
Early Biomarkers for Language Development in Infant Siblings of Children with Autism Spectrum Disorder

Adriana Mendez Leal
Neurobiology, December 2016

Boston Children's Hospital

Advisor: Charles Nelson, PhD

Mentor: April Levin, MD

Early intervention programs for autism spectrum disorder (ASD) have shown that earlier enrollment in treatment is associated with better long-term outcomes. Though ASD can be diagnosed in controlled clinical environments in infants as young as 12 months, underlying neurological changes likely begin before symptoms emerge. Recent evidence suggests there is altered brain development in children 2–6 months of age who later go on to receive ASD diagnoses, pointing to an opportunity for the isolation of a predictive biomarker for ASD available well before diagnosis using behavioral measures is possible. Such a biomarker could provide insight into the neuropathology of ASD, and might also allow for treatment earlier in the progression of the disorder, even before the onset of behavioral symptoms.

Because autism is a heterogeneous disorder, comparative analysis of EEG data from children with ASD and typically developing children is frequently limited by inter-subject variability in signal in the ASD group. However, evidence that intra-subject variability in response to a repeated stimulus is higher in children with ASD than in typically developing controls has led some researchers to suggest intra-subject variability may itself be a viable candidate for a predictive biomarker for the disorder. Furthermore, infant siblings of children with autism are twenty times more likely to develop ASD than their typically developing counterparts, and as a result they are well suited to prospective cohort studies investigating potential neural markers of ASD.

My project is a subset of the Infant Sibling Project, a study that compares neurological and behavioral development in a group of infant siblings at high risk for autism (HRA) with that of a group of low risk controls (LRC). Specifically, my project will compare intra-subject variability in EEG response to native and non-native language in 6 month old infants (as measured by inter-trial alpha phase coherence) to receptive and expressive language behavioral scores at 18 months in the HRA and LRC groups. Investigation of the relationship between EEG response variability at early ages and later developmental outcomes may establish language-specific biomarkers, and may ultimately provide insight into the viability of variability as a biomarker for ASD.

Additionally, comparing these relationships to patterns between EEG response and eventual diagnosis at 36 months will give the opportunity for differentiation between markers of atypical development or delay and ASD-specific diagnostic markers. It's possible that EEG variability is tied to general disruption of connectivity, rather than ASD itself, and examining the relationship of early EEG to behavioral outcomes will help elucidate any differences.

Investigation Into the Role of CA3c, CA2, and the Entorhinal Cortex in Pattern Separation and Completion in Spatial Memory Encoding

Alvar Faustino Paris
Medicine, 2017
Emmanuel College

Massachusetts General Hospital, Harvard Medical School

Advisor: Amar Sahay, PhD

Mentor: Kathleen McAvoy, PhD

Pattern separation is the computational process by which similar inputs are made more distinct during storage so as to minimize interference between them, while pattern completion relies upon neurons retrieving stored output patterns in a way that completes degraded or partial input signals. In the hippocampus, these phenomena respectively allow the brain to differentiate between similar spatial contexts, and to evoke a memory of an entire context from a more limited stimulus.

In the past, theoretical and lesion studies have been used in order to elucidate the roles of the subdivisions of the hippocampus. These studies have generally reinforced a model in which population-based encoding in the dentate gyrus underlies pattern separation, and area CA3 may support pattern separation or serve an autoassociation function to support pattern completion. Meanwhile, other regions within the hippocampus remain understudied, partly owing to their perception as functionally continuous with anatomically nearby areas. Such regions include CA2 (which has often been considered to be transi-

tional between CA3 and CA1) and CA3c (which has commonly been grouped with the neighbouring area CA3ab).

Recently, however, re-examination of these areas has highlighted that they are likely to be functionally unique from their anatomical neighbors. For example, CA2 has been found to be heavily involved in social memory, and studies have shown that CA3c-lesioned animals show greater deficit in spatial processing than CA3ab-lesioned animals. Thus, legitimacy has been brought to the hypothesis that CA3c and CA2 may have distinctive roles regarding pattern separation and completion.

To test this hypothesis, we investigate these subregions using catFISH (cellular compartment analysis of temporal activity by fluorescence *in situ* hybridization). We do this in both control mice and in mice where adult hippocampal neurogenesis has been enhanced, which has been shown to improve pattern separation. These mice are exposed to two spatial contexts in a temporally controlled manner before being sacrificed. By visualizing the subcellular location of *cfos* mRNA, for which transcription is upregulated when a neuron becomes active, we can then calculate the population overlap of cells active in the two separate spatial contexts. In turn, we use this as a measure of similarity in encoding, which—in conjunction with data from connected areas—provides information about whether pattern separation or completion is being supported.

To isolate the subregions in question, we define the anatomical boundaries of area CA3c and study the region within the blades of the dentate gyrus, before comparing it to CA3ab. Second, we isolate area CA2 by using *in situ* hybridization staining with RNA probes targeted against the locally expressed genes *rgs14* and *cacng5*. We also expand our investigation to the entorhinal cortex (the main input area to the hippocampus) in order to study the complete pathway of spatial context encoding.

Together, these findings will resolve in greater detail the roles that precise anatomical regions have in pattern completion and separation, as well as produce a more complete picture of information flow through the entorhinal cortex and hippocampus.

Understanding Microglial Activation in a Multiple Sclerosis (MS) Model

Audrey Effenberger
Neurobiology, 2019

Boston Children's Hospital

Advisor: Zhigang He, PhD

Mentor: Jing Wang, PhD

Multiple sclerosis (MS) is a demyelinating disease characterized by lesions in the central nervous system (CNS). These scars predominantly affect white matter, which is composed of myelinated nerve fibers, or axons, crucial to information transmission. However, the underlying cause of these lesions remains unknown, as does a comprehensive treatment.

While the nervous system is naturally capable of limited regeneration, MS lesions fail to heal over the course of the disease due to dysfunction of a key class of supporting cells: oligodendrocytes (OLs). OLs build the myelin sheaths in the CNS that are crucial to neuron function; however, they often die after injury, and are progressively depleted in MS. Though nearby oligodendrocyte progenitor cells (OPCs) can proliferate and migrate towards lesions, this endogenous response is not enough to prevent net OL loss. Without a sufficient OL population, substantial remyelination fails to occur. The exposed axons are thus nonfunctional and vulnerable to further degeneration.

Additionally, MS is characterized by inflammation, an immunological response that may contribute to further breakdown of myelin sheaths and axons. In the CNS, innate immune cells known as microglia are partly responsible. Active microglia perform phagocytosis, the uptake of particles, and are responsible for initial inflammatory response. However, not all active microglia are the same. Whereas M1 microglia promote "classical" inflammation and damage tissues, M2 microglia decrease inflammation and promote tissue repair. Natural skewing of the M1/M2 balance may compound the inability of the CNS to heal or impact OL response in unknown ways.

These features of MS pathology motivate our research in an optic nerve injury model that shares pathological similarities with MS: inflammation and remyelination failure. Because microglia have historically received little scientific attention, we seek to understand their role in demyelinating diseases like MS, as well as their relationship to OLs.

To understand how OL response is influenced by mi-

croglia, we have created a mouse optic nerve injury model where the microglial response is dampened. By virally introducing Cre-inducible diphtheria toxin receptor (DTR) into the eyes, we have generated mice where the eyes can be specifically ablated by treatment with diphtheria toxin. This localizes damage to the nerve fiber and foregoes more invasive injury methods that would trigger a full inflammatory response. Based on preliminary immunofluorescence staining of microglial cell markers, we know that microglia are present and active to a reduced extent in DTR mice.

From this model, we can begin to see if there is a causative relationship between microglial activity and OPC proliferation or differentiation. We stain tissue samples for Olig2 markers that correspond to OL lineage cells, allowing us to determine if OPCs are surviving and proliferating in the nerve one or two weeks after it has been killed. CC1 markers of mature OLs can further reveal if these cells are differentiating. If microglia play a role in post-injury remyelination and neural repair, we hope to see significant differences in the number and distribution of OPCs and OLs between normal and DTR mice.

Furthermore, we are investigating the impact of various drugs on *in vivo* microglial subtype determination. By treating mice with compounds such as retinoid receptor agonists and staining for subtype-specific markers such as CD206 and CD16/32, we can determine if these drugs shift the balance of microglia towards the M2 state. Our results should lay the foundations for future investigations and applications, from more effective MS treatments to more precise anti-neuroinflammatory drugs.

The Role of 5-HT$_{3A}$R Cells in Brain Plasticity

Emma He
Neurobiology, 2019

Harvard University

Advisor: Takao Hensch, PhD

Mentor: Anne Takesian, PhD

During development, the brain undergoes critical periods of heightened plasticity, when the neural circuits are most sensitive to external stimuli. Ongoing work in the lab has shown that a group of inhibitory interneurons that reside in cortical layer 1 (L1) and express ionotropic 5-HT$_{3A}$ receptors (5-HT$_{3A}$R cells) play an important role in the enhanced plasticity during early life. These 5-HT$_{3A}$R cells inhibit parvalbumin-expressing (PV) cells, another inhibitory cell subtype that is known to underlie experience-dependent circuit rewiring. However, it is unknown whether 5-HT$_{3A}$R cells change during periods of heightened plasticity. Our goal is to study the development of 5-HT$_{3A}$R cells within the mouse primary auditory cortex. There are two main areas of focus for our project: 1) we will determine how 5-HT$_{3A}$R cell morphology and connectivity changes across postnatal development, and 2) we will examine how early experience alters 5-HT$_{3A}$R cells.

To examine the development of 5-HT$_{3A}$R cell morphology, we are employing Brainbow, a genetic labeling technique using multicolor fluorescent proteins, to visually distinguish and reconstruct 5-HT$_{3A}$R cells in 5-HT$_{3A}$R-Cre mice. Combining this with immunocytochemistry has allowed us to determine how developing and adult 5-HT$_{3A}$R cells form synaptic connections with other major inhibitory subtypes (PV and somatostatin-expressing) and excitatory pyramidal cell somata. We will also take advantage of an approach that makes brain tissue optically transparent and light sheet fluorescence microscopy for three-dimensional 5-HT$_{3A}$R cell reconstructions. We are particularly interested in how morphological changes in 5-HT$_{3A}$R cells correlate with transitions into and out of critical periods.

To determine how early auditory experience impacts 5-HT$_{3A}$R cells, we will passively expose 5-HT$_{3A}$R-Cre mouse pups and mothers to continuous white noise, band-limited noise, or tones in a sound-attenuating chamber. Continuous white noise has been shown to restore the tonotopic critical period. If 5-HT$_{3A}$R cells are associated with the plasticity of cortical response properties, we expect them to be rejuvenated in accordance with the restoration of the critical period. Previous studies have shown that repeatedly playing pure tones to mouse pups during a critical period enlarges the area of the auditory cortex representing the frequency of that tone. Finding that this manipulation alters the connectivity of 5-HT$_{3A}$R cells during the critical period would suggest that the plasticity of these cells underlies the experience-dependent development of auditory cortical circuits. Together, these studies will increase our understanding of how 5-HT$_{3A}$R cells contribute to neural development.

Investigating Age-Related Changes in the Neuronal Circuits Underlying Multisensory Processing in *Caenorhabditis elegans*

Gaia Linfield
Integrative Biology, 2018

Center for Brain Science, Harvard University

Advisor: Yun Zhang, PhD

Mentor: Gareth Harris, PhD

Humans are constantly experiencing multiple stimuli that must be processed and integrated within their brains. The ability to integrate information obtained from different sensory modalities is central to daily function, affecting the decisions we make and our perception of our surroundings. Some individuals who suffer from neurodevelopmental disorders, such as autism spectrum disorders, are impaired in their ability to perform multisensory processing. To better understand the complex signaling pathways and neuronal circuits that regulate multisensory behaviors in mammalian systems, we use the nematode *Caenorhabditis elegans* as a model system. *C. elegans* has a well-characterized nervous system, and advanced genetic tools allow us to study the effects of manipulating the activity of a single neuron, which is very difficult to achieve in more complex systems.

Previous work in the Zhang lab has identified a network of sensory neurons and molecules that regulate an integrated response to multisensory stimuli. To map out this network, we used a behavioral assay that challenges worms to choose between two conflicting stimuli: a food lawn of *E. coli*, an attractive stimulus, and a droplet of the chemical 2-nonanone, a repellent stimulus. Using this assay, we generated a time course quantifying the percentage of worms that leave the attractive food lawn in the presence of the aversive chemical over 60 minutes. After testing worms with ablated sensory neurons and mutated genes encoding neuropeptides and other molecules, we discovered that certain mutant strains of worms were "fast" and "slow" lawn-leavers, when paired with the aversive chemical, in comparison to wild-type worms. This has allowed us to map out a circuit of neurons, neuropeptides, and molecules that communicate via inhibitory and stimulatory signals in order to modulate the processing of multiple stimuli.

We now aim to explore whether the same neurons, molecules, and genes involved in the adult pathway are implicated in younger stages, in order to better understand how age affects the wiring, function, or plasticity of neuronal circuits involved in multisensory processing. Although we know that sensory processing is different in younger versus older mammalian systems, the mechanistic details remain largely unknown. Therefore, we are currently testing younger larval stage *C. elegans* in the integrated behavioral response to elucidate these mechanistic differences. Preliminary data suggest that younger wild-type larval worms have a reduced capacity to leave the food lawn when paired with an aversive stimulus, despite moving away from the aversive cue in a similar manner to adults. After testing various mutants in younger stages, it appears that the neuronal circuit regulating this behavior in adult worms is already present in younger worms. This result suggests that a separate pathway in younger worms is suppressing the integrated behavior of which adults are fully capable. Further dissection of this pathway in larval *C. elegans* could shed light on the manner in which multisensory processing in mammals differs from one developmental stage to another. Such information is particularly important to understand since some neurodevelopmental diseases with deficits in sensory processing present at a young age.

Investigating the Ontogeny of Instinctive Behavior in Mice

Helene Lovett
Integrative Biology, 2019

Harvard University

Advisor: Catherine Dulac, PhD

Mentor: Dhananjay Bambah-Mukku, PhD

Mice rely heavily on olfactory cues to detect social chemosignals to execute sex-appropriate social behaviors. The vomeronasal organ (VNO) in the mouse nose detects pheromone cues secreted by other mice, a signal that activates neural circuits that drive instinctive behaviors such as parenting, mating, and aggression. The gene *trpc2* is required for the vomeronasal system to function, and its inactivation has been shown to cause deficits in social behavior and sex discrimination. Mice lacking functional TRPC2 protein appear unable to distinguish between male and female mice and display a loss of conspecific aggression. However, little is known as to how the neuronal circuits that drive these behaviors develop. Early in life, input from the VNO to the neural circuits involved in instinctive behavior may play an instructive role in shaping the behaviors. Alternatively, these circuits may be to a large extent hardwired, and develop independent of input from the VNO during development.

To answer this question, the Dulac lab has developed a transgenic mouse using a genetic construct adopted from bacteria that allows for the selective

activation of *trpc2* expression only when mice are fed a diet containing the derivative doxycycline (dox). The aim of my project is to optimize the reversible expression of the *trpc2* gene in order to understand the effect of *trpc2* expression at certain time points in development on adult social behaviors. I am genotyping the transgenic mice to identify those positive for all the necessary genes in the construct and then will attempt to selectively activate *trpc2* only at certain intervals during development by feeding mice dox at those times.

