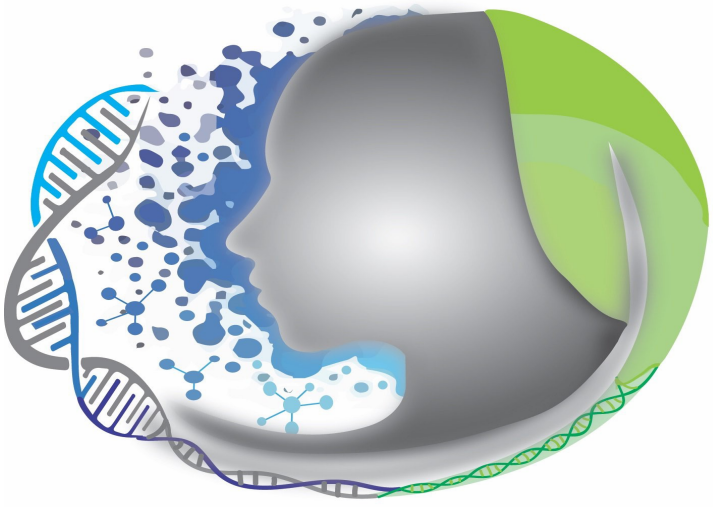
29th Annual Keck Research Conference



Precision Environmental Health

Integrating Environmental Sciences, Genomics, and Data Science to Advance Human Health

October 11, 2019

BioScience Research Collaborative
6500 Main St.
Houston, Texas

Conference Chair: Cheryl Walker, Baylor College of Medicine
Co-chair: Craig Hanis, UT Health Science Center

Sponsored by Keck Center Training Programs and Institutions of the:

Gulf Coast Consortia
QUANTITATIVE BIOMEDICAL SCIENCES

RICE Unconventional Wisdom
Baylor College of Medicine
UTHealth The University of Texas Health Science Center at Houston
The University of Texas MD Anderson Cancer Center Making Cancer History®
utmb Health
Institute of Biosciences and Technology Texas A&M Health Science Center transforming HEALTH

Keck Center Executive Committee

Chair:

Vasanthi Jayaraman, PhD
UT Health Science Center at Houston

Chairs Emeriti:

Kathleen Matthews, PhD
Rice University

Lydia Kavraki, PhD*
Rice University

B. Montgomery Pettitt, PhD*
UT Medical Branch at Galveston

Members:

Cesar Arias, MD, PhD*
UT Health Science Center at Houston

James Briggs, PhD
University of Houston

Pratip Bhattacharya, PhD
MD Anderson Cancer Center

Xiaodong Cheng, PhD
MD Anderson Cancer Center

Peter Davies, MD, PhD
Institute of Biosciences and Technology

Carmen Dessauer, PhD*
UT Health Science Center at Houston

Marc Morais, PhD
UT Medical Branch at Galveston

George Phillips, PhD
Rice University

Cheryl Walker, PhD*
Baylor College of Medicine

Ted Wensel, PhD*
Baylor College of Medicine

Richard Willson, PhD
University of Houston

Lynn Zechiedrich, PhD
Baylor College of Medicine

***Training Program Director**

GCC Oversight Committee Chair:

David Niesel, PhD
UT Medical Branch at Galveston

GCC/Keck Center Staff:

Karen Ethun
Executive Director

Melissa Glueck
Director, Keck Center

Dawn Koob
Administrative Coordinator

Carrie Cerda
Financial Administrator

Vanessa Herrera
Assistant Director, Keck Center

Suzanne Tomlinson
Director, Research Programs

The Keck Center and Gulf Coast Consortia for Quantitative Biomedical Sciences

The Keck Center

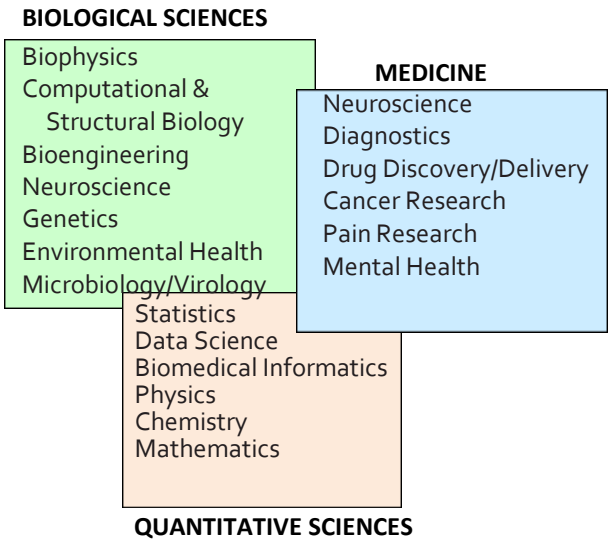
The Keck Center, established in 1990 with support from the W. M. Keck Foundation, celebrates its 29th year of supporting predoctoral and postdoctoral trainees and their mentors. From the founding institutions, Baylor College of Medicine and Rice University, the Keck Center grew in its first 10 years to six major public and private institutions in the Houston/Galveston area, including University of Houston, The University of Texas Health Science Center at Houston, The University of Texas Medical Branch at Galveston, and The University of Texas MD Anderson Cancer Center. The Institute of Biosciences and Technology of Texas A&M Health Science Center joined in 2015. Guiding the formation of this collaboration was the realization that significant advances in the biological sciences, such as the DNA sequencing of the human genome, would be driven by the integration of biology and computer science. The partners realized, however, that most biological scientists were not prepared to capitalize on novel approaches to visualization, analysis and interpretation of experimental data made possible by rapid advances in computing technology. Moreover, most researchers in computer programming and analysis systems did not have adequate knowledge about biology and biological systems. The Keck Center was explicitly designed to bridge this gap between biological and computational sciences by fostering collaborations among scientists through specially designed research and training programs.

Building on its expertise in interdisciplinary, inter-institutional programs, the Keck Center's focus has evolved to the quantitative biomedical sciences. Participants are drawn from various disciplines such as biophysics, chemistry, bioengineering, neuroscience, computer science, biochemistry, genetics, physics, mathematics, data science, biomedical informatics, environmental health, biology and statistics. Currently, the Keck Center administers training programs in biomedical informatics and data science, molecular biophysics, pharmacological sciences, computational cancer biology, precision environmental health and antimicrobial resistance.

Gulf Coast Consortia

In March 2001, the presidents of each of the six member institutions of the Keck Center signed an unprecedented agreement establishing the Gulf Coast Consortia (GCC), explicitly designed to coalesce institutional strengths in order to:

1. **train new scientists at the intersection of biological sciences with quantitative and physical sciences**
2. **build cutting-edge research infrastructure and facilities**
3. **cultivate a supportive atmosphere for the collaboration of basic and translational scientists, researchers, clinicians and students in both biological and non-biological fields**
4. **apply the resulting knowledge to prevent and treat diseases**



While the Keck Center serves as the training arm of the GCC, the research arm consists of individual, topic-focused research, including translational pain research, antimicrobial resistance, cellular and molecular biophysics, regenerative medicine, drug discovery and development, mental health research, nano x, single cell omics, and translational imaging . These consortia and newly forming clusters provide a supportive environment for the encouragement and development of research that might otherwise be beyond the reach of any one institution. New consortia form when faculty come together around a common interest, establishing a working vision and engaging a broad faculty community to pursue interinstitutional research, present conferences, acquire shared equipment and research cores and/or develop training, research or curriculum grants. http://www.gulfcoastconsortia.org/

Acknowledgements

The Keck Center thanks the following for their generous support:

National Library of Medicine (NLM)

NLM Training Program in Biomedical Informatics and Data Science T15 LM007093

Principal Investigator: Lydia Kavraki, PhD, Rice University

National Institute of General Medical Sciences (NIGMS)

Training Interdisciplinary Pharmacology Scientists T32 GM120011

Principal Investigator: Carmen Dessauer, PhD, UT Health Science Center at Houston

Houston Area Molecular Biophysics Program T32 GM008280

Principal Investigator: Ted Wensel, PhD, Baylor College of Medicine

National Institute of Environmental Health Sciences (NIEHS)

Training in Precision Environmental Health Sciences T32 ES027801

Principal Investigator: Cheryl Lyn Walker, PhD, Baylor College of Medicine

National Institute of Allergy and Infectious Diseases (NIAID)

Texas Medical Center Training Program in Antimicrobial Resistance T32 AI141349

Principal Investigator: Cesar Arias, MD, PhD, UT Health Science Center at Houston

Cancer Prevention and Research Institute of Texas (CPRIT)

Computational Cancer Biology Training Program RP170593

Principal Investigator: B. Montgomery Pettitt, PhD, UT Medical Branch at Galveston

Gulf Coast Consortia Member Institutions

Baylor College of Medicine
GIVING LIFE TO POSSIBLE

RICE

UNIVERSITY of HOUSTON

UTHealth
The University of Texas Health Science Center at Houston

THE UNIVERSITY OF TEXAS MD Anderson Cancer Center
Making Cancer History®

utmb Health
Working together to work wonders.™

INSTITUTE OF BIOSCIENCES AND TECHNOLOGY
TEXAS A&M HEALTH SCIENCE CENTER

2019 Logo Design Information

Precision Environmental Health

Integrating Environmental Sciences, Genomics, and Data Science to Advance Human Health

Logo Design by:

Mindy Engevik, PhD

Postdoctoral Fellow, Pathology
Baylor College of Medicine

Legend:

The logo is designed to symbolize personalized medicine, which encompasses genetics, environmental factors, pharmaceutical agents and the human host.

AGENDA

Oct. 11, 2019

7:50 – 8:40 Registration and Light Breakfast

8:50 – 9:00 Welcome and Introduction – **Cheryl Walker, PhD,** Professor and Chair, Center for Precision Environmental Health, Baylor College of Medicine

9:00 – 9:55 **James Eberwine, PhD**, Elmer Holmes Bobst Professor and Chair, Pharmacology, University of Pennsylvania
Single Cell Multiomics Informs Emergent Cell Biologies: From Open-Chromatin Analysis to Multigenic Functional Genomics

9:55 – 10:50 **Elizabeth Winzeler, PhD**, Professor, Division of Pharmacology and Drug Discovery, Department of Pediatrics, and Director of Translational Research, UC San Diego
The Search for Next Generation Medicine for Malaria Elimination

10:50 – 10:55 Trainee speaker: **Pavan Kota**, Bioengineering, Rice University. NLM Training Program in Biomedical Informatics and Data Science

10:55 – 11:00 Trainee speaker: **Miriam Gavriliuc**, Biology and Biochemistry, University of Houston. Houston Area Molecular Biophysics Training Program

11:00 – 11:05 Trainee speaker: **Jourdan Andersson, PhD**, Pathology, Baylor College of Medicine. Antimicrobial Resistance Training Program

11:05 – 12:05 Poster Session (odd numbered) and Networking

12:00 – 12:40 Lunch and Poster Viewing

12:40 – 12:45 Reconvene afternoon session – **Craig Hanis, PhD**, Professor, School of Public Health, UTHealth

12:45 – 1:40 **Andrea Baccarelli, MD, PhD,** Leon Hess Professor and Chair, Environmental Health Sciences, and Director, Laboratory of Precision Environmental Biosciences, Columbia University
Epigenomics and Precision Environmental Health in Human Studies

1:40 – 1:45 Trainee speaker: **Eric Smith**, Molecular and Cell Biology, Baylor College of Medicine. Precision Environmental Health Training Program

1:45 – 1:50 Trainee speaker: **Brittany Jewell**, Graduate Program in Biochemistry and Cell Biology. UTHealth. Training Interdisciplinary Pharmacology Scientists Program

1:50 – 2:50 Poster Session (even numbered) and Networking

2:50 – 2:55 Reconvene – **Craig Hanis, PhD**

2:55 – 3:50 **Richard Finnell, PhD, DABMGG,** Professor, Center for Precision Environmental Health, Baylor College of Medicine
Genomic Architecture of Neural Tube Defects: Implications for Precision Medicine?

3:50 - 3:55 Trainee speaker: **Madeline Burns**, Department of BioSciences, Rice University. NeuroEngineering: From Cells to Systems (NSF IGERT Program)

3:55 – 4:00 Trainee speaker: **David Shih, PhD**, Systems Biology, UT MD Anderson Cancer Center. Computational Cancer Biology Training Program

4:00 – 4:55 **Debra Laskin, PhD,** Distinguished Professor and Chair, Pharmacology and Toxicology, Rutgers University
Inflammatory Macrophages: Agents of Defense or Destruction in the Pathogenesis of Lung Disease?

4:55 – 5:10 Awards and Closing – **Cheryl Walker, PhD** and **Craig Hanis, PhD**

5:10 – 6:00 Reception

Speaker Abstracts
(In order of appearance)

Single Cell Multiomics Informs Emergent Cell Biologies: From Open-Chromatin Analysis to Multigenic Functional Genomics

James Eberwine, PhD
Elmer Holmes Bobst Professor of Pharmacology
Department of Systems Pharmacology and Translational Therapeutics
Co-Director of the Penn Program in Single Cell Biology
Perelman School of Medicine
University of Pennsylvania
Website Link: https://www.med.upenn.edu/apps/faculty/index.php/g275/p5441

The research efforts of Dr. Eberwine's lab are directed toward understanding the molecular basis of neuronal functioning. His experimental approach is reductionist in nature and involves analysis of gene expression in individual cells dispersed in culture, in the live slice preparation or from fixed pathological tissue specimens. His lab is currently generating molecular and bioprocess fingerprints of various cell types and disease states. When complete, the hope is that it will be possible to alter the cellular response to various challenges by altering the levels of these biological processes in a predictable manner. Dr. Eberwine is an NIH Pioneer Award recipient and has also received a McKnight Foundation Technology Innovations Award, NIH MERIT Award and Ellison Foundation Senior Scholar Award. He received his BS from Yale and his PhD from Columbia. He did postdoctoral work in molecular neurobiology at Stanford University.

Abstract: Recent advances in single cell biology have focused attention on whole cell molecular variation, such as transcriptome differences between seemingly identical cells. The functional significance of observed molecular variation at the level of single cells is unclear, however it is curious that physiological differences between cells seem agnostic to many of these molecular differences. In this context cells are not just simple building blocks of higher-level complex structures. Cells possess complex structures, which perform independent and interdependent processes that integrate to elaborate the resultant complex functions of the cell. In this presentation data from my lab, detailing how the internal genomics of a cell contribute to the characteristics of said cell and help to define its "phenotype" will be described. In particular novel assays have been developed to perform robust single cell chromatin analysis, to assess RNA granule dynamics, to quantify single mitochondrial heteroplasmy and to investigate the multigenic nature of single cell functional genomics. These studies have been undertaken to discern how these, and other subcellular structures, work together to elicit a - "more than the sum of its parts" – biology, which in the context of cellular variability, requires selective cellular and subcellular analysis of these structures. It is anticipated that the biological insights from these and other studies will contribute to the development of a dynamical "Theory of Cell Type", similar to the historical development of the "Theory of Biological Species".

The Search for Next Generation Medicine for Malaria Elimination

Elizabeth Winzeler, PhD
Professor of Pharmacology and Drug Discovery, Department of Pediatrics
University of California San Diego
Website Link: https://winzeler.ucsd.edu/winzeler/

Dr. Winzeler heads the Laboratory for Eukaryotic Pathogenesis, Drug Discovery and Chemical Biology. Her group uses systematic, data intensive methods to solve problems at the interface of host pathogen biology typically involving large collections of chemical screening data and whole genome sequencing. A focus of her lab is to conduct phenotypic screens and compound testing to accelerate the search for new malaria treatments and malaria prevention methods. She is also interested in host pathogen signaling and how malaria parasites interact with human liver cells. Dr. Winzeler is director of one of 13 scientific labs in The Malaria Drug Accelerator (MalDA), funded by the Melinda and Bill Gates Foundation to advance the development of new antimalarial drugs. In 2017 she received the Alice and CC Wang Award in Molecular Parasitology from the American Society for Biochemistry and Molecular Biology for her contributions to antimalarial drug discovery. Dr. Winzeler received her PhD in Developmental Biology from Stanford.