To determine if *trpc2* was selectively activated, the VNO is dissected to determine relative levels of *trpc2* expression. If *trpc2* expression can be rescued in mice after inactivation at an earlier time point in development, then I will be able to reversibly control the activation or inactivation of the VNO and its ability to detect pheromone cues. I will put the adult mice through a number of behavioral paradigms to test the effect of selective VNO activation. If the mice show no behavioral deficits associated with a loss of VNO function despite VNO inactivation during development, it will lend support to the idea that hardwired circuits are sufficient to execute instinctive social behaviors. If deficits are observed, this would indicate that sensory input from the environment early in life plays a critical role in programming these behaviors.

In addition to this reversible gene activation project, I am assisting with another project aimed at uncovering neural circuits deeper in the brain, which drive social behavior based on VNO input. I am scoring videos of various social behaviors and using a technique to compare the neurons activated during consecutive behavioral assays in a brain region hypothesized to be downstream of the VNO pathway. These experiments will lead to a clearer understanding of the neuronal populations involved in the behaviors, and how neural activity along specific pathways drives instinctive behaviors.

The Effect of Affective Priming on Reward Responsiveness in Individuals with High vs. Low Depressive Symptoms

Isobel Green
Cognitive Neuroscience and Evolutionary Psychology, 2017

McLean Hospital

Advisor: Diego Pizzagalli, PhD

Mentor: Poornima Kumar, PhD

Individuals with major depressive disorder (MDD) demonstrate a blunted response to rewards in the environment. Behavioral studies have quantified this blunting by showing that depressed individuals do not preferentially make choices that have previously elicited reward. Throughout the 20th century, behavioral researchers conducted experiments demonstrating that this pattern reliably set individuals with MDD apart from their healthy counterparts. In the past few decades, neuroimaging has emerged as a potent tool for linking patterns of neural activation to specific behaviors, including reward processing in both healthy and depressed people. One consistent finding is that individuals with depression show decreased reactivity in a region called the ventral striatum during reward-based decision-making. The ventral striatum is the part of the limbic system and is key to linking motivational drives to behavioral actions. Aberrations in striatal activity have been linked to key symptoms of MDD: decreased sensitivity to positive affect (anhedonia) and increased negative affect (low mood).

There exist two possible models of reward processing in individuals with depressive symptoms. The "classical" model posits that individuals with depression have a baseline impairment in reward processing that causes them to experience a chronic blunting of reward responsiveness, regardless of their stress level. However, studies have shown that the aberrant profile of reward processing seen in individuals with depressive symptoms can also be induced transiently in healthy individuals by putting them under acute stress. This context-dependency of reward responsiveness in healthy individuals suggests another possibility: individuals with depressive symptoms exist in a state of heightened stress compared to healthy controls, and that their impaired reward processing is context-dependent, reflective of a stressful external state rather than a baseline impairment in reward processing. Distinguishing between these two possibilities holds clinical importance as it points to different treatment routes for the disease.

My project this summer takes a next step towards identifying which of these two frameworks best rep-

resents reward processing in individuals with depressive symptoms. Specifically, the work will investigate reward processing in individuals with high vs. low depressive symptoms at three different time points: before a positive mood induction, after a positive mood induction, and after a delay following a positive mood induction. Therefore, the study will allow for investigations of first, baseline differences in reward processing between individuals with high and low depressive symptoms; second, the acute effect of a positive contextual prime on reward processing; and third, the extent to which an emergent increase in reward-related activation persists over time.

Based on findings of context-dependence in neural activation during reward processing tasks, my team hypothesizes that although high-depressive individuals will show significantly lower reward responsiveness before a positive mood induction, they will normalize to a level near healthy controls after acute positive mood induction. We further hypothesize that this increase in reward activation will not last; following a time delay, reward responsiveness in individuals with high levels of depressive symptoms will again fall significantly below that seen in healthy controls, representing a trait-like inability to sustain positive affect.

Fluorescence and Electron Microscope Image Registration using EM Dense Labels in Zebrafish

Jaden Freeze
Chemistry and
Physics, 2019

Harvard University

Advisor: Florian
Engert, PhD

Mentor: Mariela
Petkova

The zebrafish (*Danio rerio*) is an emerging model organism for understanding how neural circuits underlie animal behavior. Recent technological advances in light microscopy allow us to record the activity of large neuronal populations with single-cell resolution in order to functionally identify potential neural circuits involved in the behavior of zebrafish larvae. The spatial resolution of light microscopy, however, is insufficient to visualize individual connections between neurons. Therefore, to verify hypothetical circuits, we use electron microscopy (EM) as a complementary imaging technique. EM provides structural images of the fish at subcellular-level resolution, which allows us to reconstruct neuronal processes and identify connections.

The challenge to this approach is to reliably match neurons from light microscopy images to their electron microscopy counterparts. To address this challenge, we test a dual-labeling method in which we mark neurons of interest in both light and electron microscopy, making them immediately distinct from their neighbors. In light microscopy, such labeling is achieved by expressing fluorescent protein tags, while peroxidase-derived tags can be used to increase the contrast during EM visualization. The latter requires a staining process in which the peroxidase catalyzes a reaction between peroxide and 3,3'-diaminobenzidine (DAB), such that the resulting precipitate attracts osmium. This makes the region containing the peroxidase appear dark in EM images. Finally, to co-express the two labels in the same neurons, we genetically insert a construct with genes for both the fluorescent and peroxidase tags under the same promoter in the zebrafish genome.

For this project, we develop a protocol in zebrafish for visualizing a recently available peroxidase tag, APEX ("enhanced" ascorbate peroxidase). To test staining protocols, we established a transgenic line with APEX and a red fluorescent protein (TagRFP) under the *isl3* promoter, such that both are expressed in a subset of the trigeminal neurons. These are mechanosensory neurons, which are located in the trigeminal ganglion behind the fish eyes and branch out to innervate the entire fish head. The stereotyped locations and long processes of these neurons facilitate the testing of staining protocols and help us evaluate how reliably these can be reconstructed and matched to light microscopy images. So far, we have seen visual confirmation of DAB stains with light microscopy, and are expecting successful EM labelling under the electron microscope.

Dopamine Neurons and the Execution of Sequential Behavior

Jennifer Bi
Neurobiology, 2018

Center for Brain
Science, Harvard
University

Advisor: Naoshige
Uchida, PhD

Mentor: Mitsuko
Watabe Uchida

Parkinson's disease is the most common movement disorder, and the second most common neurodegenerative disorder. People affected by Parkinson's have difficulties executing sequential behaviors, many of which are habitual tasks (e.g. driving, brushing teeth, etc.). Though symptoms of the disorder are associated with low levels of dopamine neurons (DA) in the brain, the connection between the function of DA and symptoms of Parkinson's disease is currently not well understood. The goal of our research is to gain a more detailed understanding of the impor-

tance of DA in the execution of sequential behavior.

Broadly speaking, our experiment is an odor-reward behavioral assay in which mice are trained to perform a sequential task by reacting to specific odors. Each odor tells the mouse which one of two water ports to poke. The mouse is rewarded with a small drop of water (30 ms) upon choosing the correct water port to the first odor, and a large drop of water (200 ms) upon choosing the correct water port to the second odor. After the second odor, the next trial starts with a random odor after a short interval of time, and the mouse repeats the task.

Once the mice are able to stably perform the task, we then inactivate specific subsets of neurons within the brain and examine changes in behavior. Our experiment examines inactivation in three sets of neuronal populations: 1) dopamine neurons within the lateral substantia nigra pars compacta (SNC), 2) neurons within the posterior striatum, and 3) neurons within the anterior dorsal striatum. Temporal inactivation of these neuronal populations is performed by a DREADD method involving the kappa opioid receptor, KORD, which is introduced with a virus, AAV. KORD hyperpolarizes and inactivates neurons when activated by its ligand, salvinorin-B (salB). Thus, by injecting AAV-KORD into target neuronal populations, we can perform temporal inactivation of these specific neurons by then injecting salB. Through inactivation of these neuronal subsets, we hope to better understand the importance of DA in the execution of sequential behavior, and use our results to further understand Parkinson's disease.

On Object Recognition in Rats

Josh Breedon
Medicine, 2020
Emmanuel College

Harvard University

Advisor: David Cox, PhD

Mentor: Javier Masis

Your ability to identify the object that you currently hold in your hands as your cherished abstract book is one that you most likely take for granted. As conditions in illumination, orientation, and degree of occlusion vary, a particular object can project a potentially infinite number of patterns of light upon your retina. And yet, seemingly effortlessly, your visual system is able to untangle this image to allow object recognition. While our knowledge of the peripheral, parallel processing that an image undergoes is well developed, our understanding of the details of visual cortex processing is still nascent.

The use of rodents as a model system in the study of vision is becoming increasingly popular, as more and more studies demonstrate their ability to discriminate between objects. Their neuroanatomy also has many parallels with human cortex; in humans the ventral "what" stream is responsible for determining "what" objects are present in an image, and the rodent's LM, LI and LL regions are thought to play a similar role. Therefore investigation of the roles of these higher-level processing regions in object recognition may help further our understanding of the function of human visual areas.

I have trained rats to discriminate between two three-dimensional objects, using water-pump levers to provide both the method of object indication and to deliver water as reward. The rats are trained over the course of several weeks, with the difficulty of the task increasing in phases. The difficulty ranges from distinguishing the objects face on, with horizontal separation of the object also prompting the correct lever, to recognising the objects transformed by a combination of rotation and size variation. The rats progress to the next phase once they consistently perform significantly above chance, with a correct response rate above 70%. The levers are wired such that the computer automatically logs a range of data on the rats' responses, including the number of successes, failures or ignores. I used Python to extract and interpret the data from the experiments.

The ultimate aim of the project is to selectively lesion the higher visual areas of the rats, under the guidance of functional imaging techniques. These lesions will damage the neurons solely in the selected region, causing the cells to die, while leaving other regions intact. Comparison of the rats' task performance before and after lesioning will allow us to create hypotheses about the role of these brain regions in object recognition in rats. We believe that a detailed understanding of the cortical processing in this model system may also provide insight into the intricacies of the human visual system.

Relations between Sleep Spindles and White Matter Pathways in Schizophrenia

Julia Versel
Neurobiology, 2017

Massachusetts General Hospital

Advisor: Dara Manoach, PhD

Mentor: Bengi Baran, PhD

One of the core features of schizophrenia (SZ) is cognitive deficits including learning and memory. Converging evidence suggests that sleep is critical for memory consolidation. Sleep spindles, which can be recorded and detected using electroencephalography (EEG) methods, are short bursts of 12–15 Hz oscillations during stage 2 of non-REM sleep. These oscillations, generated in the thalamic reticular nucleus of the brain, are known to play a role in sleep-dependent learning and memory consolidation. Individuals with SZ have a specific reduction in sleep spindles which predicts sleep-dependent memory consolidation deficits. Therefore, understanding the neural underpinnings of sleep spindle deficits may help in finding new targets for treatment of cognitive deficits in SZ.

The primary focus of my research project is investigating the structural integrity of brain regions (white matter pathways) critical for sleep spindle generation and propagation in SZ. I will quantify structural integrity using diffusion weighted imaging. Specifically, my research question is whether there is a difference in white matter integrity between individuals with SZ and matched controls that may be associated with sleep spindle deficits and subsequently impaired memory consolidation.

For that purpose, I am using diffusion weighted imaging (DWI), a neuroimaging technique that quantifies integrity of white matter tracts in the brain by measuring the molecular diffusion of water in brain tissue. I hypothesize that patients will show abnormalities in thalamocortical tracts which will correlate with spindle activity. By relating sleep spindle activity from EEG data with white matter pathway integrity from DWI, I hope to clarify the structural mechanism involved in sleep spindle deficits in patients. This investigation will hopefully illuminate the mechanism of sleep dependent memory consolidation deficits in SZ.

The Neuronal Basis for Motor Skill Learning and Execution in Rats

Kushi Mallikarjun
Molecular and Cellular Biology, 2019

Harvard University

Advisor: Bence Ölveczky, PhD

Mentor: Steffen Wolff, PhD

Every time we tie our shoes or practice our tennis serve, we make use of precise motor skills. We are not born with these abilities; rather, we learn them gradually until we are able to execute them consistently. How our brain learns and executes these motor skills with nearly perfect consistency is an intriguing question in neurobiology. Though studying this question in humans can be rather difficult, rats serve as a viable model organism to identify the mechanisms the mammalian brain uses to learn and execute complex motor skills.

The lab has developed a paradigm in which rats learn a precise motor skill that mirrors how humans learn complex motor skills. We make use of a high-throughput system in which the rats have to press a lever twice with a distinct time interval and develop a complex sequence of movements to fill this time gap. Thus, we have a model system for a motor task in hand that allows us to study the neuronal basis of motor skills. With this system, we try to find out which parts of the brain are involved in the circuit, how they communicate, and what roles they play.

Previous experiments in the lab have shown that a lesion of the motor cortex, an important player in motor control, before the learning of this precise motor skill, prevents the rat from learning. However, if the motor cortex is lesioned after learning has already taken place, the rat is still able to execute this motor skill with precision and consistency, indicating that the motor cortex is important for the learning but not the execution of this motor skill. Thus, another region of the brain must be responsible for storing and executing this motor skill once it has been learned, and the motor cortex might act as a tutor for this region during learning. One potential tutoring target is the striatum, a subcortical area of the basal ganglia involved in motor control. Further experiments in the lab have shown the striatum to be necessary for both the learning and the execution of the motor skill.

We want to find out if the striatum is really tutored by the cortex. To address this question, we use a viral silencing technique that allows us to prevent neurons in the cortex from communicating with the striatum.

A crucial aspect of these experiments is to determine the precise injection site of the viruses used for silencing, the number and the exact location of the infected and silenced neurons. We can correlate the extent of the silencing with the impairment of learning and conclude whether already small perturbations of the connection between cortex and striatum are sufficient to impair the learning. This approach will provide important information about the communication between different parts of the motor system and how they interact to allow for the learning and execution of motor skills. This knowledge will help us to understand how this essential part of our lives is actually implemented in our brains.

Classifying Proprioceptive Neurons of the Peripheral Nervous System

Lauren Sweetland
Neurobiology, 2018

Howard Hughes Medical Institute, Harvard Medical School

Advisor: David Ginty, PhD

Mentors: Brendan Lehnert, PhD; Yang Zheng

Proprioception is the sensory ability to perceive body position and movement. Made possible through proprioceptive neurons and their accompanying terminal endings, such as the well-known Golgi tendon organ and muscle spindle fibers, proprioception is integral to many locomotive organisms. Dysfunctional proprioception is associated with a wide range of neurological and muscular impairments such as Ehlers-Danlos syndrome. Many aspects of proprioceptive neurons remain unknown, however. As a part of the peripheral sensory system, proprioceptive neurons have remained largely indistinguishable from other sensory neurons, preventing definitive research on properties such as anatomy and central nervous system connections unique to the proprioceptive classes.

Through ongoing research at the Ginty Lab, we hope to find a strategy to label subtypes of proprioceptive neurons. Recently, we have been using knock-in mouse models that utilize fluorescent proteins, such as GFP, to visualize specific neuronal classes. Furthermore, we use these mice to perform intersectional *in situ* and immunofluorescence strategies, allowing us to qualitatively assess the expression of multiple genes within one tissue sample. These approaches will help to identify gene expression levels unique to proprioceptive neurons, effectively cracking the biological code necessary to distinguish proprioceptive neurons from the billions of other peripheral neurons.