Abstract: Although phenotypic cellular screening can be used to drive antimalarial drug discovery, the process can be facilitated and improved when a compound's cellular target is known. This is especially true when appropriate high-throughput cellular assays are either lacking or are prohibitively expensive. The Malaria Drug Accelerator (MalDA) is applying a consortium-based approach to coordinately identify chemically validated targets for next generation medicines. Compounds with potent antimalarial activity are first identified using cellular assays that predict activity against different malaria parasite lifecycle stages (liver stage, asexual blood stage, sexual stage) and characterized by metabolomic and genetic profiling against known targets. We then use in vitro evolution methods to isolate compound-resistant parasites whose genomes may be analyzed using whole genome sequencing. Resistance alleles are validated using Crispr-Cas9 genome editing. We next establish whether the conditional knockdown of a putative target results in increased sensitivity to a compound. Validated targets are progressed to assay development, protein overexpression experiments, and structure determination. Altogether over 200 compounds have been evaluated and the genomes of over 400 isogenic drug resistant parasites have been created and sequenced. We have discovered a variety of new druggable targets, some of which have advanced to further assay development. The work is also defining the Plasmodium resistome—the complete set of alleles that offer resistance to xenobiotic compounds as well as compounds which resist resistance. These combined joined efforts of the MalDA consortium are laying the foundation for future efforts to create better medicines.

Epigenomics and Precision Environmental Health in Human Studies

Andrea Baccarelli, MD, PhD
Chair and Leon Hess Professor of Environmental Health Sciences
Director of the Laboratory of Precision Environmental Biosciences
Mailman School of Public Health
Columbia University
https://www.mailman.columbia.edu/people/our-faculty/ab4303#field_biography

Dr. Baccarelli's research explores epigenetic and molecular mechanisms as potential functional pathways linking exposures to environmental pollutants to human disease. His laboratory research activities are specifically focused on epigenetics, mitochondriomics, and computational epigenomics.
He is currently the PI of multiple NIH-funded projects, and since 2010, his lab has produced more than 230 publications at the intersection of epigenetics, molecular epidemiology and environmental health. His paper "Effects of Particulate Matter on Genomic DNA Methylation Content and iNOS Promoter Methylation" was named 2013 Classic Paper of the Year by Environmental Health Perspectives, the leading journal in environmental health research. The work of Dr. Baccarelli's and his lab has had significant influence on human environmental health research due to novel laboratory and statistical approaches. He is consistently recognized for his contributions to the field of environmental epigenetics.
Dr. Baccarelli received his PhD from the University of Milan and his MD from the University of Perugia, both in Italy.

Abstract: According to the World Health Organization, as much as 24% of human disease globally is caused by environmental exposures that could be averted. While this shows the huge potential benefit of disease prevention, the current model for precision medicine is mostly focused on developing targeted treatment to patients, i.e., once the disease has already developed. The concept of Precision Environmental Health posits that emerging approaches can help in the identification of individuals at risk due to environmental exposures and thus prevent disease. Today, based on an individual's exposure profiles and underlying susceptibility to disease, we can use advanced exposomic, biological, and computational methods to research and develop a prevention plan tailored precisely to that individual, rather than our current one-size-fits-all approach. For instance, while the use of big data to study environmental toxics is still an emerging research area it is growing alongside many other fields. As datasets have increased in size the ability of researchers to find relevant structure within such datasets has increased. This has been explored extensively within molecular data cohorts and is now a commonly used technique, e.g. for determining blood cell type mixtures using DNA methylation data. Less explored is the role that clustering algorithms may have in uncovering structure within environmental exposures particularly as it relates to molecular profiles and disease risk. Despite the increasing availability of data, most studies of environmental toxics have not extended to incorporate multiple correlated exposures, and the potential of molecular signatures to reflect biological risk from exposure to toxins, or signatures of the toxins themselves can benefit from similar techniques developed by compatriots in nearby fields. As large scale cohorts become more widely available and molecular data becomes more common, the study of environmental toxins and adverse health outcomes will only benefit from increasingly incorporating the tools and techniques common in other fields of big data analyses. Leveraging the resources of the wider scientific field, and placing them in context of the quetions and approaches unique to toxicology and environmental epidemiology is likely to be at the forefront of the rapid expansion into precise approaches to prevention.

Genomic Architecture of Neural Tube Defects: Implications for Precision Medicine?

Richard Finnell, PhD, DABMGG
Professor, Center for Precision Environmental Health
Baylor College of Medicine
Website Link: https://www.bcm.edu/people/view/richard-finnell-ph-d-dabmgg/9aa61609-3f9f-11e7-916f-005056a012ee

Dr. Finnell is a pediatric geneticist. He has long been involved in investigating genetic susceptibility to environmentally induced birth defects, applying stem cell technology to the detection of potential teratogenic compounds in efforts to prevent these birth defects, developing mouse models to understand the pathogenesis of the malformations, and using highly innovative approaches to treating these disabilities. During his 39+-year career, he has authored over 330 peer-reviewed publications in journals such as Science, Nature Genetics, Nature Cell Biology, PNAS and Developmental Cell. His early work with murine embryonic stem cells helped establish the dire embryonic consequences of folate deficiency during embryonic development. The Finnell laboratory is focused on how folic acid transport may be a target of selected human teratogens such as the anti-retroviral drug dolutegravir, or in other instances modifies the impact of teratogenic agents on embryonic development. This work takes advantage of his clinical training in medical genetics, as well as a background grounded in developmental and molecular biology and teratology. This lab has been fortunate to receive continuous NIH funding for decades to support innovative research on birth defect prevention.

Abstract: The presentation will highlight the promise of precision medicine that is informed by genomic analysis coupled with high level genetic counseling in reducing the occurrence of complex congenital defects. Specifically, for the first half of the talk, I will focus on the management of the pregnant epileptic mother, or women who receive anti-epileptic medication for other clinical indications. Medications such as Depakote (VPA; Valproic Acid), have been the most commonly prescribed AED globally, despite having the most significant teratogenic potential amongst the clinically available AEDs. In the Neurodevelopmental Effects of Antiepileptic Drugs (NEAD) study, 20% of VPA exposed children showed major congenital malformations (MCMs) or resulted in fetal death (Meador, et al., Neurology 2006). Yet like most teratogens, VPA and other anti-epileptic medications only compromise development in a small percentage of exposed infants. These represent challenging genetic counseling situations that could benefit from an infusion of genomic technology. In an effort to better manage AED-complicated pregnancies in order to prevent preventable birth defects, we are attempting to define a genetic signature of risk for mother-infant pairs exposed to VPA during pregnancy. Based on our recently published study (Chen et al., Cell Research 2018) that builds on the Omnigenic Model of Inheritance whereby susceptibility to various health conditions is linked not to just a few "core" genes, but rather involves almost all expressed genes throughout the genome, it is possible that NTD risk associated with VPA exposure can be quantified and explained by a defined genomic signature. The second half of the presentation will consider the role of the environment in compromising pregnancy outcomes and how identification and application of this approach to high risk mothers can one day prevent preventable birth defects.

Inflammatory Macrophages: Agents of Defense or Destruction in the Pathogenesis of Lung Disease?

Debra Laskin, PhD
Distinguished Professor and Chair of Pharmacology and Toxicology
Ernest Mario School of Pharmacy
Rutgers University
Website Link: https://pharmacy.rutgers.edu/directory/laskin-debra/

Dr. Laskin has a broad background and training in pharmacology and toxicology, with specific expertise in immunology, pulmonary physiology, lung pathology and tissue injury. Research in her laboratory is focused on elucidating inflammatory mechanisms underlying disease pathogenesis. She collaborates with a number of investigators working in different animal model systems including the lung, liver and skin. A major emphasis of her work is the analysis of lung toxicity induced by inhalation of environmental pollutants such as ozone and particulate matter, and vesicants including sulfur mustard and nitrogen mustard. In 2014 she received the Society of Toxicology Education Award and the Women in Toxicology Mentoring Award. Dr. Laskin has more than 235 publications and has written numerous influential invited reviews on macrophages and tissue injury, which are widely referenced. She received her PhD in Pharmacology/Toxicology from the Medical College of Virginia, Virginia Commonwealth University.

Abstract: A diverse group of toxicants including air pollutants and chemical warfare agents have been identified that cause acute injury to the lung that progresses to chronic disease. The pathologic response to these toxicants depends on the dose and duration of exposure and their physical/chemical properties. A common response to pulmonary toxicants is an accumulation of proinflammatory/cytotoxic M1 macrophages at sites of injury, followed by the appearance of anti-inflammatory/wound repair M2 macrophages. It is thought that the outcome of the pathogenic responses to xenobiotics depends on the balance in the activity of these macrophage subpopulations. Overactivation of either M1 or M2 macrophages can lead to injury and disease. Thus, the very same macrophage-derived mediators, released in controlled amounts to destroy injurious materials and pathogens (e.g., reactive oxygen species, reactive nitrogen species, proteases, tumor necrosis factor-α) and initiate wound repair (e.g., transforming growth factor-β, connective tissue growth factor, vascular endothelial growth factor), can exacerbate acute lung injury and/or induce chronic disease such as fibrosis, chronic obstructive pulmonary disease, and asthma, when released in excess. Understanding how lung diseases develop following toxicant exposure, and the contribution of inflammatory macrophages and mediators to the pathogenic response may lead to the development of novel approaches for treating lung diseases.

Trainee Speaker Abstracts
(In order of appearance)

Universal Pathogen Diagnostics with Nonspecific DNA Sensors and Compressed Sensing

Pavan Kota
PhD Student
Bioengineering
Rice University
NLM Training Program in Biomedical Informatics and Data Science Program

Pavan is a third year PhD student in the Department of Bioengineering at Rice University. With prior experience in both research labs and industry, he has a passion for projects with the potential for clinical translation. He is primarily interested in biosensing for rapid diagnostics and how these technologies can benefit from techniques in machine learning and signal processing. Pavan conducts research at this interface through a collaboration between his co-advisors, Dr. Richard Baraniuk and Dr. Rebekah Drezek.

Abstract: Rapid infection diagnostics enable clinicians to prescribe lifesaving therapies for critically ill patients. However, current technologies fail to provide timely, cost-efficient quantification of dozens of possible pathogens. Our interdisciplinary team is developing a novel technique to scalably quantify large panels of pathogens with only a few nonspecific DNA sensors that bind at multiple locations along whole microbial genomes. The number of binding events for each sensor provides each pathogen with a fingerprint ID number. We then apply compressed sensing, a signal processing technique that lets us easily unmix fingerprints if a few pathogens are present in the same sample.
Our sensing mechanism readily keeps pace with constantly evolving pathogens, and its simplicity motivates our ongoing research towards portable formats and clinical translation.

Measuring the Mechanical Forces During Protein Biosynthesis via EF-G Crosslinking

Miriam Gavriliuc
PhD Student
Biology and Biochemistry
University of Houston
Houston Area Molecular Biophysics Training Program

Miriam is currently a third year Biochemistry PhD student at the University of Houston. Her research is focused on studying the mechanism of Elongation Factor G mediated translocation, as well as understanding how the ribosome can quickly and accurately translate mRNA into protein.

Abstract: The ribosome is the complex molecular machine found in all living cells that is responsible for the synthesis of all proteins. The ribosome is associated with various protein factors, including the GTPase Elongation Factor G (EF-G). EF-G catalyzes tRNA and mRNA translocation on the ribosome, however, the

mechanism of this translocation remains elusive. In this talk, I will present my current work studying this mechanism and how it relates to the ribosome's ability to efficiently translate mRNA into protein. I have generated double-cysteine EF-G that is internally crosslinked with various lengths of crosslinkers, and will measure force generated by EF-G. By restricting the movement of EF-G with various crosslinkers, the force generated can be measured at different stages of translocation, giving new insight into this mechanism.

Utilizing Immunomodulating Drugs to Combat Infections

Jourdan Andersson, PhD
Postdoctoral Fellow
Pathology
Baylor College of Medicine
Antimicrobial Resistance Training Program

Dr. Jourdan Andersson is a Postdoctoral Fellow in the lab of Dr. Tor Savidge. She received her B.S in Biochemistry and Biophysics from Rensselaer Polytechnic Institute and her PhD in Human Pathophysiology and Translational Medicine from the University of Texas Medical. Her current research is focused on identifying host-targeted therapeutics to combat antibiotic resistant pathogens, particularly members of the Enterobacteriaceae family and Clostridioides difficile.

Abstract: Antibiotic resistant pathogens represent one of the most pressing public health concerns of the 21st century. Previously, in order to identify alternative therapeutics to combat pathogens, I utilized a drug repurposing screening approach and identified three non-antibiotic drugs that were broadly protective against several pathogens. Interestingly, none of the drugs were observed to have direct bactericidal/bacteriostatic effects against pathogens of interest, leading to the hypothesis that these drugs are host-directed. In this talk, I will present recent work to identify the mechanism(s) of action of these drugs and elucidate novel protective immune pathways that could potentially be targeted for the development of immunomodulating therapeutics to combat pathogens.

Investigating Lineage-specific Arsenic-induced Epigenetic Alterations during Embryogenesis

Eric Smith
PhD Student
Molecular and Cell Biology
Baylor College of Medicine
Precision Environmental Health Training Program

A native of Western New York, Eric received his B.S. in Biochemistry and Biophysics from Rensselaer Polytechnic Institute in Upstate New York. Prior to attending Baylor College of Medicine, he received a decade of training in infectious disease as a staff scientist under the mentorship of Dr. Kathleen McDonough at Wadsworth Center within the New York State Dept. of Health. Currently, Eric is a PhD student in the laboratory of Dr. Courtney Hodges in the Center for Precision Environmental Health combining in vitro 3D models and single-cell technologies to build a more detailed understanding of mechanisms underlying disease.

Abstract: Millions of people worldwide are exposed to naturally occurring arsenic through ingestion of contaminated food and water. Arsenic exposure has a broad impact, disrupting several body systems, including normal developmental processes. Furthermore, early-life arsenic exposure is associated with predisposition to cancer. To make direct connections between arsenic-induced epigenetic changes and adverse health outcomes, we will combine single-cell approaches following the addition of arsenic to embryoid bodies, an in vitro 3D model of embryogenesis. Single-cell RNA-seq and ATAC-seq approaches will map out cell-type specific effects, as well as their reversibility following therapy. This intersection of a complex 3D model of embryogenesis and single-cell approaches will allow the determination of how early environmental arsenic exposure affects development at a cutting-edge level of detail.

Bone Tumors in a Dish: Modeling Rothmund-Thomson Syndrome Associated Osteosarcomagenesis

Brittany Jewell
PhD Student
Biochemistry and Cell Biology
UT Health Science Center Houston
Training Interdisciplinary Pharmacology Scientists Pogram

Brittany is a fourth year PhD student in the lab of Dr. Dung-Fang Lee (Integrative Biology and Pharmacology) where she studies rare genetic disorders that lead to childhood bone cancer. She is a fellow of the CPRIT Innovation for Cancer Research and Prevention Program and the 2019 Charlene Kopchick Fellow.

Abstract: Osteosarcoma (OS) is the most common bone malignancy in children. Patients with Type II RTS have biallelic mutations in the DNA helicase RECQL4. This is the most penetrant cancer-causing mutation among any cancer (30%). In this talk, I will discuss our new model of osteosarcomagenesis, where we reprogrammed paired patient and parental (carrier) fibroblasts to induced pluripotent stem cells (iPSCs). We further demonstrate that these iPSCs are capable of differentiation to mesenchymal stem cells (MSCs), and then to osteoblasts, the potential cell-of-origin of OS. Because the iPSC-derived osteoblasts retain the genetic fidelity of the patient from which they were derived, we believe these cells are the ideal model to study RTS-associated OS pathophysiology. In addition, these collections of cells provide a novel model with potential for use in drug testing and personalized therapy.

Using Evolution to Address Neuroscience Questions: What Can Specialization on a Toxic Fruit Tell Us About Decision-making and Learning?

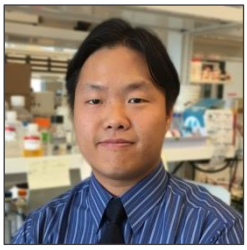
Madeline Burns
PhD Student
BioSciences
Rice University
Neuroengineering: From Cell to Systems (NSF IGERT Program)

Madeline graduated with a BS in Biology and minor in Journalism Studies from Texas A&M in 2012, then worked as a consultant in the oil and gas industry before beginning her PhD in the lab of Julia Saltz. Her

research interests are focused on understanding variation in cognition, particularly how cognitive bias and variation in environmental experience influence the behavioral outcomes of cognitive processes.

Abstract: We observe variation in decision-making and learning all the time – but how is this variation being produced? Individuals may vary in general cognitive ability but can also vary in a multitude of other ways - such as personality, perception, and environment inhabited. These differences may result in certain cognitive biases - potentially influencing cognitive processes, and thus the behavioral outcomes of those processes. So how can we parse out what sources of variation are contributing to these outcomes? To investigate, we compare two species of Drosophila: D. simulans and D. sechellia. These species are closely-related, but differ significantly in habitat use and cognitive bias, allowing us to investigate the effect of both cognitive ability and cognitive bias on decision-making and learning outcomes.

A Backdoor Attack to Circumvent Adaptive Resistance of Breast Cancer to PARP Inhibitors



David Shih, PhD
Postdoctoral Fellow
Systems Biology
UT MD Anderson Cancer Center
Computational Cancer Biology Training Program

Following an undergraduate degree in pathobiology and graduate training in molecular biology at the University of Toronto, David Shih completed doctoral training with Dr. Michael Taylor at the Hospital for Sick Children on the genomics of pediatric brain tumor. Subsequently, he joined the Department of Biostatistics and Computational Biology at Dana-Farber Cancer Institute, where he developed statistical models to identify metastatic driver alterations of brain metastases from case-control studies. He is now working with Dr. Shiaw-Yih Lin (Systems Biology) and Dr. Kim-Ahn Do (Biostatistics) at M.D. Anderson Cancer Center to study cancer vulnerabilities that result from mutational processes in cancer, including deficiencies in DNA repair.

Abstract: PARP inhibitors (PARPi) recently emerged as an approved class of anti-cancer drugs for maintenance therapy of breast cancer. PARPi are particularly efficacious against tumors with deficiencies in the homologous DNA repair pathway, often due to mutations in the genes BRCA1 or BRCA2. However, some BRCA1/2 mutant tumors eventually develop resistance against PARPi. Here, we show that a very rare population of BRCA1 mutant cells can acquire a secondary mutation that restores BRCA1 function, even before PARPi treatment. This cancer cell population are thus PARPi resistant and selectively expand after long-term PARPi treatment; concurrently, the cells acquire global transcriptional changes that result in predicted and confirmed sensitivity to MEK inhibitors. Therefore, PARPi treatment can induce a secondary vulnerability to MEK inhibitors, suggesting that this sequential treatment may be an effective strategy against BRCA1/2 mutant cancers.

Poster Presenters

in alphabetical order

First	Last	Institution	Abtract Title	Poster #
Dayana	Abdullah-Smoot	TSU	*Urban Watersheds Seasonal Deleterious Effects on Colon Cells*	1
Jonas	Actor	Rice U.	*What Do Neural Networks Learn? A Mathematical Comparison of Convolution Kernels and Image Processing Features*	2
Dreycey	Albin	Rice U.	*SeqScreen: A Biocuration Platform for Robust Taxonomic and Biological Process Characterization of Nucleic Acid Sequences of Interest*	3
Jourdan	Andersson	BCM	*Use of Host-Targeted, Immunomodulating Drugs to Combat Clostridioides difficile Infection*	4
Diner	Antunes	Rice U.	*Towards Computational Prediction of Off-target Toxicity in Cancer Immunotherapy*	5
Jennifer	Asmussen	BCM	*Identifying Drug Response Genes Using Evolutionary Action and Machine Learning in Human Cell Line Panels*	6
Sekinat	Atobiloye	TSU	*Determination of Oxybenzone and Clofibrate in Water, Sediment, and Fish from Galveston Bay / Clear Creek Areas*	7
MacKinsey	Bach	UTHSCH	*Association of the Maternal Genome with Autism Spectrum Disorder: Use of Whole Genome Sequence Data for Genome-Wide, Gene-Based Analyses and Genetic Risk Scores*	8
Nathaniel	Berg	UTHSCH	*MicroRNA-147b in Myeloid Cells Promotes Resolution of Acute Lung Injury*	9
David	Boragine	BCM	*Directed Evolution of a Potent Inhibitor of the Methicillin-Resistant Staphylococcus aureus (MRSA) Resistance Enzyme, PBP2a*	10
Cecilia	Bores	UTMB	*dsDNA Packed inside Phage Capsids: Structure and Defects Emergence*	11
Ahmad	Borzou	UTMB	*Matrix Decomposition of Omics Data and Pathway Correlations*	12
Adam	Campos Acevedo	BCM	*Generating a Molecular Understanding of the Role of Co-factors in the p53 Response*	13

First	Last	Institution	Abtract Title	Poster #
Varuna	Chander	BCM	*Investigating Mosaicism in Blood for Cardiovascular Disease Risk*	14
Nick	Charron	Rice U.	*Spectral Coarse Graining: Deep Learning of Simple, Dynamically Consistent Protein Models*	15
Samantha	Decker	BCM	*Loss of miR-290 Modifies the Risk of Embryonic Neural Tube Defects in Maternal Diabetes*	16
Darius	Devlin	BCM	*FAM170A Loss Causes Subfertility and Defective Sperm Motility in Mice*	18
Alex	Ditzel	Rice U.	*Use of a Cell-free Protein Synthesis System and Coupled Enzyme Reactions to Synthesize Natural Products*	17
Kristen	Engevik	BCM	*Identifying the Role of Purinergic and Calcium Signaling in Rotavirus Infection*	19
Melinda	Engevik	BCM	*Serotonin Promotes Epithelial Restitution Through Goblet Cell Mediated Secretion of MUC2 and TFF3*	20
Xian	Fan	Rice U.	*Methods for Copy Number Aberration Detection from Single-cell DNA Sequencing Data*	21
Sari	Galvin	TSU	*Aftermath of Hurricane Harvey: Bacterial Analysis in Houston Area Bayous*	22
Evan	Gates	MDACC	*Estimating Cellular Density in Glioma Using MR Imaging Data*	23
Miriam	Gavriliuc	UH	*Measuring the Mechanical Forces During Protein Biosynthesis via EF-G Crosslinking*	24
Katelyn	Handley	MDACC	*Assessment of CRM-1 Inhibitor Combinations in Ovarian and Endometrial Cancer*	25
Willie	Hsu	UTMB	*Measurement Of T:G Frequency at Codon 248 of Exon 7 of the TP53 Gene Using Deep Sequencing in Human Cell Lines and Tissues*	26
Ramiz	Iqbal	MDACC	*Clinical Phenotyping Using Semantic Knowledge Graphs*	27
Brittany	Jewell	UTHSCH	*A Novel Model of Osteosarcomagenesis Reveals Dysregulation of Oxidative Phosphorylation*	28

First	Last	Institution	Abtract Title	Poster #
Dongya	Jia	Rice U.	*Quantifying Cancer Metabolic Plasticity by Coupling Gene Regulation with Metabolic Pathways*	29
Angelica	Jimenez	TSU	*Bioautographic and Molecular Docking Studies of Acetylcholinesterase Inhibitory Activity of Terpenes*	30
Djene	Keita	TSU	*Activation of MAPK Signaling Pathway in Human Lung and Colon Cells Exposed to Triclosan*	31
Rachel	Keuls	BCM	*Neural Crest Development Is Temporally Regulated By miR-302*	33
Rachel	Keuls	BCM	*miR-302 Regulates Neural Tube Closure Via Inhibition of Sonic Hedgehog Signaling*	34
Pavan	Kota	Rice U.	*Signal Amplification of Nonspecific DNA Hybridization Events for Compressed Sensing of Pathogens*	32
Jiasong	Li	HMRI	*3D Label-free, Real-time, Chemistry-Sensitive Imaging to Identify Parathyroid Adenoma and Hyperplasia and Classify Parathyroid Glands and Recurrent Laryngeal Nerve During Surgery*	35
Eleni	Litsa	Rice U.	*Applying Graph Convolutional Neural Networks for Drug Metabolism Prediction*	36
Mahsa	Lotfollahi	UH	*Integrating Spatial Features into FTIR Histological Image Classification*	37
Rupali	Mankar	UH	*Comprehensive Evaluation of Myelofibrosis in Bone Marrow Using Infrared Imaging*	38
Sarah	May	UTHSCH	*Identifying a Cohort of Patients Living with HIV from a Large Electronic Health Record Database*	39
Maurício	Mengatti Rigo	Universidade Catolica do Rio Grande do Sul	*Structural in Silico Analysis of Post-Translational Modifications Present on Dnak Protein*	41
Jonathan	Mercado	BCM	*Targeting the Leishmania Hsp100 N Domain to Prevent Infective Parasite Stage Differentiation*	40
Saeid	Parvandeh	BCM	*Evolutionary Action Through Machine Learning Methods to Discover Driver Genes in Cancer Somatic Mutations*	42

First	Last	Institution	Abtract Title	Poster #
Nabina	Paudyal	UTHSCH	*Study of a Heteromeric Kainate Receptor Gluk2/k5 By Probing Single-Molecule FRET*	43
Joseph	Pecha	BCM	*Using Geographic Information Systems (GIS) and Data Science to Uncover Important Factors for Health-driven Resiliency Plans: Preparing Houston for Future Disasters*	44
Joshua	Rosario Sepulveda	BCM	*Structure and Activation Mechanisms in a Group III Metabotropic Glutamate Receptor*	47
Leila	Saadatifard	UH	*Deep Quantitative Analysis of Multiplexed ImmunoHistoChemstry (IHC) Images of Alzheimer's Rat Brain Tissue*	45
Wilhelm	Salmen	BCM	*Mechanism of Neutralization of Human GII Noroviruses by Cross-reactive Human IgA*	46
Marzia	Savini	BCM	*Tissue Cross-Talk in Longevity Regulation*	48
Seth	Scott	UTMB	*MicroRNA-122 and Poly-C Binding Protein 2 Promote Hepatitis C Virus Replication by Binding to Overlapping Targets on the 5' UTR*	49
David	Shih	MDACC	*Modeling Mutational Processes in Cancer*	50
Manmohan	Singh	UH	*High-Throughput Optical Coherence Elastography for Colorectal Cancer Detection*	51
Eric	Smith	BCM	*Lineage-specific Epigenetic Reprogramming by Inorganic Arsenic in Primordial Germ Cells*	52
Lauren	Swanson	BCM	*Prenatal Exposure to Polycyclic Aromatic Hydrocarbons (PAHs) Augments Neonatal Hyperoxic Lung Injury and Alters the Gut Microbiome in Mice: Mechanistic Role of Cytochrome P450 (CYP)1A1, 1A2, and 1B1*	53
Jessica	Symons	UTHSCH	*Defining Transbilayer Lipidomic Profiles and Determining Consequences of Biophysical Asymmetry in Mammalian Plasma Membranes*	54

First	Last	Institution	Abtract Title	Poster #
Lillian	Thistlethwaite	BCM	*Connect the Dots: A Method for Interpreting Multi-Metabolite Perturbations for the Diagnosis of Metabolic Disease Using Disease-Specific Metabolomics Networks*	55
Luis	Vega	UTHSCH	*Characterization of Relationships Between Pathogenesis and Antimicrobial Resistance-Encoding Mobile Genetic Elements in Group A Streptococcus*	56
Carlos	Vera Recio	MDACC	*The Subclonal Architecture of Triple Negative Breast Cancer*	57
Ke	Wang	UTMB	*Small-Molecule Inhibitor of OGG1 Suppresses Proinflammatory Gene Expression and Inflammation*	64
Jamie	Wright	UTHSCH	*Role of Ataxia Telangiectasia and Rad3 Related Cell Cycle Regulation in Occlusive Vascular Disease Pathogenesis in Majewski Osteodysplastic Primordial Dwarfism Type II*	58
Zheng	Xu	UTMB	*Inhibition of OGG1' Interactions with its Genomic Substrate Decreases Proinflammatory Gene Expression*	63
Ahmet	Yavuz	BCM	*Metabolic Fate of Saturated and Unsaturated Dietary Lipids*	59
Lythou	Yeo	BCM	*Elucidating the Mechanism of Developmental Toxicity of the Anti-retroviral Drug Dolutegravir*	60
Matthew	Ykema	Rice U.	*Expression and Characterization of the Human Astrovirus (HAstV) VP90 Capsid Protein for Investigating the Structural Features of the Proteolysis-Mediated Viral Maturation Process*	61
Raziyeh	Yousefi	UTMB	*Computer Simulations Show Key Role Stochasticity of Replication Fork Speed Plays in the Dynamics of DNA Replication*	62

Urban Watersheds Seasonal Deleterious Effects on Colon Cells

Dayana Abdullah-Smoot, Environmental Toxicology PhD Candidate
Department of Environmental and Interdisciplinary Sciences, Texas Southern University, Houston, Texas

Steadily rising populations are turning Houston, Texas to the urban concrete jungle it is today. With communities growing and searching for homes, the need for new construction and establishments only increases the amount of impermeable surfaces relying heavily on the native waterways. Houston is known for its complex network of waterways surrounding and interweaving throughout the city, which are referred to as bayous. They work as a storm water runoff system to help prevent flooding, but many communities use the watersheds for recreational purposes. Local patrons boat, fish and swim in the bayous, unaware of the chemicals and pollutants lurking in the water. Due to the extensive flood history in Houston and the surrounding areas, communities do not have to be near the bayou to be inundated with the polluted water. After heavy rain for only an hour or two can cause bayous to rise, streets flood and water can make the find its way into low-lying homes. Current studies have shown that there are high concentrations of As, Hg, Cr, Pb and Zn in the water. We hypothesize that the metals and chemical compounds identified could cause the water to have deleterious effects on living organisms occupying the watersheds or using its resources. My study objectives include testing the cellular response when the cells are subjected to the bayou water samples collected during the summer and fall of 2018 to compare cellular viability and to determine which of the watersheds has the greatest over cytotoxic effects.

Water samples have been collected in triplicates from Mustang Bayou, Dickinson Bayou, Horsepen Bayou, Houston Clear Lake, Cypress Creek and Cypress Creek Lake, which are six major waterways in the surrounding Houston area. The methods included buffering the water sample that will not be detrimental to the cells, then serial dilute the samples in media for varying concentrations (control or 0%, 12%, 25%, 50%, 100%) and tested the bayou water samples on the HT29 colon cells using a cytotoxicity assay and proliferation assays at 12, 24, and 48-hour time points.