This ability to single out proprioceptive neurons will open up a world of research possibilities. Novel terminal endings could be investigated; further classification of the broad "proprioceptive neuron" could be detailed; *in vivo* and behavioral experiments could finally be accomplished—the enigmatic depths of the nervous system will be explored like never before.

The Function of the Cysteinyl Leukotriene Receptor 2 in Itch Pathology

Lin Ni
Neurobiology, 2019

Harvard Medical School

Advisor: Isaac Chiu, PhD

Mentor: Tiphaine Voisin, PhD

Itch defines the irritation prompting scratching that accompanies many skin conditions such as atopic dermatitis. Originally classified as a minor subtype of pain, itch is now perceived as a sensation different from pain, a separate entity in an organism's physiological pathway. Itch, although a highly pertinent sensation, was previously an understudied phenomenon. In the last decade, however, the study of itch has gained much more interest. Focused research in this area has led to the discovery of itch-specific receptors and distinct molecules separating itch from other senses, and recent genetic profiling of somatosensory neuron populations has identified specific neuron populations showing high transcript expression levels for itch receptors. There are, however, still many facets of itch sensation that remain targets for future research.

We focus our research on one of the target receptors, Cysteinyl leukotriene receptor 2 (Cystlr2), which has been identified in a neuron population implicated in itch. Though the Cystlr2 receptor has an established role in the immune system, particularly during inflammation, the role of the receptor in itch is still unclear. It is suspected that the Cystlr2 receptor could potentially mediate the interaction between Cystlr2 receptor ligands produced by immune cells and the consequent itch response.

Experiments were performed to evaluate if a particular subset of neurons could be activated by one of the endogenous Cystlr2 ligands, LTC4. Activation of a specific population of neurons in wild-type mice, but not Cystlr2 knockout mice, would identify the Cystlr2 receptor as a potential target in alleviating atopic dermatitis and other chronic skin conditions characterized by excessive itch.

We approach identifying the function of the Cystlr2 receptor in itch through both *in vitro* and *in vivo*

experiments. *In vitro*, dorsal root ganglion (DRG) neurons are cultured from a wild-type (WT) mouse and a knockout (KO) mouse for the Cystlr2 gene. Calcium imaging is then used to determine the response of neurons from the mice to capsaicin (as a control, to eliminate dead neurons) and to LTC4. *In vivo*, differing concentrations of LTC4 are injected into both WT and KO mice. Mice are then evaluated, in five-minute increments for thirty minutes, for the number of observed scratching bouts.

Preliminary results have shown some activation of neurons in response to LTC4 in DRG neurons *in vitro*. *In vivo*, preliminary results indicate that KO mice tend to have a slightly diminished itch response to LTC4 injections as compared to WT mice, but the results have yet to be consolidated into a definitive finding.

If follow-up studies show that there is indeed a difference in responsiveness between the neurons of knockout mice and the neurons of wild-type mice to LTC4, we can establish an interaction between immune response and itch response as mediated by the Cystlr2 receptor, making the receptor a future target for treatment research on chronic skin disorders and atopic dermatitis.

Investigating the Role of KFM1 and Microglia in the Segmental Specification of Corticospinal Motor Neurons (CSMN)

Michael Liu
Human
Developmental and
Regenerative Biology,
2019

Harvard Stem Cell
Institute

Advisor: Jeffrey
Macklis, MD

Mentors: John
Hatch; Yasuhiro
Itoh, PhD

In order to better understand the targeting of corticospinal motor neurons (CSMN), we are investigating the role of KFM1 and microglial pruning in the developing spinal cord. The central nervous system (CNS) features a tremendous diversity of neurons along a spectrum of function, morphology and molecular expression. One particular subtype—CSMN—originates from pyramidal cells in layer V of the neocortex that descend through the corticospinal tract (CST). CSMN are the primary subpopulation of neurons that degenerate in ALS, hereditary spastic paraplegia (HSP), primarily lateral sclerosis (PLS) and spinal cord injury. The axons that project through the CST arise from different anatomical regions of the motor cortex. Furthermore, CSMN are a heterogeneous population of neurons. $CSMN_L$ that regulate leg movement send axons from the medial motor cortex to the lumbar spinal cord while $CSMN_C$ that regulate hand/arm movement send axons from the lateral motor cortex to the cervical spinal cord. During development, the proper pruning and specification of these long axon collaterals are necessary for the establishment of mature connections and a functional central nervous system (CNS). Aberrant pruning of axons has been implicated in several neurological disorders.

KFM1 —a small leucine-rich repeat keratan sulfate proteoglycan—is a prominent protein in the extracellular matrix (ECM). Microarray data from our lab have shown that KFM1 is expressed selectively by $CSMN_C$ and not $CSMN_L$. Biotinylated dextran amine (BDA) and fluorescence imaging have both revealed significant increase in axonal branching of the $CSMN_L$ population within the cervical cords of KFM1 knock out (KO) mice. This is an intriguing observation, because KFM1 is expressed only in the $CSMN_C$ population but the phenotype was observed in the $CSMN_L$ population. However, it remains unclear whether $CSMN_C$-derived KFM1 affects CSMN branching through direct or indirect mechanisms. Previous studies have shown that KFM1 is capable of binding and promoting the activation of macrophages. Therefore, we asked if the mechanism of specification by KFM1 was mediated by microglia—the resident macrophages of the CNS. Our initial immunohistochemistry staining demonstrated a decrease in the ramification and expression of activation markers of microglia within the cervical cords of KFM1 KO mice at P7 and P10, corresponding to the fact that the axons that form the CST are pruned by the second week of postnatal development. To further elucidate the role of microglia and KFM1 in CSMN targeting, we will perform AAV rescue experiments and interrogate the impact of microglial ablation on axon collateral branching. This project will also be accompanied by an optimization of recently developed clearing techniques to obtain high-resolution three-dimensional images of axon branching in the developing spinal cord.

An understanding of the molecular mechanisms that mediate axon branching is crucial to decoding the intricate process of CNS development and to designing treatments for aberrant circuit formation that underlie numerous pathologies. Furthermore, the exploration of axon branching and pruning will open an entirely new realm of therapies that may be leveraged for regeneration and repair of the CNS after traumatic spinal cord injury and dysfunction.

Patterns in Odor Object Recognition

Miruna Gabriela Cristuş
Applied Mathematics, 2019

Center for Brain Science, Harvard University

Advisor: Venkatesh Murthy, MD, PhD

Mentor: Elizabeth Shtrahman, PhD

Perception is the basis of our understanding of the world. Thus, we ask: how is sensory information encoded in the brain and how does it relate to behavior? In our experiments, we are trying to answer this question in mice, by seeing how olfactory information is represented in the brain and how it impacts their ability to identify odors. Because natural odors are mixtures of different chemical species, we want to analyze whether odors are perceived as the sum of their individual components (elemental processing) or as something with different holistic qualities relative to the mere components (configural or synthetic processing). There is no current consensus in the literature regarding these two outcomes; therefore, we aim to thoroughly characterize the patterns behind these results. Specifically, we want to quantify what odor inputs and properties of certain odor mixtures correspond to certain outcomes.

In order to analyze these features, we have designed a behavioral task to test odor object recognition. The odor objects in our experiment are represented by odor mixtures of three different components. The mice are thus trained to recognize such a mixture and differentiate it from other mixtures of three components that they have never encountered before. We are testing a variety of target mixtures, chosen for the chemical likeness of the components or similarity of their glomeruli responses in the brain (initial sites of neural responses). In order to measure their performance, we train water-restricted mice to respond to stimuli by licking "right" on a lickport for the target odors and "left" for the nontarget odors. This experimental setup also allows us to see the type of mistakes mice make and how their mistakes evolve during the training sessions.

Preliminary results show that mice are able to learn this task well and we are further testing a variety of odor object stimuli. We expect our results and future work to enable us to find the link between these stimuli and their perception as configural versus elemental odors.

Understanding Formation of Mirror Neurons in Humans

Mohamed Ebied
Biomedical Engineering, 2019

Harvard Medical School

Advisor: John Assad, PhD

Mentor: Antonino Casile, PhD

Mirror neurons are a class of neurons which have been identified in the ventral premotor cortex, the primary somatosensory cortex, the supplementary motor area, the primary somatosensory cortex, and the inferior parietal lobule. Since their discovery in the early 1990s, mirror neurons have become a critical consideration in our broader understanding of neurological systems, sparking significant debate in the scientific community due to their highly unique response properties. Mirror neurons are unique in their ability to respond to the execution as well as observation of goal-directed motor actions. Initially observed in monkeys, the later discovery of mirror neurons in humans is considered to be one of the most important discoveries in systems neuroscience.

Naturally, much research on this subset of neurons has followed, and the literature is dense with attempts to understand their role from an evolutionary as well as functional perspective. However, the available literature on precisely how they form is relatively sparse. My research attempts to delineate how mirror neurons form in response to new motor stimuli. In order to accomplish this, we have built a virtual reality platform which allows us to carefully train human subjects on highly specific motor tasks, minimizing the input from external stimuli. The use of a virtual reality platform also ensures that the visual associations with the motor task remain highly specific, and are not influenced by surrounding visual distractions. Based on previous experiments, it has been shown that four days is the approximate amount of time required for the brain to physiologically respond to new stimuli. Hence, the "training" period for subjects is set to four days. Subjects are scanned using fMRI before, during, and after the training phase of the experiment in order to assess the changes which are occurring in response to the new motor stimulus.

It is our hope that garnering a greater understanding of how the brain actively rewires itself to generate mirror neuron systems will help forward our understanding of their function as a whole. Since fMRI data would allow us to quantitatively understand the cascade of events which lead to the formation of these neurons, we can begin applying our understanding

of this process to the functions that mirror neurons have been hypothesized to allow—namely empathy, intention, understanding, and even complex linguistic development. We do not yet have enough data to determine what extrapolative conclusions can be drawn, but we have taken the first step in developing an accurate model of mirror neuron formation.

Disruption of Memory-Consolidating Slow Oscillations and Calcium Transients by Amyloid in Alzheimer's Disease

Shenyece Ferguson
Neurobiology, 2019

Massachusetts General Hospital

Advisor: Brian Bacskai, PhD

Mentors: Ksenia Kastanenka, PhD; Guillaume Pagnier, MS

Alzheimer's disease (AD) is a progressive neurodegenerative disorder characterized by the presence of amyloid plaques, neurofibrillary tangles, and loss of neurons in the human brain and is a major cause of dementia, or memory loss. Slow oscillations are responsible for consolidation of memories during sleep.

To determine whether aberrant slow oscillation activity is part of the disease progression, we used a mouse model overexpressing a mutant amyloid precursor protein (APP) and a mutant presenilin (PS1). Amyloid, a cleavage product of APP, is responsible for disrupting slow waves in the cortex by possibly altering electrical activity in inhibitory interneurons. To test whether interneuron activity is altered, mouse brains with interneurons that could be targeted in the context of AD were imaged with multiphoton microscopy. A fluorescent calcium indicator, GCamp, was targeted to interneurons of APP mice and wild-type littermates. Calcium imaging was performed using multiphoton microscopy to image calcium transients in interneurons.

I performed analysis by selecting regions of interest (ROIs) which corresponded to cell bodies of the interneurons using ImageJ software. The change in fluorescence over the average fluorescence was calculated for each selected cell over a 500 frame (approximately 215 second) timelapse. In MATLAB, a Fourier transform was performed to identify the amplitude and frequency of the slow oscillations of each cell. In wild-type mice, we expected to observe calcium transients that were similar to slow oscillations of interneurons and at the same frequency as slow waves. However, in APP mice, we predicted the calcium transients to be disrupted by the presence of APP, which could be responsible for the aberrant slow oscillations. The calcium transients of the wild-type mice were observed to be as expected, rhythmic and oscillatory, while in APP mice, they were sporadic and of lower power than those of the wild-type mice. Data analysis is ongoing to confirm these promising results and better our understanding of AD, slow waves, and memory formation.

Use of a Novel HEK 293 Culture Model to Examine the Role of Neuropilin-2 in Drd4 Retinal Ganglion Cells

Vimal Konduri
Molecular and Cellular Biology, 2017

Harvard University

*Advisor:* Joshua R. Sanes, PhD

Mentor: Yirong Peng, PhD

I am investigating the role of the receptor protein neuropilin-2 (Nrp2) in retinal ganglion cells (RGCs), which transmit visual information from the retina to the brain. Neuropilin-2 has been implicated in axonal guidance, dendrite development, angiogenesis, and the formation of tumors such as melanoma and glioblastoma. Neuropilin-2 is known to colocalize with the plexin receptor PlexA3 in the presence of class-3 semaphorins, and with VEGFR-2 and VEGFR-3 in the presence of vascular endothelial growth factor. The role of neuropilin-2 in the retina, which is part of the central nervous system, is currently being studied in a collaborative effort between the Sanes lab and the lab of Alex L. Kolodkin at the Johns Hopkins University. Previous studies by our lab in mice have shown that Drd4-RGCs, a specific type of direction-sensing RGC (which sense movement in the nasal direction and normally make up roughly 5% of the total RGC population), are greatly reduced in number when Nrp2 is knocked out. However, a mechanism of cell death has not been established.

I will study neuropilin-2 by examining systems in which its ligands, the semaphorins Sema3C and Sema3F, and VEGF, are overexpressed. I will use cells from the HEK 293 cell line transfected with plasmids to constitutively produce these ligands in order to simulate an overexpression, and will plate them in a co-culture with retinal ganglion cells. After co-culture, I will assess cell survival and neurite (axon and dendrite) growth using live/dead and beta-III-tubulin staining, respectively. I will carry out this assay for both RGCs in general and Drd4-RGCs (which can be isolated using FACS sorting) specifically. I will be able to assess the proportion of RGCs that survive in culture and the average length of RGC neurites, and compare

each to a control in which RGCs are co-cultured with unmodified HEK 293 cells. The results of these studies can help shed light on how neurons in the central nervous system, such as RGCs, make key decisions related to survival and development.

Organismic and Evolutionary Biology
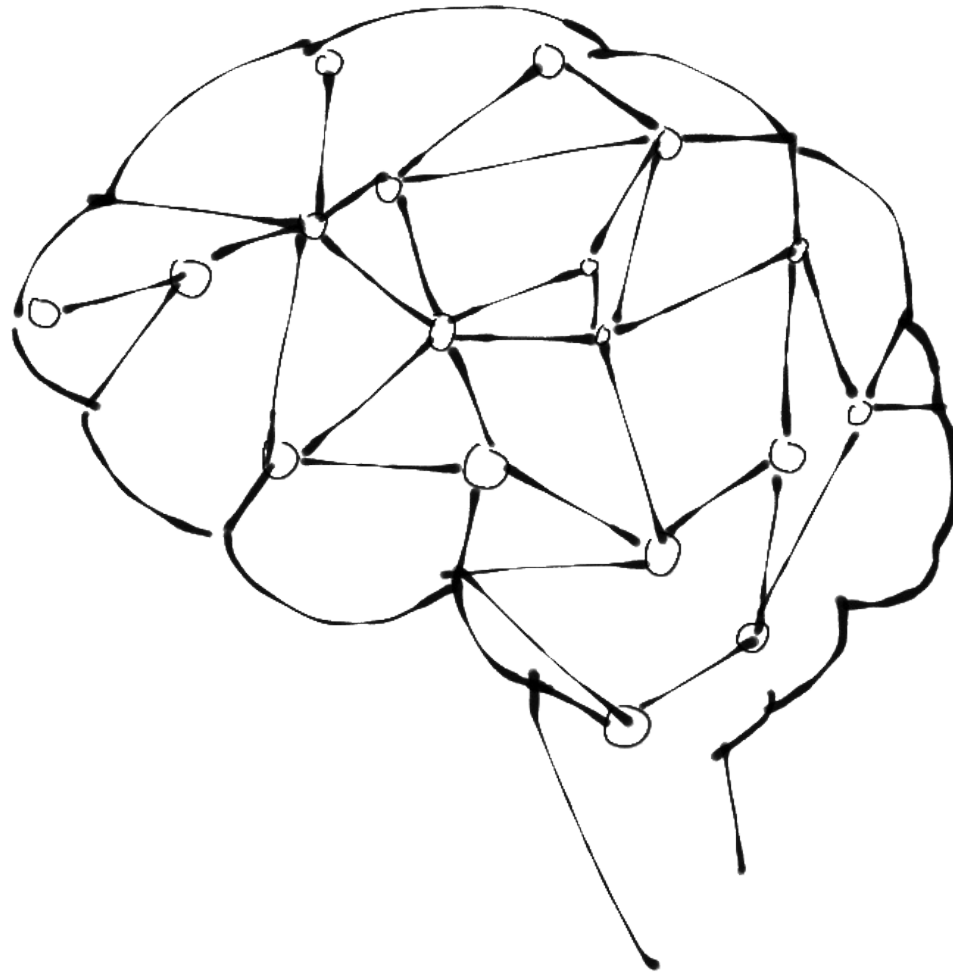
Development and Evaluation of Diagnostic Tools for Zika Virus

Cynthia Luo
Organismic and
Evolutionary
Biology, 2019

Harvard University,
Broad Institute of
MIT and Harvard

Advisor: Pardis
Sabeti, MD, PhD

Mentor: Mary Lynn
Baniecki, PhD

New cases of Zika virus are confirmed every day in the Americas, and the current outbreak has reached pandemic levels in parts of Central and South America. Zika virus, an RNA Flavivirus spread by *Aedes* mosquitoes, causes the infectious disease Zika fever. Between humans, the virus can be sexually transmitted and passed from mother to child, which may lead to brain malformations such as microcephaly. More severe infections of Zika virus in adults can also lead to Guillain-Barré syndrome, a life-threatening autoimmune disease that damages the peripheral nervous system. Low viral titers in Zika-positive patient samples and the similarity of initial symptoms to other infections such as Dengue fever and Chikungunya present multiple challenges to properly diagnosing Zika virus infection. Thus, extremely sensitive and accurate diagnostic tests for Zika virus are needed in order to provide patients with proper diagnosis and appropriate medical treatment.

Reverse-transcriptase quantitative polymerase chain reaction (RT-qPCR), a low-cost, robust technique that requires minimal training, is an ideal technique to deploy in field settings. RT-qPCR diagnostic tests for Zika have been previously developed but lack specificity to the current outbreak. Previous assays were developed using whole genome sequencing (WGS) data from an African strain of Zika, shown to be 11% genetically divergent from strains isolated in the Americas during the current outbreak. A lack of strain specificity in current diagnostics could lead to improper patient diagnosis, which delays appropriate treatment and worsens medical prognosis.

We are developing a RT-qPCR assay that is robust and highly specific to Zika virus strains across the globe. To design our assay, we first identified the most conserved region of the Zika genome by using all 80 available whole genome sequences from the National Center for Biotechnology Information (NCBI) and by creating multiple sequence alignments with Geneious software. Next, we used in-house software capable of optimizing the identification of highly conserved nucleotide regions in the Zika genome and designed new primers and probes to target these homologous regions. The specificity of the new assay design was evaluated using BLAST against the NCBI nucleotide database. We are currently evaluating the performance of the assay by assessing assay robustness, sensitivity (limit of detection), reproducibility, and cross reactivity to ensure that it will be highly specific and sensitive to detecting Zika virus in patient samples.

In parallel, we are using similar methodology to develop a multiplex RT-qPCR diagnostic test for Zika, Chikungunya, and Dengue fever. This will allow medical professionals to use one test to differentiate between all three clinically similar diseases, greatly reducing the costs and time associated with performing these diagnostic tests. We hope that these tests will be directly transferable in diagnosing Zika virus in the field and will have profound impacts on improving patient treatment for Zika infection.

Developing a SNP-Based Barcode for *Babesia microti*

Jade Moon
History and Science,
2017

Harvard University,
Broad Institute of
MIT and Harvard

Advisor: Pardis
Sabeti, MD, PhD

Mentor: Mary Lynn
Baniecki, PhD

Babesia microti is a malaria-like parasite that infects red blood cells and is spread by *Ixodes scapularis* ticks. Because the parasite is endemic to different regions across the world, the ability to identify the geographic origin of a strain can help us locate sources of outbreak, monitor the patterns and distributions of infections, and ultimately aid in elimination efforts.

To distinguish between strains, we look to the genetic variations that exist between populations; for example, a single nucleotide polymorphism (SNP) allele common in one geographical group may be rarer in another. Thus, genotyping SNPs provides a promising means of developing an inexpensive and rapid assay that can uniquely identify the geographic origin of a given strain. In *Babesia. microti* where SNPs have been determined through whole genome sequencing, we can select a small subset of SNPs that will identify parasitic infections and differentiate geographic ori-

gins. This set of SNPs will serve as unique genetic signatures or barcodes for *Babesia. microti.*

We identified informative SNPs by considering two criteria. First, the subset of SNPs must maintain clustering after performing principal component analysis (PCA). PCA ensures that our selection of SNPs captures high degrees of population diversity and differentiates between geographically distinct populations. Second, to avoid positive selection, we must only consider SNPs resulting in synonymous mutation. SNPs were genotyped using high resolution melting analysis (HRM). HRM takes advantage of the differences in melting temperatures of amplicons with different base pairings.

Developing a rapid, sensitive, and specific tool to identify parasitic origins will not only aid in correct diagnostic and treatment of infections, but also allow for examination of the phenotypic differences of different infections. Here, we developed a 40-SNP barcode consisting of 20 mitochondrial SNPs and 20 nuclear SNPs. Due to the inheritance patterns of mitochondrial and nuclear DNA, the combination of mitochondrial SNPs and nuclear SNPs allows efficient identification of geographic heritage and genomic information.

Embryonic Origins of Adult Stem Cells in the Acoel *Hofstenia miamia*

Silvia Golumbeanu
Integrative Biology, 2017

Harvard University

Advisor: Mansi Srivastava, PhD

The evolutionary basis and developmental mechanisms for regeneration in animals are not well understood. Planarians (e.g. *Schmidtea mediterranea*) and acoels (e.g. *Hofstenia miamia*) are distantly related animals both capable of whole-body regeneration; while they are evolutionarily distant, they both possess populations of pluripotent adult stem cells called neoblasts. However, the developmental origins of neoblasts are unknown. The acoel *Hofstenia miamia* is particularly useful in addressing the question of neoblast origin, as it produces abundant embryos that are amenable to experimental investigations.

We seek to understand the origins of the stem cell population by studying the embryonic development of *Hofstenia* using known adult neoblast markers—primarily *piwi-1*. Our data will allow us to distinguish between two alternative hypotheses for the developmental origin of pluripotent stem cells: either *piwi-1* is expressed from very early on in the em-

bryo's life and is subsequently restricted, or it begins expression later on specifically in the cell lineage that gives rise to neoblasts. We will also compare *piwi-1* expression to expression of other neoblast markers with possible related functions such as *vasa* and *tudor*. We are generating a staged series as a reference for embryonic development in *Hofstenia*. We are also developing methods to use *in situ* hybridization to determine the expression pattern of *piwi-1* mRNA in embryos at different stages.

As a new model organism, *Hofstenia miamia* offers a unique opportunity to study the basis of regeneration in bilaterians. Whole-body regeneration could be the ancestral state: most early lineages of metazoa are capable of performing it. Thus later diverging lineages of bilaterians seem to have restricted regenerative capacities. By documenting the previously unstudied embryonic development of Hofstenia miamia, and studying the expression of *piwi-1* in early stages of this development, we can explore the genetic mechanisms and cellular pathways necessary for maintenance and regulation of a successful adult pluripotent stem cell population. In an evolutionary context, we can use these observations to compare to gene expression during the embryonic development of later diverging lineages, and potentially explore what changes occurred that led to the restriction of regenerative capacity.
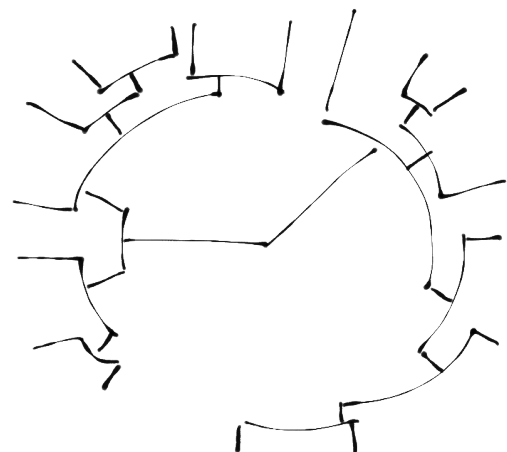
Physics and Biophysics

Slipping and Sound in the Motion of the Euler Disk

Adam Frim
Physics & Mathematics, 2018

Harvard John A. Paulson School of Engineering and Applied Sciences

Advisor: Ariel Amir, PhD

For over a century, physicists have studied the mechanics of a coin spinning on a tabletop or a larger version of the same problem known as the "Euler disk." Recent analysis has tended to focus on this system's unique feature of a finite-time singularity; that is, in a finite amount of time, the coin's rate of precession diverges. Put simply, the coin naturally "wobbles" at a rate that increases without bound, which is easily observed by listening to the coin as it reaches the final seconds of its motion as a whirring sound is produced. This unique feature has led to a number of publications with particular focus given to studying different energy dissipation mechanisms.

During the course of the coin's motion, the coin can roll with or without slipping. The case without slipping is far more easily approached; therefore, the goal of the current project was to build a mathematical model that incorporates both the rolling and slipping of the disk. Due to nonlinearity in the governing differential equations for both the rolling and slipping regimes, the problem was approached numerically. The resulting numerical model can independently transition from rolling to slipping when certain conditions are met. Such a model can provide useful simulation while remaining experimentally verifiable. Moreover, this numerical model should more accurately suggest the rate at which slippage occurs, a question that has not yet been sufficiently studied. Finally, the model could be used to suggest a mechanism by which sound is repeatedly made during the coin's motion. In particular, if this model predicts that the frequency of slippage aligns with experimentally observed sound spikes, this could be a strong indicator that the sound arises from a stick-slip phenomenon in the motion of the disk.

Constraining Dark Matter-Electron Scattering

David Gonzalez-Dysinger
Physics, 2018

Harvard University

Advisor: Cora Dvorkin, PhD

In the 1970s, astronomers noticed a discrepancy between the measured total kinetic energy of galaxies and the energy that theoretical models predicted given the total mass of all the stars and other visible matter in the galaxies. In order to account for this discrepancy, cosmologists proposed the existence of dark matter, an undiscovered form of matter that does not emit or interact with light except gravitationally, and is made up of particles outside the Standard Model of particle physics.

Dark matter particles could possibly scatter with baryons, such as protons and electrons. If elastic scattering between dark matter and baryons really occurs, it would affect the development of the universe in detectable, quantifiable ways. In particular, it would create observable changes in linear density fluctuations in the early universe. By analyzing data from the cosmic microwave background and the Lyman-alpha forest, we can place constraints on the strength of dark matter-baryon scattering. This will allow us to narrow down the range of masses that dark matter particles can have in scattering models.

I am studying the elastic scattering between dark matter and electrons in the early universe, from right after the Big Bang up until recombination, when electrons combined with protons to form hydrogen atoms. Using theoretical models, I compare the momentum exchange between dark matter and electrons to that between dark matter and protons to see if there are any regimes of dark matter mass and redshift in which the electron momentum transfer dominates over the proton momentum transfer. Studying the interactions between dark matter and low-mass particles require the modification of the equations used to study scattering with higher-mass particles. I am looking for a region in the parameter space where the dark matter-electron scattering dominates, predicting the effects of the scattering on observables, and using data to constrain this scattering.

The Statistical Physics of mRNA Dynamics

Elgin Korkmazhan Gulpinar
Chemistry and Physics, 2017

Harvard University

Advisor: Erel Levine, PhD

Mentors: Hamid Teimouri, PhD; Neil Peterman, PhD Candidate

The universe would be a very boring place if it were in equilibrium, or more precisely, in thermodynamic equilibrium. This is what is called the heat death of the universe, and it would simply mean that energy would be evenly distributed throughout all of space. With no imbalance to "drive" systems, processes such as information transfer, computation, and life would be impossible. Indeed, systems in equilibrium only account for a minuscule amount of what surrounds us. Yet, we still do not know much about non-equilibrium systems. In fact, non-equilibrium statistical physics, the field trying to answer questions about the behavior of such systems, has been declared as "one of the greatest challenges of 21st century physics." A relatively new field, it has been developed and utilized in the studies of topics ranging from black hole formation to quantum device technology.

What we are interested in is what non-equilibrium statistical physics can tell us about processes in living cells and how those processes can reveal important physics. Namely, we analyze the dynamics of mRNAs. Proteins are molecules functionally and structurally crucial for cells, and mRNAs are DNA-like molecules mainly responsible for providing the protein production machinery with the necessary code for protein synthesis. It has been shown that the amounts, positions, and gradients of particular mRNAs within and across cells are responsible for both smaller-scale events such as the localization of protein production and larger-scale changes such as the differentiation of stem cells into cells of tissues such as those of the brain.

Recently, there have been experimental observations of non-uniform gradients of specific mRNA involved in antibiotic resistance pathways in bacterial cells. The surging medical and societal cost of the failure of antibacterial drugs against resistant bacteria, or "superbugs," makes such systems very important to study. We seek the understanding of the reasons behind such mRNA gradients, their role in bacterial cells, the effective ways in which they can be perturbed, the information transfer needed for their emergence, the means of this information transfer,

the thermodynamic variables (temperature, concentration, etc.) they are related the most to, and the interplay between key players such as repressors, activators, and catalyzers. Our mathematical machinery involves simulations to reproduce what happens in real cells and theoretical analysis in order to understand them, to see how we can obtain a "big picture" of the network and perhaps interfere with it.

Currently, we are working on explaining the localization of an mRNA onto the cell membrane of *E. coli*, one of the most diverse bacterial species. Though mRNAs themselves cannot attach to the cell membrane, the domains of the proteins they help synthesize can. During protein production, the synthesis machinery moves along the mRNA and holds the mRNA and the nascent protein together. Thus, the existence of membrane binding domains on the already made part of the protein makes it possible for the mRNA to anchor to the membrane, until protein production is done and the machinery lets go of the mRNA. This mechanism suggests that membrane localization depends on factors such as the collision rate with the membrane, the "stickiness" of the protein's membrane binding domains, the size of the remaining part of the protein to be produced, and the availability of more free "anchors" on nascent proteins being simultaneously made by other synthesis machinery progressing along the same mRNA. Our aim is to understand these factors as comprehensively as possible.

Magnet Design for a Nuclear Fusion Reactor

Jacob Meyerson
Physics, 2019

MIT Plasma Science and Fusion Center

Advisor: Joseph Minervini, PhD

Mentors: Franco Mangiarotti, PhD; Taylor Sorenson; Aaron Hamilton

One of the biggest problems humanity faces is producing electricity without fossil fuels. The ability to do so on a large scale would have widespread benefits, from helping combat climate change to helping increase the independence of the United States. Nuclear fusion is an appealing option that could help solve this problem. Unlike solar or wind, it would be a constant source of electricity, and unlike nuclear fission, it would not produce long-term radioactive waste nor pose the risk of a nuclear meltdown.