We observed a linear decrease in viable cells as the concentration of bayou water sample increases. The summer 2018 samples showed that Dickinson and the Mustang Bayous have highest record of viable after the 12hr treatment at 80% and Mustang Bayou remains the highest after the 24hr treatment at 62% but the 48 hr results show that Clear Creek had the highest cell viability at 30%. The fall 2018 samples showed that Mustang Bayou have highest record of viable after the 12hr and 24hr treatments at 71% and then 68% but the 48 hr results show that Dickenson Bayou had the highest cell viability at 30%. Cytotoxic effects were observed at every time point no matter the concentration of bayou water in the sample and the highest overall cellular viability was observed in the fall 2018 samples, because the overall viability percentages were roughly 20% higher than summer samples for both the 12hr and 24hr treatments although the 48hr results stayed consistant. Future studies include treating FHC colon cells with the same bayou water concentrations to compare to the HT29 cellular response and investigate the cytotoxic cellular effect with media containing the chemical elements at the concentrations determined from the chemical analysis.

This research was primarily supported by the National Science Foundation (NSF) through Texas Southern University (TSU) under the award numbers HRD-1622993, BCS-1831205, and HRD-1829184

Poster #2

What Do Neural Networks Learn? A Mathematical Comparison of Convolution Kernels and Image Processing Features

Actor JA[1], Riviere B[1], Fuentes DT[2]
1. Department of Computational and Applied Mathematics, Rice University
2. Department of Imaging Physics, University of Texas MD Anderson Cancer Center

Corresponding author: Jonas A Actor, Department of Computational and Applied Mathematics, Rice University, 6100 Main St, MS-134, Houston, TX 77005, email: jonasactor@rice.edu

Convolutional neural networks (CNN) perform state-of-the-art image segmentation. However, such networks are considered “black boxes”; the features learned by a CNN are typically not analyzed in terms of known image processing operations. We train a UNet, a type of CNN, and analyze the convolution kernels learned at each layer along two methods of analysis: we examine the learned convolutions to see if they cluster around known image processing features, and we compare the singular values of the linear convolution operator defined by the learned kernels, to see if they perform similar actions on each input channel. We find many of the convolution kernels are explained by common image processing operations, and that the distribution of the learned kernels is relatively stable throughout the UNet. These methods characterize what actions a CNN performs during segmentation, and as such could be used by clinicians to verify CNN segmentation predictions.

JAA is supported by a training fellowship from the Gulf Coast Consortia, on the National Library of Medicine Training Program in Biomedical Informatics & Data Science (T15LM007093), supplemented by the Ken Kennedy Institute Computer Science & Engineering Enhancement Fellowship, funded by the Rice Oil & Gas HPC Conference.

SeqScreen: A Biocuration Platform for Robust Taxonomic and Biological Process Characterization of Nucleic Acid Sequences of Interest

Albin D1, Nasko D2, Elworth L3, Jacob L3, Balaji A, Muthu P4, Godbold G4, Ternus K4, Treangen T1&3

1. Program in Systems, Synthetic and Physical Biology, Rice University
2. Center for Bioinformatics and Computational Biology, University of Maryland College Park
3. Department of Computer Science, Rice University
4. Signature Science, LLC

Corresponding author: Todd Treangen, Department of Computer Science, Rice University, 6100 Main st, Houston, Texas, E-mail: treangen@rice.edu.

Rapid advancements in synthetic biology and nucleic acid synthesis, in particular concerns about its intentional or accidental misuse, call for more sophisticated screening tools to identify genes of interest within short sequence fragments. One major gap in predicting genes of concern is the inadequacy of current tools and ontologies to describe the specific biological processes of pathogenic proteins. ***The objective of this work is to design software that sensitively assigns taxonomic classifications, functional annotations, and biological processes of interest to short nucleotide sequences of unknown origin (50bp-1,000bp)***. The overarching goal is to perform sensitive characterization of short sequences and highlight specific pathogenic biological processes of interest (BPoIs). The SeqScreen software executes these tasks in analytical workflows with Nextflow and outputs results in a final report. Local and global alignments differentiate hits to taxonomically-related sequences from similar but unrelated sequences, and an ensemble approach leverages multiple tools and databases to assign a variety of functional terms to each query sequence. Final assessments are determined from the predicted functional annotations, which leverage information in pre-existing databases, as well as new custom biocurations. Machine learning models predict each biological process of interest on large protein databases before incorporation into the SeqScreen framework to streamline computational efficiency, ensure the reproducibility of results, allow for version control, and facilitate the review of the automated predictions by expert biocurators.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1842494. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. All of the co-authors were either fully or partially supported by the FunGCAT program from the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via the Army Research Office (ARO) under Federal Award No. W911NF-17-2-0089. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, ARO, or the US Government.

Use of Host-Targeted, Immunomodulating Drugs to Combat *Clostridioides difficile* Infection

Andersson JA[1,2,3], Peniche AG[4], Boonma P[5], Galindo CL[6], Yalcinkaya N[1,2], Sha J[3], Chopra AK[3], Savidge TC[1,2], Dann SM

1. Department of Pathology and Immunology, Baylor College of Medicine, 2. Texas Children's Microbiome Center, Texas Children's Hospital, 3. Department of Microbiology and Immunology, University of Texas Medical Branch, 4. Department of Internal Medicine, University of Texas Medical Branch, 5. Faculty of Medicine, King Mongkut's Institute of Technology, 6. Department of Medicine, Vanderbilt University

Corresponding Author: Jourdan A. Andersson, Department of Pathology and Immunology, Baylor College of Medicine; Texas Children's Microbiome Center, Texas Children's Hospital, 1102 Bates Avenue, Houston, TX, Email: andersso@bcm.edu

Background and objectives: With traditional drug development being insufficient to keep up with the current demand for newer drugs to combat the ever-growing threat of infectious agents, we previously utilized a drug repurposing approach and identified 3 drugs, amoxapine (AXPN, anti-depressant), doxapram (DXP, breathing stimulant), and trifluoperazine (TFP, anti-psychotic), to mitigate fatality in a murine model of pneumonic plague. Additionally, we observed that when combined with vancomycin, these drugs were also effective against a murine model of *Clostridioides difficile* infection (CDI); however the mechanism(s) of action of these drugs remain unknown. As no direct effects on *Y. pestis* or *C. difficile* growth or virulence factors was observed *in vitro*, we hypothesize that these drugs are host-targeted therapeutics, acting to elicit protective host immune responses to combat disease.

Methods: Utilizing a murine model of CDI, we evaluated the efficacy of AXPN, DXP, and TFP and compared disease severity, gut microbiome, intestinal motility, and innate immune responses. We also compared drug efficacy and innate immune responses in germ-free mice and performed cytokine and neutrophil knockdown studies *in vivo*.

Results: All three drugs provided protection against CDI, when administered at the time of infection, and decreased bacterial load and toxin titers through indirect mechanisms. Drug treatment was not observed to affect intestinal motility or resolve infection induced dysbiosis, but the microbiota proved to be necessary for drug efficacy. Whole tissue transcriptomics revealed shared enhancement of several innate immune pathways by drug treatments including several genes in the IL-1, IL-22, inflammasome, neutrophil recruitment, and antimicrobial peptide pathways. Focusing on AXPN, importantly, we observed that knockdown of IL-33 or neutrophils by antibody treatment resulted in loss of drug efficacy against CDI.

Conclusions: Overall, this data indicates that drug efficacy is related to elicitation of a controlled immune response that effectively combats infection without causing collateral host damage. This data also provides further support of the early influx of neutrophils as well as upregulation of IL-33 being protective against CDI.

Acknowledgements: This research is supported by TPAMR T32 Fellowship T32-AI141349-01 and the Jeane B. Kempner Postdoctoral Fellowship and awarded to JAA, the John S. Dunn Collaborative Research grant awarded to AKC and TCS, NIH/NIAID R56 grant AI132682 awarded to AKC and SMD, and NIH/NIAID U01 grant AI124290 awarded to TCS.

Towards Computational Prediction of Off-target Toxicity in Cancer Immunotherapy

Antunes DA[1], Abella JR[1], Bishop A[2], Jackson K[3], Lizée G[3], and Kavraki LE[1]

[1]Department of Computer Science, Rice University, Houston, TX
[2]Department of Biology, Texas A&M University, College Station, TX
[3]Department of Melanoma Medical Oncology - Research, MD Anderson Cancer Center, Houston, TX

Corresponding author: Dinler Amaral Antunes, Department of Computer Science, Rice University, 6100 Main St, Houston, TX, E-mail: dinler@rice.edu.

Cancer immunotherapy treatments aim at inducing an immune response in the patient, specifically targeting and eliminating tumor cells. Some of the most promising immunotherapies are based on the cellular immunity, which relies on the functions of T-cell lymphocytes. T-cells are "trained" to identify non-self peptides displayed at the surface of other cells by Human Leukocyte Antigen (HLA) receptors. In the context of a cancer cell, this mechanism allows circulating T-cells to recognize tumor-associated peptides being displayed, which will in turn trigger T-cell activation and tumor elimination. This "natural" immunity can be boosted for cancer treatment, through vaccination with cancer-derived peptides or adoptive transfer of engineered cancer-specific T-cells. Despite successful treatment of many patients, recent clinical trials have also reported serious off-target toxicity induced by these treatments. In a famous example, melanoma-specific T-cells mistakenly recognized a titin-derived peptide in healthy cardiac cells, causing the deaths of at least 4 patients. This recognition of unrelated peptide-targets is known as T-cell cross-reactivity, being influenced by patient-specific variability and representing a major safety concern for many immunotherapies.

Currently, there are no reliable methods, computational or experimental, to predict the cross-reactivity risk for new immunotherapies. Our objective is to implement a new computational tool to predict potential off-target toxicity in cancer immunotherapy. Our tool relies on a local database of self-derived peptides that was built using immunopeptidomics data from SysteMHC Atlas and from proteomics experiments conducted by collaborators at UT MD Anderson. Given as input the amino acid sequence of a cancer peptide that will be targeted by a new immunotherapy, this tool provides a list of highly similar, experimentally determined, self-derived peptides that represent a potential safety concern for the treatment. This prediction is based on a variety of metrics including sequence-based algorithms, and 3-D structural analysis, being consistent with recent findings on the molecular basis for T-cell cross-reactivity. Our initial implementation focused on peptide sequence analysis, using algorithms to quantify sequence identity, shared biochemical properties, shared sequence motifs, HLA binding affinity and predicted immunogenicity. These methods were tested by predicting known cases of cross-reactivity, such as the MAGEA3-Titin reaction. We are also building a local database of modeled peptide-HLA complexes, generated with APE-Gen; a modeling tool previously proposed and validated by our group. These structures will allow us to account for structural features that are determined by the specific HLA alleles of a given patient. Once concluded, our prediction tool can be used for virtual screening, suggesting unrelated peptide-targets that should be experimentally tested to determine the specificity and safety of new T-cell based immunotherapies.

This work was funded in part by the Cancer Prevention & Research Institute of Texas (CPRIT), through Grant award RP170508, and through two Fellowships from the Computational Cancer Biology Training Program (RP170593). This work was also partially supported by a training fellowship from the National Library of Medicine Training Program in Biomedical Informatics (T15LM007093).

Identifying Drug Response Genes Using Evolutionary Action and Machine Learning in Human Cell Line Panels

Asmussen J[1], Parvandeh S[1], Hsu T[1], Lagisetty Y[1], Bourquard T[1], Shapiro D[1], Katsonis P[1], Subramanian D[2,3], Lichtarge O.[1]

1. Department of Molecular & Human Genetics, Baylor College of Medicine
2. Department of Computer Science, Rice University
3. Department of Electrical and Computer Engineering, Rice University

Corresponding author: Olivier Lichtarge, Department of Molecular & Human Genetics, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas, E-mail: lichtarg@bcm.edu

The concept of personalized medicine assumes that the molecular agents associated with drug response are identifiable and that this information can be used to develop a customized course of therapy for each individual within a given population. The realization of personalized medicine has been achieved for a limited number of malignancies, but in these examples typically only a subset of the disease population possesses the pharmacologically targeted molecular agent. There is clearly a need to identify the remaining molecular agents of drug response in these and other malignancies. Recent systematic attempts to identify molecular agents of drug response to anti-cancer therapies have employed large panels of human cell lines screened against FDA-approved compounds and compounds in clinical development[1–3]. Human cell line panels have been documented to be representative of primary tumors based on the comparison of their shared molecular features, lending confidence to this type of *in vitro* experimental approach[1]. Novel and known drug-response associations have been identified through the use of human cell line panels, although it is unclear if these associations are predictive in a clinical setting.

For this study, we will employ evolutionary action and three computational methods to identify drug response genes (DRGs) in panels of human cell lines treated with chemical compounds. Evolutionary action (EA) uses calculus to score a somatic mutation based on the amino acid residue at which the mutation occurs, taking into account its evolutionary importance, as well as the type of amino acid substitution at that residue[4]. We hypothesize that DRGs will exhibit a biased distribution of EA scores in sensitive or resistant cell lines. DRG identification methods will include: 1) EPIMUTESTR, a k-nearest neighbor machine learning algorithm, 2) EA Integral, a statistical analysis method and 3) EA Wavelet, a representational learning algorithm. Post-experimental analysis of high confidence genes identified by each method may provide information related to the mechanism of drug response that was not previously recognized.

Initial work has focused on a panel of 94 molecularly characterized human lung cancer cell lines treated with 222 drugs. Preliminary experiments have yielded a maximum drug response prediction accuracy of 0.648, where we used EPIMUTESTR to identify DRGs and elasticNet, random forest, and recursive partitioning through leave-one-out cross-validation for drug response prediction. We are currently working to improve this accuracy by incorporating an EA Pathway feature into the DRG prediction process, which will be piloted using Reactome pathways. Accuracy may also be increased by transitioning to a larger cell line panel with integration of the EA Pathway feature.

This work is supported by a training fellowship from the Gulf Coast Consortia, on the NLM Training Program in Biomedical Informatics & Data Science (T15LM007093)

1. Iorio, F. *et al*. *Cell* (2016), 2. Garnett, M.J. *et al. Nature* (2012), 3. Barrentina, J. *et al. Nature* (2012), 4. Katsonis *et al. Genome Res.* (2014)

Determination of Oxybenzone and Clofibrate in Water, Sediment, and Fish from Galveston Bay / Clear Creek Areas

Atobiloye S[1], Turner J[1], Davis F[1], Relangi P[1], and Yakubu MA[2]

1. Graduate Student, Department of Environmental & Interdisciplinary Sciences, Texas Southern University
2. Professor, Department of Environmental & Interdisciplinary Sciences, Texas Southern University

Corresponding Author: Sekinat Atobiloye, Department of Environmental & Interdisciplinary Sciences, Texas Southern University, Houston, Texas, E-mail: s.atobiloye0421@student.tsu.edu.

Environmental chemical contaminants have been of major concern as they can cause adverse health effects in single or multiple forms, even at trace levels. They are variously derived from different activities, and include pesticides, personal care products, pharmaceuticals, and byproducts of various industrial processes. There are growing awareness and concern about them because of the possibilities of their interacting with biological systems to cause serious health issues. They are often referred to as contaminants of emerging concern (CEC). Galveston Bay and Clear Creek receive water from several sources including runoffs which may contain a significant number/concentration of contaminants with a high possibility of human exposure. The potential adverse effects of the contaminants on humans, human environment, and the ecosystem can range from acute to chronic and are of significant importance. This study determined the presence of selected chemicals of concern: oxybenzone and clofibrate in water, sediment, and fish samples from the Galveston Bay and Clear Creek areas. Samples were collected from the identified locations, extracted using standard protocols, and analyzed using high-performance liquid chromatography coupled with an ultraviolet/visible detector with modified EPA methods. Results quantified, and bioconcentration factor (BCF) calculated. The results show that oxybenzone and clofibrate are present at different concentrations in samples from the locations analyzed. The average concentration of oxybenzone in water and sediment samples are 1.85 ± 0.82 and 1.30 ± 1.55 part per million (ppm) respectively. Oxybenzone was detected at an average of 38.49 ± 14.64 ppm in the fish tissues/organs, this concentration is approximately 2,081% and 2,961% of its concentration in the water and sediment samples. Clofibrate was detected at a concentration of 0.45 ± 0.31 and 2.28 ± 3.76 ppm in the water and sediment samples. Its mean concentration of 7.76 ± 2.46 ppm found in the fish tissues/organs is 17 and 3 times greater than the levels in water and sediment. Thus, the results revealed that oxybenzone and clofibrate are present in the locations analyzed and bioconcentrate in fish tissues/organs with BCF of 20 and 18, respectively. The BCF values indicate high probability of biomagnification of the contaminants in higher animals. Oxybenzone being the one with the highest biomagnification capacity may possibly produce more toxic interactions. The data from this research is an impetus for the continuous monitoring of the levels of CEC in the Galveston/Clear Creek areas bodies of water and the aquatic environment.