Fusion energy is produced when two hydrogen isotopes' nuclei fuse to form a helium nucleus, a reaction which releases energy. For these reactions to occur, hydrogen isotopes must be heated up to very high

temperatures—hundreds of millions of degrees—to form hot plasma. Since the plasma is so hot, you cannot contain it in a box—it would melt through any material. However, one possible way of confining the plasma is through the use of high magnetic fields. The magnetic field acts as an invisible force that suspends the plasma in vacuum.

Producing these high magnetic fields is a challenge. A method of doing so is through the use of superconducting materials. At extremely cold temperatures—at around 5–20 K, near absolute zero—superconducting materials have zero resistivity. Therefore, the materials can handle large amounts of current, consequently creating a large magnetic field. Creating these magnets is challenging. They must produce the appropriate magnetic field, they must be able to withstand the stress due to the force from the magnetic field, and they must remain cold while million degree plasma floats a meter away.

Researchers at MIT's Plasma Science and Fusion Center are working on designing these magnets. Yet, their in-depth simulations of various aspects of the magnet (for example, the magnet's stress concentrations) take weeks to complete. Along with two other students, I am working on a code that will calculate first-order approximations of the overall magnet design—the stresses of the magnet, the magnet's geometry, the cooling of the magnet, the structural composition of the magnet. By using the code, the researchers will hopefully be able to narrow down the areas in which they investigate.

Towards Electrically Controlled High Temperature Interlayer Exciton Superfluids

Kateryna Pistunova
Physics, 2018

Harvard University

Advisor: Philip Kim, PhD

Mentors: Luis Jauregui, PhD; Andrew Joe

The concept of reduced dimensionality in physical systems has long been studied as a purely abstract although mathematically interesting idea. Such systems were predicted to exhibit unique electronic and optical properties, making them completely distinct from their 3D counterparts. However, the success of graphene, a single-atom-thick layer of carbon, has proven it possible to realize 2D systems experimentally. Held together by weak electrostatic interaction, often called van der Waals bonding, graphene layers can easily be cleaved from bulk graphite. The rise of graphene has led to the emergence of a whole family of crystalline 2D materials, among them an insulator boron nitride (BN) and transition metal dichalcogenides (TMDCs). Although TMDCs have been known for decades as inorganic semiconductors, the ability to isolate their single layers opened up new opportunities in nanoelectronics, optoelectronics, and photonics.

In conventional semiconductors, electrical transport is realized with either negatively charged electrons or positively charged holes (an absence of an electron in a place where it could exist in the crystal structure). Due to electrostatic interactions, these oppositely charged carriers can form excitons, bound states with particle-like behavior. A characteristic feature of excitons is their lifetime, the time it takes for electron and hole to recombine. Normally, the close proximity of electrons and holes in conventional semiconductors results in very short lifetimes, such as a few picoseconds in GaAs. However, spatial separation of an electron and a hole, which form a so-called "space indirect" exciton, leads to much longer lifetimes due to the increased distance between oppositely charged carriers. Our project aims to form spatially indirect excitons localized in vertically stacked monolayers of n- and p-type TMDCs (that is, with high electron or hole concentration, respectively). Due to the

reduced dimensionality of TMDC monolayers, electrons and holes are confined to a plane, resulting in an enhanced electrostatic interaction between them. Consequently, the two-dimensional nature of monolayer TMDCs facilitates the formation of spatially indirect interlayer excitons with prolonged lifetimes predicted to be longer than microseconds. The extended lifetimes provide a platform for exploring exciton Bose-Einstein condensation (BEC) phenomena, a state of matter in which a large portion of particles occupy the same energy state and behave as a single entity, exhibiting macroscopic quantum phenomena. Such exciton BEC is expected to form at record-high temperatures and has never been realized in TMDCs before.
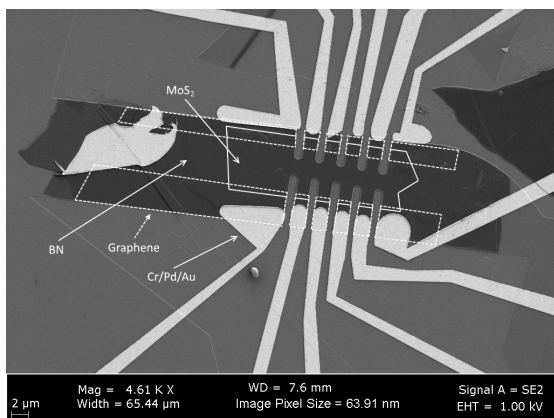
In order to experimentally study BEC phenomena in TMDCs, high-quality electrical contacts need to be demonstrated to control the chemical potential and electric field between TMDC layers. However, due to TMDCs' semiconducting nature, a potential energy barrier is formed at the metal-semiconductor junction, increasing the contact resistance, prohibiting current flow, and lowering the device performance. Our goal is to fabricate devices with the right metal/TMDC combinations to achieve low-resistance contacts. We obtain monolayers of TMDCs and few-layer BN by mechanical exfoliation using the Scotch tape method. Electron beam lithography, followed by metal evaporation, is used to make pre-patterned contacts on BN. The second BN is then picked up with polycarbonate stamp on a glass attached to a micromanipulator. The van der Waals interaction makes it possible to pick up monolayer TMDC with BN and place it on top of pre-patterned contacts. After the device fabrication, the electrical and magnetic measurements are performed in a cryostat with temperatures ranging from 300K to 1.4K. If successful, the low contact resistance would allow the study of high-temperature interlayer exciton BEC, as well as many other quantum transport phenomena.

Heat Shock on *C. elegans*

Myles Ingram
Biophysics, 2018

Harvard University

Advisor: Erel Levine, PhD

Mentor: Kyung Suk Lee, PhD

Stress responses across organisms have many similarities. The nematode *Caenorhabditis elegans* is a simple model organism with relatively predictable responses to various stresses. Their response to heat shock is especially useful to study because it involves stopping the unfolding of proteins. Researching the heat shock response pathway, along with other stress response pathways, has led to a deeper understanding of the biological mechanisms involved in complicated human diseases such as Alzheimer's and Parkinson's.

The heat shock response of *C. elegans* has been well-characterized by subjecting these worms to uncomfortable amounts of heat and subsequently monitoring their hormesis (positive response to low stress exposure) and their regulation of heat shock proteins (HSP).

However, few fatal heat shock experiments have been conducted to describe the survival of *C. elegans* and little experimentation has been done to describe the motility of *C. elegans* after a fatal amount of heat shock. This manner of killing will expand upon what is already known about *C. elegans* and their heat shock response.

In these experiments, we aim to characterize both the survival and motility of four *C. elegans* strains subjected to heat shock, leading to the death of a significant subpopulation. Survival in wild-type worms after heat shock was compared to that of two different mutants—one lacking the primary thermosensory neurons, AFD, and another lacking the thermo-sensory interneuron, AIY—while monitoring the expression of *hsp-16.2*, an intestinal heat shock protein. Without AFD neurons, worms cannot sense extreme temperatures and without the AIY neuron, worms cannot sense extreme cold. We counted the number of dead, alive but not motile, and motile worms at five different time points to describe survival and motility. We test for motility by tapping the worms on the nose, a well-known *C. elegans* reflex point.

To deepen our understanding of extreme heat shock in relation to motility and survival, we also stressed the worms with mild heat shock and starvation before the fatal heat shock. We stress the worms with

mild heat and starvation to understand the effect of hormesis and to investigate the correlation between different stresses. In addition to evaluating motility and survival, we also monitor the expression of *hsp-16.2* in the two mutant strains (ERL35 and ERL36) and the reporter strain (TJ3001) after the fatal heat shock. With these experiments, we hope to demonstrate that the inactivation of certain neurons included in the heat shock response can have a positive or negative effect on both survival and motility in *C. elegans* exposed to a fatal dose of intense heat. Our work will broaden the literature concerning the roles of heat shock proteins and regulation mechanism of their expression.

Quantifying Thermodynamic Properties in Graphene Heterostructures

Shaan Ajay Desai
Physics, 2018

Harvard University

Advisor: Philip Kim, PhD

Mentor: Jonah Waissman, PhD

Modern condensed matter physics has succeeded in creating electronic devices of exceptional purity, allowing for the discovery of the fractional quantum hall effect and electron hydrodynamics (Dirac fluid), among other phases of matter. While these successes are attributed to electronic transport and resistance measurements, many underlying properties of these phases remain unknown and unverified. It is for this reason that thermodynamic measurements are crucial to the study of condensed matter physics. Quantities such as specific heat, magnetization, and thermal conductivity can reveal fundamental properties underlying material and electronic phase transitions that cannot otherwise be detected using transport systems.

The experimental challenge that lies in determining thermodynamic properties is the scale of nano devices. The specific heat of a two-dimensional electron system in which the quantum hall effect was discovered, for example, is dominated by the 3D crystal structure surrounding the electrons, making for low noise-to-signal ratios and hardly detectable signals. However, the recent discovery of graphene and related two-dimensional materials such as hexagonal boron nitride now provides a new way forward. These 2D materials can be stacked together simply via van der Waals attraction, which allows both the creation of clean electronic devices (as it eliminates our need to grow layers), and heat to be trapped in an individual 2D material layer, eliminating the need for multi-layered stacks.

Through electrically isolating graphene with boron nitride we are able to create gated samples where changes in capacitance can be sensitively measured since boron nitride can be made relatively thin and still maintain its electrically isolating properties. Measuring fluctuations in capacitance informs us of how the density of states varies as the chemical potential of graphene changes. This is useful since it enlightens us as to the geometry of these unique electronic phases. By using amplifier circuits, thermal oscillations can be converted to alternating voltages, which allows us to determine the fluctuations in displacement current across the capacitor. In using this technique, we should theoretically be able to detect how thermal fluctuations influence displacement current, and in turn, how they influence changes in entropy. The end goal is to make repeatable measurements of entropy variations in graphene at ultra-low temperatures and be able to derive and predict how these variations take shape.

Whole Brain Imaging to Determine Neuronal Dynamics in *C. elegans*

Benjamin Sorscher
Physics & Mathematics, 2018

Center for Brain Science, Harvard University

Advisor: Aravinthan D.T. Samuel, PhD

Mentor: Vivek Venkatachalam, PhD

How is it that networks of interconnected neurons enable animals to process sensory inputs and encode appropriate behavioral outputs? Our own brains contain many billions of neurons, yet the roundworm *Caenorhabditis elegans* is capable of strikingly complex processing with a mere 302 neurons. *C. elegans* can translate a diversity of sensory stimuli (temperature, odor, electric field, touch) into one of several mutually exclusive, stereotyped behaviors (forward and backward locomotion, left and right turns, and reversals). By studying the dynamics of the *C. elegans* nervous system we may begin to understand the fundamentals of how this complex encoding of behavior takes place.

Thanks to advances in confocal microscopy, it is now possible to image the whole brain of freely moving *C. elegans* at single-cell resolution. Further, a pan-neuronal promoter enables us to express fluorescent calcium indicators in every neuron of the *C. elegans* nervous system. Calcium ions play a crucial role in intercellular signaling, and thus reflect the activity of an individual neuron. Combined, these techniques afford an unprecedented peek at the whole *C. elegans* brain as it processes information. In order to

study the global dynamics of neural activity, I record whole-brain movies as the animal interacts with sensory stimuli and behaves in its environment. From these movies, I extract a time series of activity for each neuron, and look for patterns of activity between neurons. Using principal component analysis, I can reduce the dimensionality of the pan-neuronal activity data and identify dominant signals across many neurons that correspond with specific behavioral states. I can then employ statistical techniques like clustering and empirical dynamic modeling to identify structure in the neural activity data and isolate potential signaling pathways and sensorimotor circuits. By studying the whole-brain dynamics that translate sensory stimuli into the neural antecedents of behavior in *C. elegans*, I hope to begin to understand how neural networks are able to solve the problem of encoding sensory information into behavior.

Exploring the Phase Diagram of Thin-Film BSCCO

Margaret Panetta
Physics & Astrophysics, 2017

Harvard University

Advisor: Philip Kim, PhD

Mentor: Frank Zhao

When thinned to a few atomic layers, superconductors behave differently than they do in their bulkier forms. While electrons in thin film superconductors still form "Cooper pairs" that normally condense into a superconducting state, when a magnetic field is applied to these materials, their superconductivity is destroyed even at the lowest experimentally achievable temperatures. A two-dimensional "Bose metal" state, the existence of which is debated, theoretically lies between the insulating and superconducting regimes allowed for a material capable of superconductivity. We are using resistance measurements to investigate this transition for thin films of $Bi_2Sr_2CaCu_2O_{8+x}$, called BSCCO, a copper-oxide (cuprate) high-temperature superconductor.

BSCCO consists of layers of copper oxide intercalated with layers of its other constituent materials, allowing production of extremely thin films incorporating only a few iterations of the fundamental repeating crystal unit. We will measure the resistance of a thin-film sample down to the millikelvin range under an applied magnetic field, seeking evidence of the quantum metallic state at one thousandth of the BSCCO superconducting transition temperature of 85K. To do this, we will use a novel stencil-mask technique to evaporate gold electrical contacts onto few-layer

flakes of BSCCO. We will then chill these samples in a cryostat and supply a current, tracking the associated voltage drop between contacts at varying temperatures and field strengths. Starting with thicker flakes, we will decrease their thickness over the course of iterated device fabrication to explore what electrical contact to thin BSCCO is possible. Production of devices aims for the single-layer limit, which will be most useful for possible observation of a Bose metal state far below the superconducting transition.

Stem Cell and Regenerative Biology
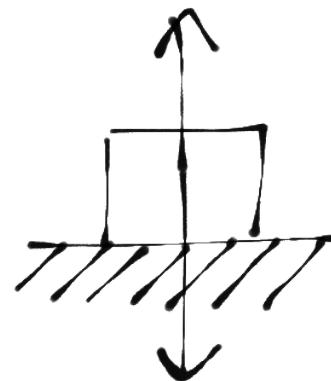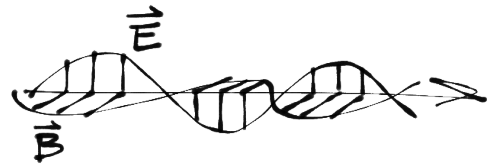
NEUROD6 as a Novel Therapeutic Target for Parkinson's Disease

Alex Pai
Human Developmental and Regenerative Biology, 2019

Harvard Stem Cell Institute

Advisor: Lee Rubin, PhD

Mentor: Tim Ahfeldt, PhD

Dopaminergic (DA) neurons play an important role in cognitive and motor function. In the midbrain, DA neurons are prominently found in the substantia nigra pars compacta (SNpc) and the ventral tegmental area (VTA). Of interest is the degeneration of DA neurons in the SNpc as a phenotype of Parkinson's disease (PD). Though DA neurons in the SNpc and VTA share a common neural progenitor, VTA DA neurons are much less affected in PD. An RNA-sequencing experiment conducted using DA neuron populations extracted from different brain segments of reporter mice identified *Neurod6* as a gene that is differentially expressed between DA neurons in the SNpc and the VTA. The role of *Neurod6*, which is exclusively expressed in VTA neurons, is not well understood outside its role as a transcription factor that regulates cell fate, but recent studies suggest that *Neurod6* plays a role in mitochondrial biogenesis and maintenance that may explain the resistance of VTA DA neurons from the cellular stresses associated with PD.