Source of Funding: None

Association of the Maternal Genome with Autism Spectrum Disorder: Use of Whole Genome Sequence Data for Genome-Wide, Gene-Based Analyses and Genetic Risk Scores

Bach MA
Department of Epidemiology, Human Genetics and Environmental Sciences (EHGES), The University of Texas Health Science Center at Houston (UTHealth), School of Public Health

Mentor: Laura E. Mitchell, UTHealth
Co-Mentor: Mohammad H. Rahbar, UTHealth

Corresponding author: MacKinsey A. Bach, Department of EHGES, UTHealth, School of Public Health, Houston, TX, E-mail: Mackinsey.A.Bach@uth.tmc.edu.

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder with poorly understood etiology despite an increase in prevalence in recent decades. Maternal genetic effects can contribute to offspring phenotype by modifying the *in utero* environment of the developing fetus. Prior genome-wide association studies (GWAS) of maternal genetic effects in ASD had null results and may have been underpowered to detect associations at the variant level. Additionally, prior epidemiologic studies generally support associations between maternal metabolic conditions, including obesity, diabetes, and hypertension, and ASD in offspring. However, the potential effects of maternal pre-clinical metabolic conditions and/or high genetic susceptibility to metabolic conditions on ASD risk in offspring have not yet been investigated. The objective of this study is to better characterize maternal genetic effects in ASD using a genome-wide approach and targeted genetic risk scores (GRS) for metabolic conditions.

Whole genome sequencing data from the Simons Foundation Autism Research Initiative, Simons Simplex Collection (SSC), which includes n=2,644 families, will be used to conduct analyses of maternal genetic effects in ASD. The proposed studies will use a case-control study design, in which mothers are cases and fathers are controls. Genome-wide gene-based analyses of maternal genetic effects will be investigated using variant set mixed model association tests. Genetic variants will be weighted based on functional annotations and grouped into gene-based variant sets. Gene Ontology term enrichment will then be used to identify pathways/processes related to ASD via the maternal genome. The targeted GRS analyses will use previously published GRS for body mass index, diabetes (type 1 and type 2), and systolic blood pressure derived from GWAS summary statistics to calculate GRS for the metabolic conditions in the SSC population. Linear and threshold effects of maternal GRS will be assessed for each metabolic condition on ASD in offspring. Interactions between maternal GRS and other environmental variables will also be explored.

The fetal environment is determined or modulated by the maternal genome. Identifying maternal genotypes associated with ASD in offspring may shed light on the complex etiology of ASD. Results from this study may be used to expand prenatal screening and counseling to include information about ASD.

Funding Acknowledgement: Research reported in this publication was supported by the National Institute of Environmental Health Sciences of the National Institutes of Health under Award Number T32ES027801. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

MicroRNA-147b in Myeloid Cells Promotes Resolution of Acute Lung Injury

Berg N[1], Yuan X[1], Wang Y[1], Eltzschig E[1]

1. Department of Anesthesiology, McGovern Medical School, University of Texas Health Science Center at Houston

Corresponding Author: Nathaniel Berg, Department of Anesthesiology, University of Texas Health Science Center at Houston, 6431 Fannin St., Houston, TX, Nathaniel.K.Berg@uth.tmc.edu

Rationale: Acute Respiratory Distress Syndrome (ARDS) is characterized by dysregulated inflammation of the lungs that results in decreased ability for ventilation and oxygenation of blood. ARDS occurs in a significant number of high-risk surgical and critically ill patients and is a leading organ injury associated with mortality. Currently, there exists no specific targets for ARDS. Here, we sought to identify microRNAs that could serve as novel therapeutic targets for the resolution of ARDS.

Methods: Using murine intratracheal injection of LPS to model acute lung injury, we performed a miRNA array and qPCR profiling over the time course of inflammation. Antisense Locked Nucleic Acids were used to pharmacologically inhibit miR-147b and the generation of conditional miR-147b knock-out mice were used to investigate the role of miR-147b during lung injury. Weight loss and mortality were recorded and histological acute lung injury was scored by a blinded pathologist. BAL fluid albumin and cytokines were measured with ELISA assay. To study the therapeutic potential miRNA-147b mimics and scramble control oligonucleotides were packaged into dioleoyl phosphatidylcholine (DOPC) and then delivered to mice intravenously (5 ug/injection) 2 and 4 days after the initiation of LPS-induced lung injury.

Results: miR-147b was found to be highly expressed in the lungs during the recovery phase of LPS-induced acute lung injury. Pharmacologic inhibition and genetic deletion of miR-147b resulted in increased mortality and delayed recovery in weight loss. miR-147b knockout mice had increased BAL fluid albumin and cytokine levels and increased histological evidence of lung injury. Mice treated with miR-147b packaged DOPC nanoparticles demonstrated improved weight gain during the recovery phase of acute lung injury and had decreases in BAL fluid albumin and histological evidence of acute lung injury compared to mice treated with scrambled control mimics.

Conclusions: Our findings suggest that miR-147b promotes the recovery of acute lung injury. Furthermore, treating mice with miR-147b packaged nanoparticles after the development of acute lung injury enhanced recovery.

Directed Evolution of a Potent Inhibitor of the Methicillin-Resistant *Staphylococcus aureus* (MRSA) Resistance Enzyme, PBP2a

Boragine D_1, Adamski C_1, Mehta S_2, Palzkill $T_{1,2}$

1. Department of Biochemistry and Molecular Biology, Baylor College of Medicine
2. Department of Pharmacology and Chemical Biology, Baylor College of Medicine

Corresponding Author: Timothy Palzkill, Department of Pharmacology and Chemical Biology, Baylor College of Medicine, One Baylor Plaza, Houston, Tx, timothyp@bcm.edu

Abstract:

Antibiotic resistance has manifested into a global health epidemic. One of the most widespread human pathogens, Methicillin-resistant *Staphylococcus aureus* (MRSA) encodes a novel penicillin-binding-protein, PBP2a. Production of PBP2a by MRSA confers resistance to nearly all β-lactam antibiotics by continued peptidoglycan cell wall synthesis, even at high concentrations of antibiotic. The transpeptidase domain of PBP2a shares structural homology with class A β-lactamases, bacterial enzymes that inactivate β-lactam antibiotics. Class A β-lactamases, such as TEM-1, are inhibited by protein-based inhibitors named β-lactamase inhibitory proteins (BLIPs) and BLIP-II potently inhibits TEM-1 with a K_D 0.48pM. It was previously found that BLIP-II also weakly binds to PBP2a in the low micromolar range (K_D 1.5μM), in contrast to BLIP-II's potent inhibition of TEM-1. We hypothesize that through directed evolution, we can re-engineer BLIP-II for altered specificity to potently bind PBP2a. A directed evolution approach using phage display affinity selection was used to identify a BLIP-II double mutant, N50A:Y113H, that enhanced the binding affinity to PBP2a 30-fold to a K_D of 50nM. An additional directed evolution cycle, starting with the tighter binding BLIP-II$_{N50A:Y113H}$ template, was then performed to select for PBP2a binding while driving the binding affinity down to low nanomolar range. To date, we have identified several BLIP-II$_{N50A:Y113H}$ mutants that exhibit an additional ~3-5-fold enhancement in binding affinity to PBP2a and have also identified a G205W mutant that improved binding an additional 50-fold with a K_D ~1nM. This is a >1,500-fold enhancement in binding affinity to PBP2a compared to wild-type BLIP-II. However, the BLIP-II$_{N50A:Y113H}$ mutants resulted in >5,000-fold loss in binding to TEM-1, K_i 0.09nM, thereby indicating a change in selectivity compared to wild-type BLIP-II. Each BLIP-II mutant resulted in enhanced binding affinity to PBP2a, while simultaneously weakening the binding affinity to TEM-1. Thus, the evolved BLIP-II variants have shifted from high affinity for □-lactamases and low affinity towards PBPs, towards high affinity for PBPs with reduced affinity for □-lactamases. These results suggest that BLIP-II can be further optimized and serve as a scaffold for developing potential PBP2a inhibitors.

Funding by:
NIH RO1 AI32956-25 to TP
Gulf Coast Consortia: HAMBP T32 GM008280-29 to DB

dsDNA Packed inside Phage Capsids: Structure and Defects Emergence

Bores C[1] and Pettitt BM[1]
1.Biochemistry and Molecular Biology, UTMB, Galveston TX

Corresponding author: Cecilia Bores, Biochemistry and Molecular Biology, UTMB, 301 University Blvd, Galveston, TX Email: cbquijan@utmb.edu.

DNA packaging and ejection are two critical moments in dsDNA bacteriophages lifecycle and their understanding is decisive for the effective application of phages as an alternative to antibiotics. The forces needed to pack the DNA molecule to near crystalline density (~0.5g/ml) combined with the geometrical constrains of the phage capsid determine the conformation of the confined DNA. Most theoretical studies that have been performed to better understand how the conformation of the DNA inside bacteriophage capsids is considered DNA as a perfect elastic rod and predict highly ordered structures. However, the emergence of more disordered conformations exhibiting defects such as knots, kinks, loops, … that hinder both the insertion and extraction of the DNA molecule in the phage capsid is also plausible. This raises the central question of the present work - how much order and disorder are reasonable (or required) when DNA is confined inside the phage capsid?

We have performed Molecular Dynamics simulations using oxDNA model for dsDNA and a purely repulsive harmonic wall representing the proteic capsid of the bacteriophage to mimic the packing process in phage ϕ29. We have thoroughly analyzed the DNA conformation by means of density profiles and correlation functions during packing finding different results depending on how fast the DNA is being inserted. DNA structure predicted by these simulations show patterns that agree with experiments, cryoEM and X-ray diffraction, but many features in a more realistic capsid model – presence of multivalent ions, torsional forces, and local attractive/repulsive sites in the capsid or an elongated shape - might contribute to the emergence of these or other characteristics.

We gratefully acknowledge the Robert A. Welch Foundation (H-0037) and the National Institutes of Health (GM- 066813) for partial support of this work.

Matrix Decomposition of Omics Data and Pathway Correlations

Borzou A[1], Sadygov R G[1]

1. Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, 301 University Blvd, Galveston, TX, Email: ahborzou@utmb.edu

Non-alcoholic fatty liver disease (NAFLD) is the most chronic liver disease, and its diagnosis and treatment are currently challenging. About 30% of the population in western countries are affected by the disease. Proteomics provides a high-throughput means to study proteome in NAFLD. We start with the basic principles and build proteomics analysis tools with applications in prevention/diagnosis/drug study of NAFLD.

We assume that a biological system is governed by the thermodynamics laws of physics and its entropy in terms of the protein abundance is the same as in every other physical system. Consequently, the distribution function can be derived through entropy-probability relation. Furthermore, we assume that protein networks exist such that the abundances of some proteins with some weights add up to a constant in every replicate or experiment of the same biological entity. These constraints, which are derived from data, reduce the number of possibilities or the degrees of freedom of the system. Next, we construct the partition function as the sum over the entropy-based distribution function in the domain of possible modes. The underlying interactions among the proteins of the system, and also a classification of the samples in the dataset(based on their occurrence probability) can be derived from our partition function.

To demonstrate the method, we analyze the plasma proteome of 48 patients at different stages of liver disease. Fig. 1 shows all of the observed proteins where an edge indicates significant interaction according to our partition function. Such network analysis allows us to study the effects of a stage of the disease on the protein co-regulations. This can potentially lead to identification of biomarkers. Fig. 2 shows the separation of the patient samples along the y-axis according to their distance(determined by our partition function) from the maximal physical entropy state. This can be seen as one of the diagnostic techniques for NAFLD.

Research reported in this publication was supported in part by the NIGMS of the NIH under award number R01GM112044. AB was partly supported by a training fellowship from the Gulf Coast Consortia, on the NLM Training Program in Biomedical Informatics & Data Science T15LM007093.

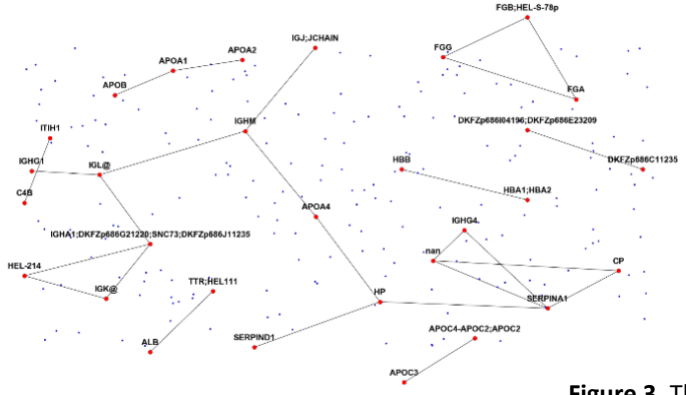


Figure 1. *The identified plasma proteins of all 48 patients. The blue dots represent proteins which do not interact significantly with any other measured protein. Note that most of the connected proteins belong to the same family. This shows the ability of our method in finding co-regulated proteins.*

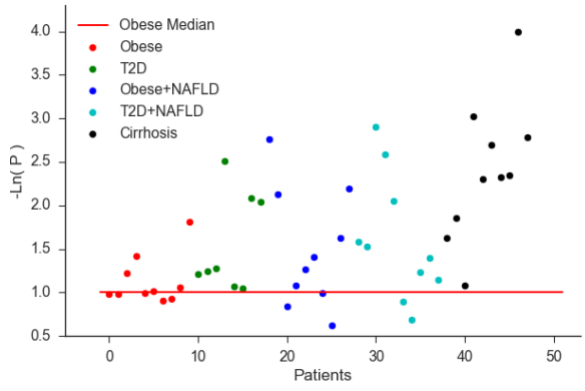


Figure 3. The y-axis shows the distance from the maximal entropy state equivalent to the minus of the log of the probability of that condition. The 48 patients are shown along the x-axis. The color labels are clinically determined stages of NAFLD. Our method is capable of separating the patients with liver disease (the last 30 points on the x-axis) from healthy patients. Notice that the Cirrhosis patients with the most severe liver disease are separated more than the rest. Patient samples with other health issues can be similarly classified. This will be particularly useful in cases with no suitable clinical diagnosis method.

Generating a Molecular Understanding of the Role of Co-factors in the p53 Response

Campos Acevedo A[1], Jianpeng M[1,2], Qinghua W[1]

1. Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, Texas, USA
2. Department of Bioengineering, Rice University, Houston, Texas, USA

Corresponding author: Qinghua Wang, Department of Biochemistry and Molecular Biology, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas. E-mail: qinghuaw@bcm.edu

The p53 gene is one of the most commonly mutated genes in human cancer. About 50% of human cancer contains a p53 mutation while the rest is caused by alterations in components of the p53 pathway. p53 has a central role in various important pathways that are responsible for maintaining cellular integrity. However, p53 can be influenced by several other groups of proteins during DNA damage response, such as the formation of complex p300-JMY-TTC5. A high-resolution structure of the complex is crucial to unveil, at a molecular level, an understanding of the interplay between these factors and their role in the p53 response. There are structural studies focused on the individual components of this complex; however, there is no structural evidence of the complex as a whole. With the development of a new approach for 3D refinement to improve Cryo-EM map resolution, we will identify the crucial structural elements of this complex.