To explore our hypothesis that in humans *NEUROD6* has neuroprotective effects, we plan to use an *in vitro* model utilizing a tyrosine hydroxylase (TH) reporter line of human pluripotent stem cells (hPSCs) transduced with a lentivirus to allow for inducible expression of *NEUROD6*. We will induce expression of *NEUROD6* in differentiated neurons, which fluoresce red under the microscope, and use a cell stress and death assay to quantify survival under various mitochondrial stresses. If *NEUROD6* is found to be neuroprotective, targeted expression in SNpc DA neurons to confer resistance may represent a novel therapeutic avenue in PD.

Investigating the Role of Non-Classical Progesterone Receptors in the Regulation of HLA-G Expression in EVT

Andrew Mazzanti
Human Developmental and Regenerative Biology, 2017

Harvard Stem Cell Institute

Advisor: Jack Strominger, MD

Mentors: Leonardo Ferreira, PhD; Torsten Meissner, PhD

HLA-G is a major histocompatibility complex (MHC) protein that, unlike most other types of HLA, is not involved in peptide presentation to T cells. Instead, it interacts with NK cells to prevent their cytolytic activity. NK cells are innate immune cells that can target and kill cells infected by virus or cells found to be foreign or "non-self." Suppression of the NK cell-mediated immune response is potentially dangerous, and accordingly, the expression of HLA-G is limited to specific tissues and cell types, including the placenta. Extravillous trophoblasts (EVT) are fetal-derived cells that invade maternal tissue during the course of pregnancy. Though fetal-derived cells express paternal antigens, and EVT is exposed to the maternal immune system, NK cells do not lyse EVT due to the expression of HLA-G. Progesterone signaling has been shown to increase the expression of HLA-G in EVT. However, EVT are devoid of the classical progesterone receptors PR-A and PR-B. Non-classical receptors must be investigated to better understand how progesterone regulates the expression of HLA-G in EVT.

The approach to this research has three steps. First, candidate non-classical progesterone receptors will be identified in consultation with the current literature. Candidate genes will be knocked out from JEG-3 cells, a cell line model of EVT, via the CRISPR-Cas9 system. Next, HLA-G expression will be measured at the mRNA and protein levels both before and after treatment with progesterone and various progesterone analogs to determine which if any of the gene knockouts impact the ability to respond to progesterone. Finally, expression of knocked out genes will be restored via an expression vector construct to rescue the knockout phenotype. At the same time, overexpression of gene candidates can be assessed for any impact on HLA-G expression. Currently, two knockout lines are being screened via FACS and western blot while two more knockout lines are being gen-

erated in parallel. Overexpression experiments will follow confirmation of successful knockouts.

If a knockout line fails to upregulate HLA-G after progesterone exposure, that gene will be a likely candidate for the non-classical progesterone receptor. Since progesterone-mediated upregulation of HLA-G is essential to the survival of the fetus and to successful pregnancies, better understanding the signaling pathway could lead to the development of treatments to reduce the incidence of pre-term birth and increase the rate of successful pregnancies.

TDP43 Mislocalization in ALS and Related Neurodegenerative Diseases

Christopher Cantrell
Human Developmental and Regenerative Biology, 2018

Harvard Stem Cell Institute

Advisor: Kevin Eggan, PhD

Mentor: Daniel Mordes, MD

TDP43 is a gene involved in ALS and other neurodegenerative conditions. The protein product of this gene, TDP43, is normally located in the nucleus of neurons, but when one has a neurodegenerative disease, the protein is mislocalized into the cytoplasm of the cell. TDP43 has various functions, including mRNA transport and transcriptional repression. This protein mislocalization results from protein imbalances in the neurons. My project is to build a model of this neurodegenerative condition and to then determine what causes the TDP43 protein to mislocalize from the nucleus into the cytoplasm of neurons. I have also been developing methods to quantify the change in cellular localization using fluorescent microscopy and analytic software called CellProfiler. Additionally, I have been exploring if other pathways in neurodegenerative diseases cause a similar TDP43 mislocalization.

My experiments involve the use of cell lines that are easier to work with than neurons, which is necessary in order to build my model of neurodegenerative disease conditions. For instance, I have used human bone cancer (u2os) cells, human embryonic kidney cells, SHSY5Y neuroblastoma cells, and SKNFI neuroblastoma cells. I have gained experience in working with and growing stem cells, and have also become familiar with the methods to generate neurons from stem cells. I have also gained proficiency in the various techniques required to successfully conduct experiments involving tissue culture, including methods that will allow visualization of the protein of interest.

I have taken fluorescent images of both untreated cells and cells treated with various drugs, including proteasome and autophagy pathway inhibitors, and have worked with the CellProfiler program to create an analysis pipeline that can show whether or not my neurodegenerative disease model works. The value I use to analyze my results is the cytoplasmic fluorescence divided by the nuclear fluorescence (C/N value); thus, the higher the C/N value, the more cytoplasmic TDP43 there is relative to the TDP43 in the nucleus of the cell. The analysis software helps me calculate the C/N value for each individual cell, and then I qualitatively evaluate the results via histograms and quantitatively evaluate them using a 2-group t-test. The t-test can prove that there is a statistically significant difference between the untreated and treated cells' C/N values, and can therefore prove that TDP43 mislocalization to the cytoplasm is occurring.

Preliminary experiments demonstrate that proteasome inhibition causes TDP43 mislocalization in u2os cells, and the same seems to be true in neuroblastoma cells. However, it does not appear that the use of autophagy inhibitors causes the same result. I will continue to test my model to ensure that it effectively models TDP43 mislocalization under neurodegenerative conditions via immunofluorescence imaging and the CellProfiler analysis program. Also, my experiments will include the use of neurons I differentiate from human stem cells. If my analysis consistently shows that my model works, I will proceed with my project to conduct tests of association to attempt to discover which related pathways cause TDP43 mislocalization, as that would have implications for ALS and other neurodegenerative diseases.

A Cas9-Mediated Genome-Wide Screening Approach to Identify Novel Regulators of Primed Pluripotency in Epiblast Stem Cells

Christopher Li
Human Developmental and Regenerative Biology, 2018

Harvard Stem Cell Institute, Massachusetts General Hospital

Advisor: Konrad Hochedlinger, PhD

Pluripotency is a feature of embryonic stem cells (ESCs) and describes the cellular potential for multi-lineage differentiation—the ability to give rise to cell types of all three embryonic germ layers. *In vitro*, ESCs can self-renew indefinitely while maintaining pluripotency. Recent investigation of the signaling mechanisms that regulate configurations of pluripotency has revealed two distinct states—naive and primed—each with unique epigenetic, molecular, and functional characteristics.

In mice, naive embryonic stem cells are derived from the pre-implantation blastocyst, while primed epiblast stem cells (EpiSCs) are derived from the post-implantation epiblast. Intriguingly, human ESCs are thought to be more closely related to primed EpiSCs than to naive ESCs in terms of their developmental potential as well as their signaling requirements and epigenetic properties, which therefore offers an invaluable model to study primed pluripotency using transgenic tools readily available in mice but not in humans. Although the molecular and signaling pathways that sustain naive pluripotency in murine embryonic stem cells (mESCs) are rather well described, the mechanisms regulating primed pluripotency remain incompletely understood. We are particularly interested in the mechanism of exit from pluripotency in EpiSCs, as well as in the similarities and differences between the maintenance of human and murine primed pluripotency.

The CRISPR-Cas9 system for genome editing is a robust method to systematically assay loss of function phenotypes and is a powerful tool in mammalian genetics. To further elucidate the molecular mechanisms controlling pluripotency, we are performing a comprehensive Cas9-mediated genome-wide screen to identify novel regulators of the signaling pathways that promote or permit exit from pluripotency in EpiSCs. We exploited the use of an Oct4-GFP reporter of the pluripotent state, which is rapidly and completely downregulated when EpiSCs exit pluripotency, and derived a TetO-Cas9/Oct4-GFP EpiSC line from E6.5 post-implantation epiblasts. After Cas9 induction followed by infection with a lentiviral genome-wide gRNA library, we withdrew growth factors necessary to sustain pluripotency. Cells maintaining Oct4-GFP reporter expression are likely to carry knockouts for genes required for differentiation, and subsequent isolation of these cells and sequencing of integrated gRNAs will identify the mutated genes of interest, which prevent EpiSC commitment.

This work will help to reveal novel regulators of pluripotency and to dissect the mechanism by which primed pluripotency is maintained in murine EpiSCs, and these studies will inform future efforts to generate human naive ESCs. Our findings should further our understanding of the pathways that govern stem cell pluripotency, and hold significant implications for clinically relevant strategies to modulate stem cell fate in regenerative medicine.
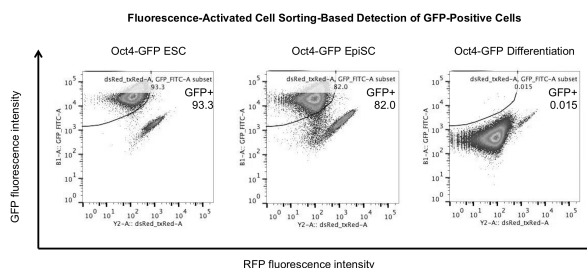

Figure 1: GFP expression in Oct4-GFP ESCs, EpiSCs, and EpiSC-differentiated cells as measured by flow cytometry. These data validate our use of an Oct4-GFP reporter of pluripotency, which is rapidly and completely downregulated upon EpiSC exit from pluripotency.

Effects of DNA Damage on Fibrotic Cell Cycle Arrest in Kidney Organoid Models

Edgar Garcia
Human Developmental and Regenerative Biology, 2018

Brigham and Women's Hospital

Advisor: Joseph Bonventre, MD, PhD

Mentors: Ryuji Morizane, MD, PhD; Navin Gupta, MD

Chronic kidney disease (CKD) is an illness affecting ten percent of the worldwide population, resulting from continuous or repeated injury to the kidney. Kidney fibrosis, a hallmark of CKD, is scarring that results from failure to regenerate kidney tubules after injury. However, little is known about the fibrotic pathway and the exact mechanisms by which one could suppress it.

Mouse models have related fibrosis to cell cycle arrest in the G2/M checkpoint (Yang et al., 2010). I believe that DNA damage from chemotherapeutic treatments likely results in G2/M arrest and leads to fibrosis.

Current animal models are limited in their ability to effectively model human pathology, especially in the kidney where there is considerable heterogeneity of solute transporters. Therefore, kidney tissue developed from human pluripotent stem cells (hPSCs) is an attractive model to explore mechanisms of renal fibrosis linked to DNA damage in human tissues. Using hPSC-derived human kidney organoids as highlighted in Morizane et al.'s (2015) Nature Biotechnology paper, I hope to explore the associations between DNA damage, G2/M checkpoint arrest, and kidney fibrosis.

To do so I used cisplatin, a chemotherapeutic drug that causes proximal tubular injury and DNA damage. I began by finding the concentration and interval of cisplatin that best induces fibrosis and minimizes generalized nephrotoxicity. After differentiating the hPSCs, I treated the organoids with 5 μM cisplatin for 24 hours twice weekly. I collected samples, and their corresponding controls, after 1, 3, and 5 treatments. I then assayed for fibrosis at the mRNA level using quantitative reverse transcription PCR (RT-PCR) and at the protein level using immunohistochemistry (IHC). I saw an upregulation of fibroblast signaling (p<0.001) after treating with cisplatin at least three times.

After showing the induction of a fibrotic-like state in the kidney organoids, I went on to characterize the cell cycle distribution within the treated samples. I used phospho-histone H3 (HH3) staining to assay for cell cycle distribution within the cisplatin-treated organoids. HH3 is differentially phosphorylated at the various stages of the cell cycle, giving it distinct staining patterns for each of the different stages of the cycle. Using this knowledge, it appeared that there was greater skew towards G2 arrest within the cisplatin treated samples when compared to their negative controls.

Finally, I sought to modulate the fibrosis pathway and characterize the fibrotic response resulting from these modulations. I did so by using a small molecule inhibitor that induces cell cycle arrest in G2/M. Following these experiments, I performed similar staining and quantitative RT-PCR to assay for changes in fibrotic levels.

These experiments thus far support the hypothesis that there is a link between DNA damage, G2/M arrest, and fibrosis. With the results from the final phases of my experiments, I hope to fully determine the molecular relationship between fibrosis and CKD.

The Role of the ARRDC3 in Fructose Absorption in the Small Intestine

Ellen Zhang
Human Developmental and Regenerative Biology, 2019

Harvard Stem Cell Institute

Advisor: Richard Lee, MD

Mentor: Shannon Carroll, PhD

The intake of fructose in modernized societies has increased dramatically in the past decades. Excess fructose may be a significant factor in metabolic disorders like diabetes and non-alcoholic fatty liver disease. Therefore, a complete understanding of fructose metabolism has important implications for human health.

Following ingestion, fructose is transported from the small intestinal lumen into the enterocyte via glucose transporter 5 (GLUT5). GLUT5 is a specific transporter for fructose and is expressed exclusively on the apical membrane of the enterocyte (Douard & Ferraris, 2008). Fructose exits the enterocyte and enters the portal bloodstream through the activity of glucose transporter 2 (GLUT2), which transports both monosaccharides and is expressed on the basolateral membrane of the enterocyte.

Previously, Dr. Lee's lab found a protein called thioredoxin-interacting protein (Txnip) that interacts with both GLUT5 and GLUT2. Utilizing a Txnip knockout mouse, they found that Txnip is necessary for fructose absorption in the small intestine of mice. However, another protein closely related to Txnip called Arrestin domain containing 3 (Arrdc3) was also found to interact with GLUT5 and GLUT2. This discovery suggests that Arrdc3 may also regulate fructose absorption in the small intestine in a manner similar to Txnip; alternatively, Arrdc3 may have other effects on fructose absorption.

Little is known of Arrdc3's function or expression pattern, although the gene for Arrdc3 is linked to obesity in men. If Arrdc3 has a role in fructose absorption in the small intestine, we hypothesize that Arrdc3 will be expressed in the enterocytes of the small intestine. We will test this hypothesis using two different models. First, we will utilize an intestinal human cell line (Caco-2 cells) to study

ARRDC3 mRNA expression *in vitro* via quantitative PCR. Secondly, we will utilize a ARRDC3/lac-z reporter mouse to track expression of Arrdc3 in the small intestine. This will require isolation and β-galactosidase staining of the small intestine followed by histological sectioning and microscopy.

This research will give valuable information about Arrdc3 expression in enterocytes and the small intestine which is necessary for continued research on the role of Arrdc3 on fructose absorption. This research will also provide a foundation for understanding Arrdc3 expression in the enterocyte is regulated.

Enforced Expression of E-Selectin Ligand on Mesenchymal Stem Cell-Derived Exosomes via Exofucosylation

Gerardo Castillo
Molecular and Cellular Biology, 2018

Brigham and Women's Hospital, Harvard Medical School, and Program of Excellence in Glycosciences

Advisor: Robert Sackstein, MD, PhD

Mentor: Gisela Pachon, PhD

Exosomes are cell-derived microvesicles which are secreted from many different cell lines, including human mesenchymal stem cells (hMSCs), and have been shown to play a role in intracellular signaling. Specifically, prior studies have shown that exosomes derived from hMSCs reduce myocardial ischemia, induce peripheral tolerance, trigger autoimmune responses, and repair acute kidney injuries. Therefore, it is possible that isolated hMSC-derived exosomes can be used as a therapy at specific sites and tissues.

Prior studies have also shown that it is possible to direct certain cells to specific sites given the presence of E-selectin and a potent E-selectin ligand, which allows for transmigration of the cell through the vascular endothelial wall. The multistep paradigm of transmigration is dependent on tethering/rolling interactions (step 1 of cell migration) which involves the binding of ligands expressed on cell surfaces, mainly E-selectin ligands, to selectins expressed on vascular endothelium. Research conducted in the Sackstein lab has data to suggest that hMSCs lack E-selectin ligands that are important for transmigration. However, these cells do express native CD44 and could be modified via a process known as exofucosylation through glycosyltransferase programmed stereosubstitution (GPS) to enforce E-selectin-binding sialyl Lewis X (sLe^x) decorations, thereby creating the most potent E-selectin ligand known as hematopoietic cell E-/L-selectin ligand (HCELL).

Creating HCELL on exosomes has many clinical applications; in therapeutics, stem cell transplants have mixed results due to the variability rate of cells to transmigrate across the vascular endothelial to the bone marrow. Glycan engineering to create HCELL expression allows the efficient vascular delivery of specific cells due to the nature of strong step 1 interactions that HCELL can undergo with E-selectin.

Here, we sought to analyze the effect of exofucosylation and the creation of E-selectin ligands on exosomes using α-(1,3) fucosyltransferases. The current study involved the isolation of exosomes derived from hMSCs through differential ultracentrifugation. Exosomes were bound to 4 μM aldehyde/sulfate latex beads overnight via passive adsorption. Identification and characterization of exosomes was performed through the use of a panel of cell surface markers, such as CD81 (a specific biomarker of exosomes), CD63, CD73, and CD44 (specific markers of hMSCs) by flow cytometry. Exosomes were exofucosylated with α-(1,3) Fucosyltransferase VI (FTVI) via GPS and HECA-452 (mAb which recognize sLe^x structures) reactivity was analyzed to identify the enforced expression of sLe^x determinants on exosomes. This study supports the utility of cell surface glycan engineering to enforce the migration of hMSCs-derived exosomes to sites of tissue injury or inflammation.

Designing a New High-Throughput Screen for Molecules that Contribute to Cell Fate Determination

Katie Kixmoeller
Chemistry and Physics, 2018

Harvard Stem Cell Institute

Advisor: Douglas Melton, PhD

Mentor: Nadav Sharon, PhD

Type 1 diabetes is a common disease in which the insulin-producing beta cells of the pancreas are destroyed by an autoimmune response. Patients can't produce insulin, so they can't respond properly to an increase in blood glucose. Regenerative medicine presents the possibility for a cure. A few years ago, the Melton lab published a protocol for creating human stem cell-derived beta cells with the goal of providing rejection-free beta cell transplants to patients with severe diabetes. The protocol mimics normal pancreatic development, guiding stem cells along the process of differentiation into beta cells. It is an astonishing accomplishment, but many challenges

remain before these cells can be given to patients. One significant difficulty is that the yield of beta cells remains low. Also, the hallmark of functional beta cells is glucose-stimulated insulin secretion (GSIS), and the cells produced can't perform GSIS as well as normal beta cells.

The goal at each stage of the protocol is to have a homogenous population of cells so that the final product is homogenous beta cells. However, at one of the intermediate stages of the protocol, we identify two populations of cells, pancreatic progenitors and more-differentiated endocrine cells. This heterogeneity lowers the yield of the protocol, so we need to adjust the ratio between the two populations.

High-throughput molecular screening could help solve this problem by identifying molecules that affect this ratio. However, acquisition of cell fate is a gradual and complex process. It involves a combination of molecular signals regulating the expression of hundreds of genes. So, it is extremely unlikely that a single molecule alone could cause a cell to acquire a particular fate. Traditional high-throughput screens are poorly suited to identify molecules that have only a partial effect for two reasons. First, the output is generally binary (cell alive or dead, protein present or absent). Second, these screens are limited in the number of genes that can be tested at a time.

I set out to design a new high-throughput screen to identify molecules that contribute to determining cell fate. Although each molecule alone wouldn't determine cell fate, some combination of them could be used in the lab's protocol to adjust the ratio between cell populations.

Recent developments in single-cell RNA sequencing technology allow analysis of multiple genes across multiple samples. My screen harnesses this technology for high-throughput screening. Additionally, the screen's output is sequencing data, allowing for more sensitive detection of gene expression levels than traditional outputs such as fluorescence. My project is to design, test, and calibrate this new screen with the goal of deploying it experimentally in the fall.

This screen could also be applied to other research questions, both in the Melton lab and beyond. We also plan to look at cells that have reached the end of the lab's protocol but do not perform GSIS as well as normal beta cells. This screen has great potential to improve the differentiation protocol, which is an important step toward clinical applications of stem cell derived beta cells.

Regulation of Cardiac Regenerative Capacities

Krystal Phu
Human Developmental and Regenerative Biology, 2019

Harvard Stem Cell Institute

Advisor: Richard Lee, MD

Mentors: Ahmed Mahmoud, PhD; Christian Shigley

Heart attacks are a leading cause of death in adults worldwide because they lead to the loss of heart muscle cells, known as cardiomyocytes. Cardiomyocytes in adult humans regenerate very slowly, and the inability to replace dying cardiomyocytes leads to decline of the pumping performance of the heart. In contrast, newborn mammals, or neonates, can fully regenerate damaged cardiomyocytes, though this capability is lost shortly after birth. This regeneration occurs by cell division of other pre-existing cardiomyocytes and not by stem cells. It has been speculated that a critical step in cell division of cardiomyocytes is disassembly of the rigid sarcomere structures that provide pumping function but could otherwise impede mitosis.

One theory for the loss of regenerative capability in adult human hearts is that neonatal cardiomyocytes mature into high performance adult heart cells that are less able to divide. Previous work in the Lee lab has found transcription factors that are expressed differentially between the neonatal and adult cardiomyocytes. Of these, several transcription factors have previously been associated with sarcomere formation or disassembly. Although these factors have other known functions, their capacity to regulate sarcomeres may be important for cardiomyocyte division potential.

To evaluate the impact of these factors, we use an N-1 combinatorial approach via lentiviral transduction into adult rat cardiomyocytes, which normally are very resistant to cell division. Then we immunohistochemically stain cardiomyocytes for sarcomere-related markers and utilize imaging techniques to measure morphological differences in sarcomere structure between cells receiving the candidate factors in the experimental conditions and those receiving empty lentiviral vectors in the control condition. We hypothesize that in conditions wherein key transcription factors associated with sarcomeres are necessary for cardiomyocyte division, those factors cause sarcomeres to be disassembled prior to the mitosis event.

If our experiment performs as hypothesized, we

will have identified factors capable of disassembling sarcomeres that can participate in cardiomyocyte division. Future study will test the impact of expression of these factors on sarcomere assembly in the context of adult cardiomyocyte proliferation *in vivo*. This work could represent a step toward transiently reprogramming adult cardiomyocytes into a regenerative state.

Hematopoietic Stem Cell Gene Therapy for Very Early Onset Inflammatory Bowel Disease Caused by IL10 Receptor Deficiency

Linda Qin
Human
Developmental and
Regenerative Biology,
2019

Harvard Stem Cell
Institute, Boston
Children's Hospital

Advisor: David A.
Williams, MD

Mentor: Christian
Brendel, PhD

Loss of function mutations in IL-10 or the IL-10 receptor lead to life-threatening very early onset inflammatory bowel disease (VEO-IBD). Patients with VEO-IBD present with severe symptoms from a very young age and are largely unresponsive to conventional anti-inflammatory treatments. Allogeneic hematopoietic stem cell transplantation (HSCT) is the only available curative treatment option, but the limitations of finding suitable HLA-matched donors and the underlying risks of acute toxicities and graft-versus-host-disease (GVHD) are significant deterrents.

However, preliminary data from our group have shown that, in mice, the transfer of the IL10R gene into IL10R-deficient bone marrow precursors followed by HSCT can prevent the development of IBD. This suggests that lentiviral gene therapy of autologous hematopoietic stem cells (HSCs) may carry potential to be a safe and effective alternative. Additionally, recent data indicates that transfer of healthy anti-inflammatory macrophages may be sufficient to prevent colitis in IL10R-deficient mice. These data suggest that a feasible therapeutic option may entail Cas9-mediated gene editing in patient derived induced pluripotent stem cells (iPS) to correct the defective gene, differentiate iPS cells into mature macrophages *in vitro*, and subsequently inject cells into the patient. Both approaches require suitable *in vitro* test systems to evaluate and optimize the efficacy of gene transfer or gene editing in a simple system and in a time-effective manner, but currently neither human nor murine cell lines with a complete IL10R knockout do exist.

We will first create human knockout cell lines of both IL10Rα and IL10Rβ using CRISPR/Cas9. Clonal lines with gene inactivation by frameshifts or by complete deletion of the coding region will be established by single cell cloning. The knockout will be confirmed by western blot and the ability or inability to phosphorylate STAT3 in response to IL10 stimulation. Such a knockout model will be useful in testing lentiviral and alpharetroviral gene transfer vectors with different configurations to identify the most promising candidates for *in vivo* mouse experiments. Furthermore, the knockout cells will provide a comprehensive method to optimize a toolset for gene editing which will later be used to correct patient derived iPS cells. This includes identification of the most effective Cas9 nuclease type, the specific guide RNAs, and the size and location of the repair template to achieve specific and efficient expression of the IL10 receptor. These knockout cell lines will eventually be used to test the species cross reactivity of components of the IL10-receptor signaling pathway and to establish the required expression strength of IL10R chains required for full signaling competence, and will greatly facilitate the testing and selection of gene therapy vectors and gene editing tools required to develop novel therapies.

Towards Vascularizing Cerebral Organoids *In Vitro*

Lucy Nam
Molecular and
Cellular Biology,
2017

Wyss Institute for
Biologically Inspired
Engineering

Advisor: Jennifer
Lewis, ScD

Mentor: Mark
Skylar-Scott, PhD

The creation of *in vitro* organ models of human development is an important and emerging area of study, with applications in drug screening, disease modeling and tissue engineering. In particular, induced pluripotent stem cells (iPSCs) can self-organize into three-dimensional tissue-like architectures, or "organoids," enabling the real-time observation of complex developmental processes. If patient-specific iPSCs are used in deriving the organoids, specific pathologies can be recapitulated and studied *in vitro*. However, to date, existing organoids have been avascular, and subsequently develop necrotic cores that limit the size and complexity of the resulting tissues.

To construct vascularized cerebral organoids, or cerebrovascular organoids, iPSC cultures must be differentiated into both neural (ectoderm) and vascular (mesoderm) tissues. However, generating pure populations of tissues from distinct germ layers using

a single media condition is unlikely to be successful. Thus, we used genomic engineering to generate two separate populations of iPSCs: a wild-type population, and a second population that underwent differentiation into vasculature via a drug-induced upregulation of the transcription factor ETS-variant 2 (ETV2).

When these cells were co-cultured in proneural media conditions in the presence of doxycycline, and implanted into a Matrigel extracellular matrix, the wild-type iPSCs form cerebral organoids, while the dox-inducible ETV2 iPSC population differentiates into endothelium that express both CD31 (PECAM-1) and vascular-endothelial cadherin (VE-Cadherin). These endothelial cells further undergo vasculogenesis and angiogenesis to generate a three-dimensional microvascular network that surrounds the developing cerebral tissue.

Using this method, we aim to perfuse the microvascular network that surrounds the cerebral organoids to provide the requisite oxygen, nutrients and waste removal for the developing tissue to thrive. We will achieve this by embedding the cerebrovascular organoids in-between two pin-casted microchannels that are connected to an external peristaltic pump for perfusing fresh, oxygenated media through the tissue. Once the ETV2 cell-derived microvasculature connects across the two microchannels, the media flow will be re-directed through the endothelial network to nourish the tissue. We will monitor the subsequent growth and development of the tissue to identify whether a primitive blood brain barrier (BBB) will develop, and whether we can prevent the formation of a necrotic core as the tissue further grows.

In Vitro Modeling of the Gastro-Hypothalamic Axis to Elucidate Appetite and Satiety Signals

Luke Smith
Biomedical Science, 2017
Emmanuel College

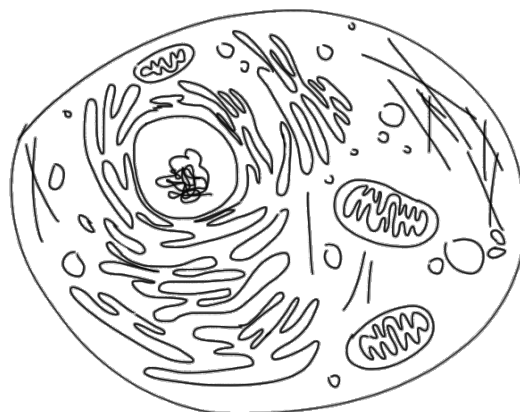Harvard Stem Cell Institute

Advisor: Qiao Zhou, PhD

Despite comprising a small percentage of the adult human brain, the hypothalamus is a complex and essential neural element regulating a variety of systems, including energy homeostasis. It may regulate a "set point" for body weight. Neurons in this region of the brain are well placed to sense neural and physiological signals from the alimentary tract via hormones and the vagus nerve, relaying this to high order structures that influence behaviour. Unlike other neurons, in which signals decay in milliseconds, many hypothalamic neurons have been shown to be capable of retaining signals over extended timescales, allowing for integration of temporally distant events. Loss of this ability has been seen to cause weight gain in the case of POMC and PVH neurons or loss for AGRP neurons in mice. Understanding connectivity and mechanisms within the hypothalamus, and of its reciprocal interactions with the alimentary tract, are fundamental to developing our knowledge of disorders in weight such as obesity.

Obesity is no longer seen simply within a paradigm of eating too much and exercising too little but instead as a complex interplay of factors. The project centres on the use of human embryonic stem cells to generate human hypothalamic neurons. The ability to do this is central to developing an *in vitro* model of how the gut and hypothalamus communicate, and to identifying novel factors which may have an effect on weight. I will stain the resultant cells to ensure they are expressing elements indicative of hypothalamic lineages and test their electrophysiology.

Upon confirmation of their nature, I will begin investigating the effect of secretions of enteroendocrine cells (produced in the lab through directed re-differentiation) on their behaviour. The hormones secreted by enteroendocrine cells have a putative importance in regulating satiety and eating behaviours. An *in vitro* model would allow for screening of compounds for their possible effects in the hypothalamus; this could yield currently unknown endogenous molecules to pursue in mouse models for their effects on eating. Ultimately, we hope to identify putative targets which can be exploited pharmaceutically to either suppress appetite (in obesity) or stimulate it (such as in patients undergoing chemotherapy).

Evaluation of Candidate Genes for Controlling Limb Size in Vertebrate Regeneration

Rachel Oshiro
Molecular and
Cellular Biology,
2019

Harvard Stem Cell
Institute, Brigham
and Woman's
Hospital

Advisor: Jessica
Whited, PhD

Mentors: Donald
Bryant, MS; Victor
Luria, PhD

Worldwide, ten million individuals suffer from an amputation; currently, the treatments available often struggle to replace the lost limbs. The ultimate goal is to stimulate human limb regeneration. To begin to devise approaches to fulfilling this goal, a thorough understanding of what occurs at the molecular level that enables animals to regenerate is likely to aid in this goal. Unlike humans, axolotls (Mexican salamanders) can regenerate their limbs after amputation. However, the process by which they do this is not entirely understood. We do know that the process involves the formation of a blastema which is thought to be composed of cells that have become "dedifferentiated" (lost their specific function) but might also include contributions from stems cells, another type of undifferentiated cells. The new limb then grows from the blastema cells. Interestingly, nerves are necessary for proper regeneration to occur.