We propose a combination of experimental studies using Cryo-EM to determine the atomic structures of the p53 activation complex and sub-complexes. The development of new computational tools to extract high-resolution images from cryo-EM data will help to overcome the *ill-posed* problem (high level of noise and incompleteness of the data). Limitations that arise when processing Cryo-EM data include micrograph artifacts, relative particle orientation, and heterogeneity. To minimize these limitations, it is crucial to develop new computational algorithms to increase the objectivity of our map interpretation. To accomplish this, we will design a new computational method that includes a novel real space penalty function that incorporates sparseness and smoothness into the refinement process. This novel refinement method will helps to alleviate conformational heterogeneity and non-uniform angular sampling in cryo-EM analysis.

Here we present the expression and purification of the JMY-p300-TTC5-p53 complex as well as the JMY-TTC5 sub-complex. A negative stain map of the JMY-TTC5 sub-complex was obtained by electron microscopy. The low-resolution map provides a significant contribution to future high-resolution studies to develop a more complete understanding of the regulation that controls p53 activity and may lead to improved therapeutic strategies for treating cancer.

Cryo-grids optimization will be required to get an atomic structural map of the complex and sub-complex. With the atomic structure maps, we will perform the refinement using our new 3D refinement method that is under development to suppress bias and increase the objectivity in the Cryo-EM analysis.

*This research was funded by the CPRIT training fellowship Program in Computational Cancer Biology Training Program of the Gulf Coast Consortia (CPRIT Grant No. RP170593).

Investigating Mosaicism in Blood for Cardiovascular Disease Risk

Chander V[1], Gibbs R[1], Milosavljevic A[1]

1. Department of Genetics and Genomics, Baylor College of Medicine

Corresponding author: Varuna Chander, Genetics and Genomics, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas, E-mail: varuna.chander@bcm.edu

Abstract:

Common diseases constitute a significant health burden, affecting a large segment of the population and the leading cause of death in the United States. Despite this impact, the underlying genetic causes of common diseases cancer have been poorly understood. Advancements in sequencing technologies have facilitated the discovery of somatic driver mutations that cause clonal expansion of hematopoietic stem cells (HSCs) in the blood. This common age-related biological process termed as Clonal Hematopoiesis of Indeterminate Potential (CHIP) is associated with significantly increased risk of developing blood cancer and cardiovascular disease (CVD), two common causes of human mortality. Despite this impact to a major proportion of the aging population, there is a considerable gap in knowledge about the factors that cause clonal expansions in healthy individuals, or how they lead to cardiovascular disease.

To advance our understanding, we apply sensitive computational methods to perform somatic variant detection within the cardiovascular disease enriched cohorts: Heart Care Cohort and ARIC (Atherosclerosis Risk in Communities (ARIC) cohort data at the Human Genome Sequencing Center (HGSC). We will employ two sensitive variant calling pipelines for somatic variant calling and assess performance. We will identify somatic mutations in genes associated with CVD events with a potential role in CHIP and perform additional phenotype association studies.

Additionally, we are interested in unexplored mechanisms: inherited variants, structural variations and will identify likely pathogenic variants for follow-up using molecular studies. Furthermore, since age is the strongest risk factor for CHIP and DNA methylation is one of the best characterized epigenetic modifications during aging, future sequencing studies are underway to study the potential epigenetic effect on clonal expansions in blood. Integrating genetic and epigenetic layers of information, coupled with experimental validation will enable identification of novel driver genes and reveal insights into the mechanisms of clonal expansions in blood cells.

Statement of Significance:

Age is the dominant risk factor for CVD, but the mechanistic basis for why age predisposes to CVD is not entirely understood. Besides, more than 60% of patients with CVD have either zero or one conventional risk factors. Through our proposed research, we address an important biological link between aging, CVD, cancer, with tremendous potential impact on the public health of aging populations. With sensitive computational methods, we will identify genetic causes of cardiovascular disease and better screen for disease to enable proactive clinical decision making. Ultimately, this knowledge will benefit biomedical scientists, clinicians and patients where insights will inform clinical care in the aging population.

This work is supported by the NLM Training Program in Biomedical Informatics and Data Science T15LM007093, Program Director Dr. Lydia Kavraki.

Spectral Coarse Graining: Deep Learning of Simple, Dynamically Consistent Protein Models

Charron NE[1,2], Feliks Nüske[1,3], Jiang Wang[1,3], Lorenzo Boninsegna[1,3] Patel A[4,5], Clementi C[2,3]

1. Department of Physics, University of Houston, Houston TX
2. Center for Theoretical Biological Physics, Rice University, Houston TX
3. Department of Chemistry, Rice University, Houston, TX
4. Department of Electrical and Computer Engineering, Rice University, TX
5. Department of Neuroscience, Baylor College of Medicine, Houston TX

Corresponding author: Nick Charron, Physics and Astronomy, Rice University, 6100 Main St, Houston, TX 77005, E-mail: nec4@rice.edu

Coarse grain models of proteins offer promising gains in both computational efficiency for molecular simulations and the development of simple physical interpretations. Recent efforts have focused on formulating the development of coarse grained force fields as a supervised learning problem, taking advantage of deep learning techniques for handling highly non-linear multibody effects produced by imposing coarse grained representations. Yet, while traditional methods formulated on the basis of force matching have shown to successfully reproduce the essential thermodynamics of the underlying all-atom representations, these methods are generally unable to reproduce the dynamics characterizing metastable transitions. In this work, we present a deep learning method that utilizes spectral information from simulation data to preserve essential dynamics of the original system. We follow a Koopman-motivated approach based on weak integration of the distribution dynamical eigenequation characterized by an underlying generator in order to optimize the dynamical consistency between fine grain and coarse grain systems. Through this method, we can recover coarse grain representations of the free energy landscape that preserve essential dynamical information. This approach may also be combined with force matching methods, forming the basis for a potentially powerful data driven approach to coarse graining.

This work was funded by NSF grants: CHE-1265929, CHE-1738990, PHY-1427654, Welch Foundation grant C-1570m and NLM training grant 5T15LM007093-27 . The researchers acknowledge the support of the Rice Center for Theoretical Biological Physics and the Gulf Coast Consortia.

Loss of miR-290 Modifies the Risk of Embryonic Neural Tube Defects in Maternal Diabetes

Decker SC[1], Parchem JG [2], Keuls RA[1], Paikari A[2,3], Blelloch R[3], Parchem RJ[1]

1. Baylor College of Medicine, Houston, TX
2. University of Texas Health Science Center at Houston (UTHealth), Houston, TX
3. University of California, San Francisco, San Francisco, CA

Corresponding author: Samantha C. Decker, Neuroscience, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX, Samantha.Decker@bcm.edu.

Pre-gestational diabetes is an established risk factor for miscarriage and fetal anomalies, including neural tube defects (NTDs). This increased risk is a concerning one, due to diabetes' prevalence increasing globally and NTDs already being a leading cause of infant mortality. The mechanisms of diabetic embryopathy are not well understood and are considered multifactorial, resulting from the interaction between a hyperglycemic environment and genetic predisposition. We show that genetic deletion of the embryonic stem cell microRNA family, miR-290, leads to cranial NTDs in embryos from diabetic, but not non-diabetic, pregnant mice. Among diabetic mice, we observed a dose- dependent increase in embryonic NTD frequency associated with miR-290 loss with KO embryos showing higher rates of NTD than heterozygous littermates. Previously, miR-290 expression was shown to be down-regulated in the embryo starting at ~E7.5 prior to neurulation and is not expressed in embryonic tissue by the end of cranial neural tube closure (E9.5) but is maintained in the extra-embryonic tissues (placenta) and yolk sac. This expression profile suggests miR-290 is regulating genes in the embryo during neural tube closure but is also potentially regulating genes in extraembryonic tissue that may confer the disease environment-specific phenotype we have observed. Future directions of this project include characterizing this model using differences between littermates in a diabetic mother to determine potential cellular defects causing NTD. There are many well characterized pathways implicated in NTD and we plan on using immunofluorescence to determine differences in cell death, proliferation, and differentiation between KO embryos and their WT littermates. We also plan to use RNA sequencing and predicted miR-290 targets to identify candidate genes that may be responsible for this gene-dose and maternal environment dependent effect.

TPEHS fellowship: T32 ES027801

FAM170A Loss Causes Subfertility and Defective Sperm Motility in Mice

Devlin DJ[1,2], Nozawa K[2,3], Ikawa M[4,5,6], and Matzuk MM[2,3,7,8*]
1. Interdepartmental Program in Translational Biology and Molecular Medicine, Baylor College of Medicine
2. Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX, USA
3. Center for Drug Discovery, Baylor College of Medicine, Houston, TX, USA
4. Research Institute for Microbial Diseases, Osaka University, Suita, Osaka, Japan
5. Graduate School of Pharmaceutical Sciences, Osaka University, Suita, Osaka, Japan
6. The Institute of Medical Science, The University of Tokyo, Minato-ku, Toyko, Japan
7. Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX, USA
8. Department of Pharmacology and Chemical Biology, Baylor College of Medicine, Houston, TX, USA

Corresponding author: Martin M. Matzuk, Department of Pathology & Immunology, Baylor College of Medicine, One Baylor Plaza, Houston, TX, Email: mmatzuk@bcm.edu

Family with sequence similarity 170 members A and B (FAM170A and FAM170B) are two testis-specific, acrosome-localized paralogous proteins conserved throughout mammals. While *in vitro* experiments in previous literature suggested FAM170B plays a role in mouse sperm acrosome reaction, the role of FAM170A in the testis has not been explored. In this study, we have used CRISPR/Cas9 to generate null alleles for each gene and homozygous null male mice were mated to wildtype (WT) females for six months to assess fertility. *Fam170b* KO males were found to produce normal litter sizes, had no significant difference in sperm counts, and sperm morphology appeared normal. In contrast, mating experiments revealed significantly reduced litter sizes from *Fam170a* KO males compared to controls. Though both *Fam170a* KO and control males had normal testis weights, CASA experiments revealed a significant reduction in *Fam170a* KO sperm count. Although overall percent motility was similar to control males, KO males had dramatically reduced progressive motility. In addition, light microscopy of *Fam170a* KO sperm revealed abnormal sperm head morphology and a bent neck. Data from mating Fam170a/Fam170b double KO (dKO) males to WT females revealed that dKO causes more severe subfertility. This observation suggests that both FAM170A and FAM170B operate within the same pathway, though FAM170A has a more vital impact on male fertility. This work will further our molecular understanding of sperm function and could help improve male infertility diagnoses.

Research was supported by Eunice Kennedy Shriver National Institute of Child Health & Human Development grants P01 HD087157 and R01 HD088412, the Bill & Melinda Gates Foundation grant OPP1160866, and training fellowships from the Interdepartmental Program in Translational Biology and Molecular Medicine grant T32 GM088129 and the Training Interdisciplinary Pharmacology Scientists (TIPS) Program grant T32 GM120011.

Use of a Cell-free Protein Synthesis System and Coupled Enzyme Reactions to Synthesize Natural Products

Alex Ditzel[1], George N. Phillips Jr.[12]
1. Department of BioSciences, Rice University
2. Department of Chemistry, Rice University

Corresponding author: Alex Ditzel, Department of BioSciences, Rice University, 6100 Main Street, Houston, TX, Email: AlexDitzel@rice.edu

Natural products are a great source of pharmaceuticals, providing a majority of all small molecule drugs that exist today. However, due to the difficulty in producing them, pharmaceutical companies shifted away from natural products and towards synthetic chemistry because of the ease of screening large libraries of synthetic molecules. Unfortunately, very few of these synthetic molecules were biologically active in a useful way, so many companies want to shift focus back towards natural products, but it can be difficult to produce a range of products and screen them. Creating natural products through organic synthesis can take years of effort and synthesis *in vivo* in heterologous hosts can also be difficult and time consuming. Therefore, in order to allow for easier screening and production of natural products, I will be demonstrating the use of a novel cell-free system to assemble natural products *in vitro* using coupled enzyme reactions. This allows for higher throughput screening with the help of robotic automation while also allowing for complete control of the constituents of the system so that metabolic engineering becomes greatly simplified or eliminated. This may allow for the production of products that are very difficult to assemble within cells and may allow for more molecules to be screened for biological activity than traditional cell-based expression methods.

Supported by a training fellowship from the Gulf Coast Consortia, on the Houston Area Molecular Biophysics Program (Grant No. T32 GM008280) and by NIH grants R01 GM115261 and R01 GM121060.

Serotonin Promotes Epithelial Restitution through Goblet Cell Mediated Secretion of MUC2 and TFF3

Engevik M[1,2], Chang-Graham A[3], Danhof H[3], Engevik KA[3], Britton R[3], Hyser J[3], Versalovic J[1,2]
[1]Department of Pathology & Immunology, Baylor College of Medicine
[2]Department of Pathology, Texas Children's Hospital
[3]Department of Molecular Virology & Microbiology, Baylor College of Medicine

Background: Although serotonin is best characterized in the central nervous system, recent studies are beginning to shed light on the critical role for serotonin in intestinal homeostasis. Serotonin is primarily produced and secreted by intestinal enterochromaffin cells and elevated serotonin levels are noted in response to intestinal tissue damage; indicating that serotonin release may mediate important repair mechanisms. Recently serotonin-receptor 4 (5-HTR4) was identified on goblet cells; a cell type known to release the mucin protein MUC2 and the wound-healing peptide trefoil factor 3 (TFF3). To date, few studies have explored the link between serotonin, goblet cells, and wound repair. In a microbial-centered approach, our lab has previously identified a single commensal bacterium, *Bifidobacterium dentium*, which in gnotobiotic mice is able to increase luminal and plasma serotonin and elevate intestinal MUC2 and TFF3. As a result, we hypothesize that *B. dentium* is able to stimulate serotonin production by enterochromaffin cells and that serotonin activates serotonin receptor 4 (5-HTR4) on goblet cells to promote MUC2 and Trefoil Factor 3 (TFF3) release. We speculate that TFF3 will activate its receptor CXCR4 to promote actin cytoskeleton rearrangement and restitution.
Methods & Results: Human intestinal enteroids, also known as organoids, were propagated from jejunal biopsies derived from healthy individuals. The enteroid system is a biologically relevant *in vitro* model of the intestinal epithelium as it contains all epithelial cell types and maintains segment specificity. Addition of *B. dentium* metabolites to human intestinal enteroid monolayers resulted in serotonin release by 30 min in a dose-dependent manner. Human enteroids typically harbor only ~1-3 enterochromaffin cells. To enhance the levels of enterochromaffins cells, human intestinal enteroids were lentivirus transduced to stably engineer doxytetracycline-inducible expression of neurogenin-3 (NGN3), a transcription factor that drives enterochromaffin cell differentiation. In this model, *B. dentium* metabolites elicited a more robust serotonin response. These two model provide strong support for the ability of *B. dentium* secreted products to drive serotonin release from enterochromaffin cells. Next, we sought to identify whether serotonin could activate goblet cells and promote wound repair. Importantly, RNAseq of human intestinal enteroids revealed that this culture system harbored the molecular machinery proposed in our model. We observed high levels of goblet cell markers (MUC2, TFF3, AGR2, CDX2, MEP1B, FCGBP, KLK1), expression of 6 of the serotonin receptors including 5-HTR4, and expression of the TFF3 receptor CXCR4. Application of serotonin or *B. dentium* metabolites to enteroid monoalyers resulted in secretion of MUC2 and TFF3. To identify the key components of wound repair, human intestinal enteroids were transduced with the calcium sensor GCaMP6S and the actin sensor LifeAct. To mimic intestinal wounds, human enteroid monolayers were scratched and cell migration was monitored by live imaging. *B. dentium* metabolites, exogenous serotonin and exogenous TFF3 all increased repair rates of enteroid scratches compared to that of media controls. Inhibition of the TFF3 receptor CXCR4 in presence of both *B. dentium* metabolites and serotonin delayed repair, indicating the requirement of TFF3 in serotonin-mediated epithelial repair. **Conclusions**: Collectively this work indicates that microbe-induced serotonin activates 5-HTR4 on goblet cells to stimulate MUC2 and TFF3 release. In turn, TFF3 mediates epithelial repair. This work points to commensal microbes and serotonin as potential modulators of TFF3 and has broad implications for developing novel therapeutic strategies for treating wounds.
Funding: Danone YIGPRO v20190708