The Whited lab has created animals with "miniaturized limbs" or mini-limbs by repeatedly cutting the limb blastema to prevent proper growth, resulting in a limb substantially smaller than normal in a large fraction of animals. These can be used to understand what drives limb regeneration and what mechanisms integrate limb size and body size. To discover the genes which may control these processes, the lab performed mRNA sequencing, to determine expression levels of mRNA, from mini-limb samples and control sibling samples. The data was analyzed to determine what transcripts were "downregulated" or decreased and what transcripts were "upregulated" or increased in mini-limbs in comparison to the controls. The thresholds were made with expression levels increased by a factor greater than 2 and expression levels decreased by a factor less than 0.5 but greater than 0. Several genes, notably those encoding the CATK (downregulated), FGFR1 (downregulated), and SEM6D (upregulated) proteins, seemed promising as their expression levels varied greatly between the control and the mini-limb axolotl; preexisting studies from other contexts suggest testable hypotheses about the possible roles of these factors in our system.

By using RNA *in situ* hybridization, we plan to validate where these transcripts are expressed in mini-limbs and the controls. It can then be determined what would happen if expression was stopped or increased allowing us to elucidate the transcripts functions within the regenerating limb and possibly how appendage size is matched to body size. We also plan to examine internal tissue composition using immunohistochemistry allowing the visualization of the different types of cells within the limbs. Using this technique, the mini-limbs can be studied in terms of nerves as despite the fact that mini-limbs have fewer targets since they are smaller than regular limbs, they should have the same number of nerve cells.

In the end, we will compare the mini-limbs and normal limbs in order to elucidate what determines and restricts size in regeneration and growth and have an increased understanding of the function of nerves within the regrowing limb.

Using *In Vitro* Mouse and Human Spinal Muscular Atrophy (SMA) models to Investigate the Effect of Survival of Motor Neuron (SMN) Deficiency on Satellite Cells

Soumyaa Mazumder
Molecular and
Cellular Biology,
2019

Harvard Stem Cell
Institute

Advisor: Lee Rubin,
PhD

Mentors: Rebecca
Gibbs; Tim Ahfeldt,
PhD

The leading genetic cause of death for infants, Spinal Muscular Atrophy (SMA) is a neurodegenerative disease characterized by the progressive loss of motor nerve cells in the spinal cord. Affecting approximately 1 in 6,000 infants, SMA is caused by a mutation in the Survival of Motor Neuron 1 gene (*SMN1*). This mutation leads to the insufficient expression of the Survival of Motor Neuron (SMN) protein, which subsequently results in severe muscle weakness and often death.

Although SMN is expressed by all cells, most therapies focus on targeting motor neurons (MNs), as it is thought that MNs are the most severely affected. However, there is considerable evidence from this lab and others to suggest that in addition to the death of MNs, SMN deficiency may have deleterious, independent effects on other cell types, such as muscle itself.

Given these findings, as well as results suggesting that muscle weakness in SMA mice may be in large part caused by postnatal muscle growth abnormalities, this project focuses on understanding the effect of SMN deficiency on a population of cells responsible for postnatal muscle growth and regeneration: satellite cells. During postnatal growth, satellite cells (muscle stem cells) proliferate and differentiate to form muscle fibers. Previous studies have suggested that the loss of SMN may diminish the ability of different cell types to proliferate. Based on my mentor's preliminary experiments, this project aims to investigate if SMN deficiency specifically affects satellite cell proliferation. Our main hypothesis is that the loss of SMN decreases satellite cell proliferation, resulting in a decrease of myogenic precursor cells needed to support postnatal muscle growth and a reduction of the satellite cell population.

To test this hypothesis, we are using an *in vitro* SMA mouse model to identify SMN-dependent defects in satellite cell proliferation. Using mouse embryonic stem cells (mESCs) from SMA and wild-type (WT) mice, I will differentiate mESCs into satellite cells and myofibers by adding myogenesis-promoting factors at appropriate time points. While the entire differentiation process takes place over several weeks, cells from both the SMA and WT line will be fixed weekly and immunostained for markers of satellite cells, myoblasts, myogenic precursor cells, and myofibers. In addition to immunostaining, western blotting and qPCR will also be performed to measure expression of key genes expressed by presomitic mesoderm cells, satellite cells, and other types of muscle cells. These techniques will allow us to track the differentiation of mESCs as they change over time.

In addition to maintaining an SMA mESC line in culture, we are developing an *in vitro* human SMA model. One key project is to develop a PAX7 red fluorescent tdTomato reporter for satellite cells. The transcription factor PAX7 is a key marker of satellite cells, and thus such a reporter would be useful for tracking human induced pluripotent stem cell (iPSC) differentiation and ESC differentiation into satellite cells. In particular, this reporter would be useful for developing future SMA patient-derived iPSC lines.

If our *in vitro* models support the hypothesis that SMN deficiency impairs satellite cells, this would suggest satellite cells as an important, novel therapeutic target for treating SMA.

Evolutionarily Conserved Role of C5aR1 in Cardiac Regeneration

Yamen Abbas
Human Developmental and Regenerative Biology, 2017

Harvard Stem Cell Institute

Advisor: Richard Lee, MD

Mentor: Aysu Uygur, PhD

According to the American Heart Association (AHA), heart failure—a chronic, progressive condition in which the cardiac muscle fails to pump blood efficiently enough to meet the body's need—is a growing epidemic, with more than 870,000 new cases diagnosed annually in the United States. Heart failure (HF) is characterized by progressive loss of cardiomyocytes, which eventually compromises the contractility of the remaining cardiac muscle. Unlike lower vertebrates such as salamanders and zebrafish, mammals fail to regenerate the heart following injury. Thus, developing therapeutic methods that replace the lost cells is an attractive strategy for the treatment of HF. Interestingly, one similarity between regeneration-competent animals is a weak immune systems. The link between the immune system and regeneration has been investigated by many groups; inflammatory response, which is tightly regulated by the immune system, has been shown to either drive or inhibit regeneration depending on context. In our lab, C5aR1, a complement protein, has been identified to be highly upregulated in the hearts of zebrafish, axolotls, and neonatal mice post apical resection injury. In addition, our lab has shown that its activation is necessary for neonatal mice cardiac regeneration. My work focuses on characterizing C5aR1, identifying downstream targets, and performing gain-of-function experiments on adult mice. Thus far, my work shows that contrary to what we would expect, C5aR1 is mostly expressed in cardiomyocytes and endothelial cells, rather than macrophages.

Characterization of the Stem Cell Population in the Acoel *Hofstenia miamia*

Andreé Franco-Vasquez

Human Developmental and Regenerative Biology, 2017

Harvard University

Advisor: Mansi Srivastava, PhD

Regeneration is a fundamental feature of animal biology. Thomas Hunt Morgan's extensive cutting experiments with the planarian *Schmidtea mediterranea* first established it as a regenerative organism capable of whole-body regeneration worth studying further. Since then, endless studies have been conducted in order to completely characterize the process of regeneration in this species, which has turned planarians into the model organism for studying regeneration. Regeneration studies with *Schmidtea*, including the characterization of molecular signaling involved in the process, have revealed some mechanisms underlying regeneration. However, the lack of additional model organisms capable of whole-body regeneration has prevented conclusive comparative analysis of the mechanisms involved in this process.

Hofstenia miamia, commonly known as the three-banded panther worm, is an acoel species capable of whole-body regeneration and has been recently established as a new model system. Several parallels have been discovered between it and *Schmidtea* in terms of regeneration mechanisms, including the presence of pluripotent stem cells, called neoblasts. Neoblasts have been characterized extensively in planarians but have been studied minimally in *Hofstenia*. We aim to identify additional neoblast markers in the new organism by characterizing the expression profile of genes enriched in a neoblast-specific RNAseq dataset through fluorescent *in situ* hybridization. Our preliminary studies have revealed several neoblast markers, some that are shared with *Schmidtea* and some that appear unique to *Hofstenia*, which we are investigating functionally. The parallels in regeneration between the three-banded panther worm and planarians, two species separated by 550 million years of evolution, are numerous and will be crucial for the comparative study of regeneration. Understanding the similarities of the molecular mechanisms controlling this process in both species will lead to the ability to compare regeneration across phyla. Such evolutionary analyses could lead to the identification of currently unknown conserved pluripotency factors that could be essential features of animal biology.

Acknowledgements

We extend our deepest thanks to the following individuals, without whom the Summer Undergraduate Research Village would not have been possible.

Program Staff

Gregory Llacer, *Director of Harvard College PRISE; Director of the Office of Undergraduate Research and Fellowships (URAF)*

Jennifer Shephard, *Coordinator of BLISS; Special Initiatives Program Manager in the Office of the FAS Dean for Social Sciences*

Marais Young, *Coordinator of PRIMO; Programs Operations Manager in Harvard Business School*

Jeffrey Berg, *Coordinator of SHARP; Assistant Director of the Office of URAF*

Christy Colburn, *Coordinator of SURGH; Assistant Director of the Global Health & Health Policy Undergraduate Program*

Emily Maguire, *Coordinator of SURGH; Assistant Director of the Global Health & Health Policy Undergraduate Program*

Jozef Sulik, *Coordinator of Undergraduate Research in the Office of URAF*

Cammi Valdez, *Assistant Director of the Office of URAF*

Program Assistant Fellows

Justin Dower, *Lead Program Assistant Fellow*
Joel Bateman
Jennifer Chiang
Vimal Konduri
Hueyjong Shih
Madeleine Snyder
Kaan Yay

House Proctors

Luke Pizzato, *Lead Proctor*
Kaitavjeet Chowdhary
Madeline Cooper
Niamh Durfee
Michelle Guo
Jessica Hermann, *Emmanuel Proctor*
Gurbani Kaur, *Proctor/Administrative Fellow*
Josh Stallings

PRISE Faculty Advisory Board

Scott Edwards, *Professor of Organismic and Evolutionary Biology*

Howard Georgi, *Mallinckrodt Professor of Physics*

Evelynn Hammonds, *Chair, Barbara Gutmann Rosenkrantz Professor of the History of Science; Professor of African and African American Studies*

Jay Harris, *Dean of Undergraduate Education; Harry Austryn Wolfson Professor of Jewish Studies*

Georgene Herschbach, *ex officio, Former Dean of Administration for Harvard College*

Gregory Llacer, *ex officio, Director of Harvard College PRISE*

Robert Lue, *Professor of the Practice of Molecular and Cellular Biology; Director of Life Sciences; Director of HarvardX; Director of the Bok Center for Teaching and Learning*

Joan Reede, *Associate Professor of Medicine; Dean for Diversity and Community Partnership, Harvard Medical School*

John H. Shaw, *Dudley Professor of Structural and Economic Geology*

Dona Lee Wong, *Associate Professor of Psychiatry, Harvard Medical School*

House Staff

Paul Hegarty, *Leverett House Building Manager*

Dick Nerden, *Quincy House Building Manager*

List of Fellows

BLISS Fellows

Charlotte Anrig
Anthony Chen
Irene Limb
Krysianna Papadakis
Hansenard Piou
Sebastian Reyes
Mara Roth
Chace Shaw
Quinn Sluzenski
William Strang
Harrison Tanzola
Susan Wang
Liana Yamin
Eun Jin Yu
Anna Zannetos

PRIMO Fellows

Osama Arif
Caroline Brettler
Holly Breuer
Antonia Chan
Jiafeng Chen
Benjamin Cohen
Brian Cornyn
Kevon Edmondson
Lauren Nicholson
Daniel Nightingale
Christina Qiu
Alexander Rohe
Angela Yang
Hellary Zhang
Grant Zhao
Rohan Reddy
Steven Tan
Yilin Chen, *Wellesley Fellow*
Christina Chiu, *Wellesley Fellow*
Lingjun Xia, *Wellesley Fellow*

SHARP Fellows

Lev Asimow
Tez Clark
Jess Clay
Edith Enright Claire
Robert Hopkirk
Sarah Judd
Brieanna Martin
Thomas Peterson

Caleb Shelburne
Henry Shreffler
Aaron Suduiko
Zachary Wong

SURGH Fellows

Gita Bhattacharya
Kristen Fang
Kristy Hong
Henna Hundal
Isaac Hung
Laura Kanji
Michael Kennedy-Yoon
Julianna Lee
Sarah Rahman
Patrick Tan
Jonathan You
Jenna Zhang

PRISE Fellows

Yamen Abbas
Matthew Aguirre
Sahar Ashrafzadeh
Naomi Asimow
John Austin
Jennifer Bi
Alan Bidart
Julian Braxton
Erica Budina
Christopher Cantrell
Gerardo Castillo
Pooja Chandrashekar
Gui Zhen Chen
Anne Cheng
Bonirath Chhay
Mara Coyan
Miruna Cristuş
Jett Crowdis
Thomas Culp
Bluyé DeMessie
Carl Denton
Shaan Desai
Natasha Dhamankar
Christian Dolliff
Mohamed Ebied
Audrey Effenberger
Deanna Emery
Feyisayo Eweje
Shenyece Ferguson

Julio Fierro
Jaden Freeze
Felipe Flores
Andreé Franco Vasquez
William Fried
Adam Frim
Edgar Garcia
Michael Giles
Silvia Golumbeanu
David Gonzalez-Dysinger
Madeline Granovetter
Isobel Green
Elgin Korkmazhan Gulpinar
Emma He
James Hotchkiss
Jennifer Hu
Myles Ingram
Alexander H. Jin
Jessica Kim
Mingu Kim
Youbin Kim
Katie Kixmoeller
Peter Kraft
Deepika Kurup
Hannah Larson
Charles Law
Brandon Lee
Matthew Leifer
Valia Leifer
Catherine Li
Christopher Li
Jason Li
Victoria Lin
Gaia Linfield
Michael Liu
Eliza Llewellyn
Helene Lovett
Cynthia Luo
Bahlakoana Mabetha
Kushanth Mallikarjun
Leah Marsh
Ian Maynor
Soumyaa Mazumder
Andrew Mazzanti
Joshua Meier
Keegan Mendez
Adriana Mendez Leal
Jacob Meyerson
Zhuang Miao
Jade Moon
Alexander Munoz
Lucy Nam
Lin Ni
Nicolas Ontiveros
Muhammed Ors
Rachel Oshiro
Artidoro Roberto Andrea Pagnoni

Alex Pai
Margaret Panetta
Handong Park
Krystal Phu
Kateryna Pistunova
George Qiao
Linda Qin
Sai Shanthanand Rajagopal
Leah Rosen
Andre Sanchez
Suproteem Sarkar
Jacob Scherba
Yong Shen
Benjamin Sorscher
Lauren Sweetland-Martin
Ramtin Talebi
Mai-Linh Ton
Lillian Tsai
Kristin Tsuo
Daniel Um
Santiago Vargas
Julia Versel
Kruti Vora
Enxin Wan
Brandon Wright
Scott Xiao
Siddharth Yarlagadda
Kimberley Yu
Christina Zeina
Ellen Zhang
Christine Zheng
Joshua Breedon, *Emmanuel Fellow*
Ashavari Ghose, *Emmanuel Fellow*
Alvar Paris, *Emmanuel Fellow*
Oliver Philcox, *Emmanuel Fellow*
Luke Smith, *Emmanuel Fellow*