Identifying the Role of Purinergic and Calcium Signaling in Rotavirus Infection

Engevik KA[1], Chang-Graham A[1], Engevik MA[2], Perry JL[1], Hyser JM[1]

[1]Department of Molecular Virology and Microbiology, Baylor College of Medicine

[2]Department of Immunology and Pathology, Baylor College of Medicine

Corresponding author: Kristen Engevik, Department of Molecular Virology and Microbiology, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas, kengevik@bcm.edu

Background: Calcium (Ca^{2+}) is a ubiquitous messenger that influences numerous cellular processes, and therefore Ca^{2+} signaling is tightly regulated by cells. Ca^{2+} signaling dysregulation results in severe and potentially life-threatening diseases, which is exemplified by rotavirus (RV) infection. RV is an enteric virus that causes life-threatening diarrhea in children, resulting in ~198,000 deaths each year. Pathophysiological consequences of RV infection are widely studied, yet host Ca^{2+} signaling pathways and the mechanisms by which RV exploits them to cause diarrhea remain incompletely characterized. We have found that RV infection increases Ca^{2+} signaling both within infected enterocytes and in surrounding uninfected cells through paracrine signaling. This manifests as intercellular Ca^{2+} waves that originate from the infected cell and propagates through uninfected cells mainly through ADP and the P2Y1 receptor. **Hypothesis:** We hypothesized that ADP purinergic signaling activate Ca^{2+} dependent pathways in various intestinal cell types, including goblet cells, enteroendocrine cells and macrophages. **Methods and Results:** We generated cell lines and human intestinal enteroids (HIEs) stably expressing cytosolic genetically-encoded calcium indicators to characterize calcium signaling throughout RV infection by time-lapse imaging. We found that P2Y1-mediated signaling was critical for activation of secretory epithelial cells, including induction of serotonin secretion of enterochromaffin cells and mucus secretion from goblet cells in human intestinal enteroids (HIEs) and mucin-producing cell lines. In monkey kidney MA104, RV-infection chemoattracted RAW and bone-marrow derived macrophages, which also harbor the P2Y1 receptor. This effect was blunted by pharmacological inhibitors of the P2Y1 receptor. Consistent with our *in vitro* findings, we observed that murine RV infection promoted secretion of serotonin, mucin and accumulation of macrophages. These effects of minimized in the presence of P2Y1 inhibitors and in P2Y1 knock out mice. **Conclusion:** Collectively these findings point to the role of purinergic signaling and Ca^{2+} waves in the pathophysiology of RV-infection. Understanding the role ADP signaling via P2Y1 receptor plays in RV will provide mechanistic insights into the homeostatic function of purinergic signaling in the GI tract.

Funding: R01 DK115507 (Hyser), T32 DK007664 (Engevik)

Methods For Copy Number Aberration Detection From Single-cell DNA Sequencing Data

Fan X[1], Edrisi M[1], Navin N[2], Nakhleh L[1]
1. Department of Computer Science, Rice University
2. Department of Genetics, the University of Texas M.D. Anderson Cancer Center

Corresponding authors: Xian Fan and Luay Nakhleh, Department of Computer Science, Rice University, 6100 Main St, Houston, TX, E-mail: {xian.fan,nakhleh}@rice.edu

Single-cell DNA sequencing technologies are enabling the study of mutations and their evolutionary trajectories in cancer. Somatic copy number aberrations (CNAs) have been implicated in the development and progression of various types of cancer. A wide array of methods for CNA detection has been either developed specifically for or adapted to single-cell DNA sequencing data. Understanding the strengths and limitations that are unique to each of these methods is very important for obtaining accurate copy number profiles from single-cell DNA sequencing data. Here we review the major steps that are followed by these methods when analyzing such data, and then review the strengths and limitations of the methods individually. In terms of segmenting the genome into regions of different copy numbers, we categorize the methods into three groups, select a representative method from each group that has been commonly used in this context, and benchmark them on simulated as well as real datasets. While single-cell DNA sequencing is very promising for elucidating and understanding CNAs, even the best existing method does not exceed 80% accuracy. New methods that significantly improve upon the accuracy of these three methods are needed. Furthermore, with the large datasets being generated, the methods must be computationally efficient. The study was supported by the National Science Foundation grant IIS-1812822 (L.N.). X.F. was supported in part by a Computational Cancer Biology Training Program (CPRIT Grant No. RP170593).

Aftermath of Hurricane Harvey: Bacterial Analysis in Houston Area Bayous

Galvin S[1], Rosenzweig JA[2]
1. Department of Environmental and Interdisciplinary Science, Texas Southern University
2. Department of Biology, Texas Southern University

Corresponding author: Shari Galvin, Department of Environmental and Interdisciplinary Science, Texas Southern University, 3100 Cleburne St., Houston, TX, Email: s.galvin5961@student.tsu.edu.

Houston, also known as the "Bayou City," contains 2500 miles of waterways. In August of 2017, Hurricane Harvey released an unprecedented amount of water over Houston and surrounding communities, causing extreme flooding of area bayous. The flooding of waste water treatment plants posed a health risk to the public. Gut bacteria, such as *E. coli*, can be used as an indicator to measure the safety of water for recreational, industrial, and agricultural purposes. Although indicator bacteria may not directly cause illnesses, they are indicators of potentially harmful pathogens in water-bodies. The greater Houston area bayous that were chosen for this study were Greens, Buffalo, Halls, White Oak and Hunting. To assess the impact of this major flooding event on bacterial populations, we conducted a comparative study between water samples collected during the summer (pre-Harvey) and fall (post-Harvey) months of 2017. Water samples, collected from 15 sites, were tested for the presence of bacteria by quantifying total and enteric bacterial counts. Total bacterial loads were determined using broad medium (i.e., Nutrient Agar) and plate counting. To analyze enteric bacteria, selective and differential medium (i.e. MacConkey) was used along with plate counting. The indicator bacteria, *E. coli*, was isolated from water samples by membrane filtration according to the methodology outlined in USEPA Method 1603. Additionally, meta-genomic analysis of representative bayou water samples was conducted to comprehensively evaluate bacterial population distributions pre- and post-Harvey. Our results indicate that Hunting, Buffalo and Halls bayou had significantly higher bacterial loads in the summer (pre-Harvey) when compared to the fall (post-Harvey). However, the overall total bacterial loads of Greens bayou were significantly higher in the fall samples when compared to the summer. Except for Buffalo Bayou, temporal analysis of water quality revealed a noticeable increase in the presence of *E. coli* in post-Harvey samples when compared to pre-Harvey samples. Pathogenic bacteria such as *Clostridium spp*, environmental *Chlamydiae*, and *Legionella spp* where present in higher numbers in the representative pre-Harvey samples when compared to post-Harvey. Conversely, *Pseudomonas spp.*, was present in the environment at 3-fold higher levels in the fall when compared to the summer. This genus contains species that can survive for months in water with minimal nutrients which could account for the higher number in the fall.

This research was primarily supported by the National Science Foundation (NSF) through Texas Southern University (TSU) under the award numbers HRD – 1622993, BCS-1831205, HRD-1829184 and funding source National Science Foundation RISE (HRD-1345173).

Estimating Cellular Density in Glioma using MR Imaging Data

Gates EDH[1,7], Lin JS[1,8], Weinberg JS[2], Prabhu SS[2], Hamilton J[3], Hazle JD[1], Fuller GN[4], Baladandayuthapani V[5], Fuentes D[1], and Schellingerhout D[3,6]

Department of: [1]Imaging Physics, [2]Neurosurgery, [3]Neuroradiology, [4]Pathology [5]Biostatistics, [6]Cancer Systems Imaging, University of Texas MD Anderson Cancer Center, Houston, TX
[7]University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences
[8]Department of Bioengineering, Rice University, Houston TX

Corresponding author: Dawid Schellingerhout, Neuroradiology and Cancer Systems Imaging, UTMDACC, 1515 Holcombe blvd., Houston, TX. E-mail: Dawid.Schellingerhout@mdanderson.org

Introduction: Increased cellular density (CD) is a hallmark of cancerous growth and a key feature in histologic analysis and grading of gliomas. Being able to estimate cellular density throughout a tumor would be a valuable tool to probe how tumors infiltrate and transition from diseased to healthy brain. The purpose of this work is to do so using a machine learning model trained on preoperative magnetic resonance imaging and spatially specific tissue samples.

Methods: Data was collected as part of an IRB-approved clinical trial for adult, treatment-naïve, glioma patients. Up to 5 separate tissue biopsies were collected per patient, which were targeted at areas of heighted perfusion and/or restricted diffusion. The sampling coordinates of each tissue biopsy were recorded using stereotactic cranial neuronavigation software. Tissue samples were sectioned, stained with H&E, and cell density measured via image analysis software in nuclei/mm^2.

Four clinical categories of imaging were obtained for every patient on a 3T MRI scanner: Anatomic or conventional, diffusion weighted, perfusion (dynamic susceptibility contrast), and permeability (dynamic contrast enhanced, DCE). After processing, a total of 23 images or parametric maps were available for each patient. A 5 mm spherical region of interest (ROI) was placed at the sampling coordinates of each biopsy and the average intensity was recorded. Corresponding "virtual biopsy" ROIs were placed in contralateral normal white matter and imaging measurements were similarly extracted. Virtual biopsies were assumed to have a cell density of 2912 ± 673 nuclei/mm^2 based on literature.

The best single image feature was selected from each family of imaging: anatomical, diffusion, perfusion, and permeability using random forest variable importance rankings. The final random forest model using this fixed four-variable set was trained to predict cell density. We evaluated the model performance using correlation between predicted and observed CD in five-fold cross validation.

Results: 23 evaluable patients with 52 image guided biopsies were included in the final analysis. The cell density among all tissue samples was 6237 ± 3552 nuclei/mm^2 (mean +/- sd). The best predictor from each family for the random forest was: T2, fractional anisotropy, cerebral blood flow, and area under time-intensity curve (from DCE). The random forest trained on these four images had the best performance with R^2 of 0.585.

Conclusions: Cellular density in biopsy samples can be estimated using a combination of conventional, and advanced MRI sequences. Being able to non-invasively estimate cell density with imaging data has the potential to guide biopsy procedures and therapeutic intervention towards areas of high cellular density to improve diagnosis and treatment for glioma patients.

Funding: Evan Gates is supported by a training fellowship from the Gulf Coast Consortia, on the NLM Training Program in Biomedical Informatics & Data Science (T15LM007093).

Measuring the Mechanical Forces During Protein Biosynthesis via EF-G Crosslinking

Gavriliuc M[1], Wang Y[1]
1. Department of Biology and Biochemistry, University of Houston, 3455 Cullen Blvd, Houston, TX 77204

Corresponding author: Dr. Yuhong Wang, Department of Biology and Biochemistry, University of Houston, 4800 Calhoun Road, Houston, Texas, E-mail: ywang60@uh.edu

The ribosome is the complex molecular machine found in all living cells that is responsible for the synthesis of all proteins. The ribosome is associated with various protein factors, including the GTPase Elongation Factor G (EF-G). EF-G is responsible for catalyzing tRNA and mRNA translocation on the ribosome, however, the mechanism of this translocation remains elusive. A recent crystallographic study has implied large conformational changes of EF-G during translocation. Previous studies observed only the elongated, post-conformational state; however, a compact, pre-translocation state has recently been seen. The question regarding the biological relevance of these conformational changes remains. To answer this, we have generated double-cysteine EF-G that is internally crosslinked, to itself, with various lengths of crosslinkers. If the large conformational change does occur in solution, then translocation will be affected by the crosslinking. To observe the effects of the crosslinking on translocation, we are measuring the force EF-G generates when crosslinked with different crosslinker lengths. Restricting the movement of EF-G will also restrict the force that can be generated by EF-G. This study will help to gain a better understanding of the mechanism of translocation and how the ribosome can efficiently translate mRNA into protein.

Funded by the National Institutes of Health grant T32GM008280 to Baylor College of Medicine

Assessment of CRM-1 Inhibitor Combinations in Ovarian and Endometrial Cancer

Handley KF[1*], Ram PT[2], Sood AK[1]

1. Department of Gynecologic Oncology and Reproductive Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas
2. Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas

*Corresponding author: Katelyn Foster Handley, MD, Department of Gynecologic Oncology and Reproductive Medicine, The University of Texas MD Anderson Cancer Center, 1881 East Road, Unit 1908, Houston, TX 77054, Email: KHandley@mdanderson.org.

The search continues for ideal drug combinations and chemotherapy adjuncts to battle ovarian and endometrial cancers.

CRM1 inhibitors are selective inhibitors of nuclear export and have demonstrated antitumor effect in ovarian cancer mouse models both as monotherapy as well as in combination with chemotherapies such as topotecan and paclitaxel. In vitro studies also establish antitumor activity of CRM1 inhibitors in endometrial cancer cell lines. We aim to assess the efficacy of a CRM-1 inhibitor, selinexor, alone and in combination with newer drugs such as bevacizumab, PARP inhibitors, and immune checkpoint inhibitors in ovarian and endometrial cancers.

We will study the antitumor effects of these drugs in combination with selinexor in orthotopic ovarian and endometrial cancer mouse models. We will also assess other potential drug combinations using high throughput drug library screening and evaluate the effects of selinexor in combination with PARP inhibitors in vitro utilizing MTT. We will then use bioinformatics to understand the mechanism behind the synergy of effective drug combinations.

Funding: KFH is supported by a training fellowship from the Gulf Coast Consortia, on the Computational Cancer Biology Training Program (CPRIT Grant No. RP170593).

Measurement Of T:G Frequency At Codon 248 Of Exon 7 Of The *TP53* Gene Using Deep Sequencing In Human Cell Lines And Tissues

Hsu W[1], Hancock KJ[2], Khanipov K[1], Pimenova M[1], Fofanov Y[1], Sowers LC[1]

1. Department of Pharmacology and Toxicology, University of Texas Medical Branch
2. Department of Surgery, University of Texas Medical Branch

Corresponding author: Lawrence Sowers, Department of Pharmacology and Toxicology, University of Texas Medical Branch, 301 University Blvd, Galveston, TX, 77555, E-mail: lasowers@utmb.edu

Colorectal cancer (CRC) is one of the most commonly diagnosed cancers with a lifetime probability of 4.7% and 5.0% in women and men, respectively. Factors believed to be important in CRC etiology include diet, exercise, and a history of colon inflammation. Among these three causes, the understanding of how inflammation leads to CRC is limited. This project aims to understand chemically on a molecular level how inflammation leads to tumor development, which can ultimately lead to rational prevention strategies.

One of the most commonly observed mutations in DNA damage events is the C to T transition, which occurs frequently at CpG sites. These transitions arise when 5-methylcytosine (5mC) is deaminated to thymine, generating a thymine:guanine (T:G) mispair. Failure to detect and repair this DNA damage intermediate leads to the formation of an irreversible T:A mutant. Mutations at methyl-CpG sites can lead to dysregulation of transcription factors. Transition mutations at CpG sites are a common event in many high-profile genes in a variety of cancers. Our study focuses on the *TP53* gene at codon 248, where amino acid variation via deamination of 5mC abrogates *TP53* DNA binding function, leading to tumorigenesis. While previous studies have detected DNA mutations at this codon in both inflamed and cancerous colon tissue, there lacks a sensitive method to detect these DNA damage events before tumor development.

In an effort to develop a method to detect C to T transition mutations using Next Generation Sequencing, we profiled the error landscape of the *TP53* exon 7 region by sequencing an oligonucleotide (Top C) and a duplex (C:G) composed of two oligonucleotides annealed together. Additionally, an experimental data set was gathered using the colorectal cell line HCT116 (**Figure 1**). All sample sets showed a remarkably similar error profile across the sequence region, with an 87% agreement in error profile pattern. Using serial dilutions of oligo-nucleotide duplexes with either a T:A or C:G at the first position of codon 248 revealed a lower limit of detection of 1:2000, which was used to validate a DNA damage signal present in HCT116 at the same position ($p < 2.2 \times 10^{-16}$).

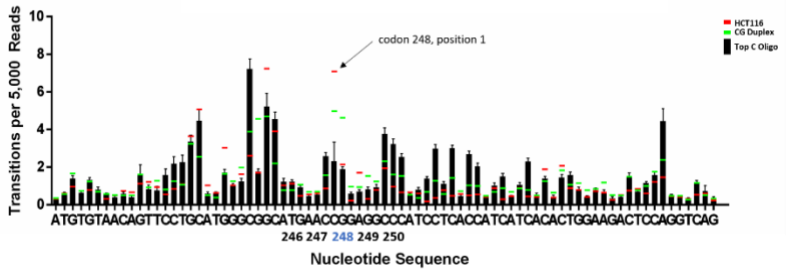


Figure 1. Averaged transition error per 5,000 reads. The bar graph represents Top C oligonucleotide base call error. The green and red line plots represent a triplicate run of a C:G duplex and a single experimental run of HCT116 colorectal cell line, respectively.

This projected was funded by CCBTP.

Clinical Phenotyping Using Semantic Knowledge Graphs

Iqbal R[1], Ken Chen1, Bernstam E[2]
1. University of Texas MD Anderson Cancer Center
2. UTHealth School of Biomedical Informatics

Corresponding Author: Ramiz Iqbal, Department of Quantitative Sciences, University of Texas MD Anderson Cancer Center, 1400 Pressler St, Houston, TX Email: riqbal@mdanderson.org

Clinically phenotyping a patient as having a particular disease is often a difficult task to conduct at scale and a persistent problem for healthcare systems. Automating the task of phenotyping is confounded by the need for deep expertise in particular domains and the terminology itself. The Unified Medical Language System (UMLS), a controlled vocabulary developed for the task of extracting biomedical concepts, has been utilized to overcome some of the aforementioned issues. However, the UMLS is known to be fraught with inconsistencies and errors which has affected the ability of researchers to use this vocabulary to automate tasks.

We propose to construct a knowledge graph where the relatedness of UMLS terms is determined by their semantic relationship in biomedical literature rather than where they are in the UMLS hierarchy. Grounding the relationship between the terms on the literature filters the UMLS for the most utilized terms for a specific disease. The approach we suggest would both help overcome the variability inherent to the UMLS terminology and the lack of deep domain knowledge needed to phenotype a patient. We will also construct the knowledge graph such that it learns from new biomedical literature and the task of phenotyping patients improves over time as well as adapts to fluctuations in the state of knowledge.

A patient would be phenotyped by comparing the UMLS terms in a patient's EHR to the knowledge graph to yield a probability of the patient having a phenotype, e.g., breast cancer. This process can be conducted across thousands of records in far less time than the gold standard of having a physician review the patient's record.

Lastly, our method is based on literature and can be expanded to phenotype other disease, therefore it is widely applicable to any healthcare system and disease.

This work is supported by the NLM Training Program in Biomedical Informatics and Data Science T15LM007093, Program Director Dr. Lydia Kavraki.

A Novel Model of Osteosarcomagenesis Reveals Dysregulation of Oxidative Phosphorylation

Jewell BE[1,2], Xu A[1], Zhou R[1,2], Zhu D[1], Lu L[3], Zhao R[1], Wang LL[3], Lee D-F[1,2]

[1] Department of Integrative Biology and Pharmacology, McGovern Medical School at UTHealth (Houston, Texas, USA)
[2] The University of Texas MD Anderson UTHealth Graduate School of Biomedical Sciences (Houston, Texas, USA)
[3]Department of Pediatrics, Baylor College of Medicine, Texas Children's Cancer and Hematology Centers, Houston, USA

Corresponding author: Brittany Ellis Jewell, Houston, Texas, Email: brittany.jewell@uth.tmc.edu

The purpose of this study is to develop a new model to study osteosarcoma in order to elucidate cellular pathways important in osteosarcomagenesis that are potential therapeutic targets. Osteosarcoma is the most common bone malignancy in childhood and early adolescence. Thirty percent of patients with the rare genetic disorder Type II Rothmund- Thomson Syndrome (RTS) develop osteosarcoma. Patients with RTS have biallelic mutations in the RECQL4 gene, which encodes an ATP dependent DNA helicase. Unfortunately, many attempts to model RTS associated osteosarcoma have not been successful. Here, we describe a patient-derived induced pluripotent stem cell (iPSC) model capable of recapitulating the process of osteosarcomagenesis. Briefly, RTS patient fibroblasts are reprogrammed to iPSCs, then subsequently differentiated to mesenchymal stem cells. These cells are validated and further differentiated to early and late osteoblasts, the potential precursors of osteosarcoma. Because the iPSC-derived osteoblasts retain the genetic fidelity of the patient from which they were derived, we believe these cells are the ideal model to study the molecular mechanism associated with RTS-associated osteosarcoma. Utilizing the patient-derived osteoblasts, we next performed RNAseq and subsequent GSEA and DEseq2 method analyses to explore the transcriptome of RTS osteoblasts and found dysregulation of oxidative phosphorylation in RTS osteoblasts. Interestingly, we have found that specific vacuolar ATPases are transcriptionally upregulated in patient cells. It has been reported that vacuolar ATPases function to increase proton concentration within the mitochondria and also within the extracellular environment. Understanding key mechanisms of osteosarcomagenesis, such as dysregulation of oxidative phosphorylation is critical to making steps toward novel therapeutics.

This research is supported by:
Gulf Coast Consortia, Training Interdisciplinary Pharmacology Scientists (TIPS) Program (NIH Grant No. T32GM120011)
UTHealth Innovation for Cancer Prevention Research Training Program Pre-doctoral Fellowship (CPRITGrant # RP160015)
Charlene Kopchick Fellowship, MD Anderson UTHealth GSBS

Quantifying Cancer Metabolic Plasticity by Coupling Gene Regulation with Metabolic Pathways

Jia D[1], Lu M[2], Jung KH[3], Park JH[3], Paudel BB[4,5], Quaranta V[4,5], Levine H[1, 6,7,*], Kaipparettu BA[3,8,*], Onuchic JN[1,9,10,11,*]

[1]Center for Theoretical Biological Physics, Rice University, Houston, TX 77005;
[2]The Jackson Laboratory, Bar Harbor, ME 04609;
[3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030;
[4]Department of Biochemistry, Vanderbilt University, Nashville, TN, 37232;
[5]Quantitative Systems Biology Center, Vanderbilt University, Nashville, TN, 37232;
[6]Department of Bioengineering, Northeastern University, Boston, MA 02115;
[7]Department of Physics, Northeastern University, Boston, MA 02115;
[8]Dan L. Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, TX 77030;
[9]Department of Biosciences, Rice University, Houston, TX 77005;
[10]Department of Physics and Astronomy, Rice University, Houston, TX 77005;
[11]Department of Chemistry, Rice University, Houston, TX 77005

*To whom correspondence should be addressed. Email: jonuchic@rice.edu (José Onuchic), kaippare@bcm.edu (Benny Abraham Kaipparettu), or h.levine@northeastern.edu (Herbert Levine).

Abstract:
Metabolic plasticity allows cancer cells to adjust their metabolic phenotypes in hostile environments. Both glycolysis and oxidative phosphorylation (OXPHOS) can be adapted by cancer cells. However, it remains largely unknown how cancer cells orchestrate different metabolic phenotypes. Current approaches mostly focus on either metabolic pathways or gene activities without addressing their extensive cross-talk. To systemically quantify cancer metabolic plasticity, we establish a theoretical framework by coupling gene regulation with metabolic pathways. We demonstrate a direct association between the activities of AMPK and HIF-1, master regulators of OXPHOS and glycolysis respectively, with the activities of three major metabolic pathways: glucose oxidation, glycolysis and fatty acid oxidation (FAO). We further demonstrate the existence of a highly aggressive hybrid metabolic phenotype in which cells use both glycolysis and OXPHOS, and a metabolically inactive phenotype that is employed by drug-tolerant cancer cells, through integrating mathematical modeling, data analysis with experiments.

Acknowledgments:
This work is supported by the National Science Foundation (NSF) grant for the Center for Theoretical Biological Physics NSF PHY-1427654, by NSF grants PHY-1605817 and CHE-1614101, by National Cancer Institute (NCI) grants R21CA173150, R21CA179720, and R03CA212816 and by the John S. Dunn Foundation Collaborative Research Award. J.N.O. is a Cancer Prevention and Research Institute of Texas (CPRIT) Scholar in Cancer Research. D.J. is supported by a training fellowship from the Gulf Coast Consortia, on the Computational Cancer Biology Training Program (CPRIT Grant No. RP170593).

Bioautographic and Molecular Docking Studies of Acetylcholinesterase Inhibitory Activity of Terpenes

Jimenez AD[1], Saleh MA[2], Anthony KP[2]

1. Department of Environmental and Interdisciplinary Sciences, Texas Southern University
2. Department of Chemistry, Texas Southern University

Corresponding author: Dr. Mahmoud Saleh, Texas Southern University, 3100 Cleburne Street, Houston, Texas, 77003. mahmoud.saleh@tsu.edu

Plant based alternative medicine have been used to combat an array of diseases such as Alzheimer disease (AD). Great interest has been displayed in plant-derivatized bioactive compounds that decrease Acetylcholinesterase (AChE) activity, which is closely related to AD. Terpenes are commonly extracted from a plant's essential oil, which are rich in bioactive compounds that are used in alternative medicine. In this study, a total of 84 terpenes were screened for 2, 2-diphenyl-1-picrylhydrazyl antiradical activity, β-carotene bleaching antioxidant, and ferric reducing power antioxidant activity. Terpenes which exhibited strong activity were then chosen to be investigated for acetylcholinesterase inhibitory activity and molecular docking studies. Nine monoterpenes and one sesquiterpene were tested for AChE inhibition by bioautographic high-performance thin layer chromatography (HPTLC) assay. Two methods were employed, diazotization and Ellman's reagent. In diazotization, terpenes that inhibited AChE appeared pale on a purple background. In Ellman's reagent, the inhibitor appeared pale in a yellow background. For this qualitative analysis, Donepezil HCl (a potent AChE inhibitor drug) was used as a positive control for identification of true AChE inhibition by the terpenes. In both methods, β-ionone, cis-jasmone, thymoquinone, and cedrol, were confirmed positive inhibitors. (-) α-Pinene, and r-carvone, were confirmed to be a negative inhibitor of the enzyme in both methods. α-Ionone was an inhibitor in diazotization method. In Ellman's reagent method, eugenol, thymol, and carvacrol, showed positive inhibition. Monoterpenoid phenols will show a false negative in diazotization method, thus Ellman's reagent method must be performed to confirm their activity. Molecular modeling was used to study the binding affinities of active terpenes with human acetylcholinesterase, this was performed using SYBYL- X- 2.1.1. Based in the docking studies, the terpenes displayed the same behavior as in the enzymatic inhibitory assay with acetylcholinesterase. (-) α-Pinene, and r-carvone did not have any bond formation and did not show inhibitory activity, while thymoquinone, β-ionone, α-ionone, cis-jasmone, eugenol, thymol, carvacrol, and cedrol, did have binding interactions, and did positively inhibit acetylcholinesterase. Docking studies showed consistency with the experimental bioautographic studies of terpenes.

This research was funded, in part, by Award Number P031B090216; Title III, Part B, Historically Black Graduate Institutions (HBGI) (CFDA No. 84.031B) and Center of Biomedical and Health Research Excellence (CBHRE) – Environmental Toxicology Core 2G12MD007605-24.

Activation of MAPK Signaling Pathway in Human Lung and Colon Cells Exposed to Triclosan

Keita D [1], Shishodia S[2]

1. Department of Environmental and Interdisciplinary Sciences, Texas Southern University
2. Department of Biology, Texas Southern University

Corresponding author: Djene Keita, Department of Environmental and Interdisciplinary Sciences, Texas Southern University, 3100 Cleburne Drive, Houston, Texas 77004,
E-mail: djene.keita@tsu.edu

Triclosan (2,4,4′-trichloro-2′-hydroxy-diphenyl ether) (TCS), is a widely used antimicrobial agent. Due to its antibacterial properties, triclosan is added in personal care products, hair products, household items, sports equipment as well as textiles and furniture, and many more. There is great potential that humans may be exposed to triclosan through oral and via contact with the lungs with constant exposure or usage of products containing this compound. There is concern about the bioaccumulation of triclosan. Pharmacokinetic studies in humans have indicated that triclosan can be absorbed from the gastrointestinal tract and skin. The presence of triclosan has been determined in human body fluids and tissues, including plasma, breast milk, urine, brain and liver tissues as well as hair and nails. With the continuous exposure to triclosan, it may produce inflammatory responses by activating pro-inflammatory signaling pathways in cells. The present study aimed to determine how triclosan, a ubiquitous compound, regulates mitogen-activated protein kinases (MAPKs) activity. The findings of this study will offer a significant contribution to the mechanism of action on the effects of triclosan exposure in the lungs when inhaled or in the gut when ingested, which may have deleterious effects. A549 and Beas-2B lung cells and HT-29 colon cells were used to study MAPK pathway. The cells were cultured, and viability determined for 24hrs, 48hrs and 72hrs by MTT and Live-Dead fluorescence assay. Expression and phosphorylation of extracellular signal-regulated protein kinases (ERKs), c-Jun amino-terminal kinase (JNK) and p38 MAPK, were investigated using Western immunoblot analysis. Nuclear translocation of nuclear factor-κB (NF-κB) p65 subunit was also examined using Western immunoblot analysis. Cytokines were measured by ELISA, and oxidative stress evaluated by microscopy. Activation of JNK, ERK, and p38 pathways by triclosan was cell type-specific. Triclosan activated JNK pathway in both lung cell types through different mechanisms. Triclosan induced phosphorylation of JNK through MKK7 in Beas-2B cells. However, activation of MKK4, not MKK7, promoted phosphorylation of JNK in A549 cells. In HT-29 colon cells, triclosan suppressed the phosphorylation of JNK. Triclosan also induced the expression of NF-κB in all cells in a time-dependent manner. Furthermore, triclosan affected the production of proinflammatory cytokines and reactive oxygen species (ROS) in all cell lines. Triclosan could potentially affect inflammation in the cells by affecting the production of proinflammatory cytokines which are mediated through activation of MAPK and NF-κB signaling pathways. Future research involves determining gene expression by measuring mRNA levels in response to triclosan.
Funding Source: This research was primarily supported by the National Science Foundation (NSF) through Texas Southern University (TSU) under the award numbers HRD-1622993, BCS-1831205, and HRD-1829184.

Neural Crest Development Is Temporally Regulated By miR-302

Keuls R[1] and Parchem R[2,3,4]

[1] Baylor College of Medicine Graduate School for Biomedical Sciences, DDMT Program, Houston, TX Rachel.Keuls@bcm.edu
[2]The Center for Cell and Gene Therapy, Baylor College of Medicine, Houston, TX USA Parchem@bcm.edu
[3]Department of Neuroscience, Baylor College of Medicine, Houston, TX USA
[4]The Stem Cells and Regenerative Medicine Center, Baylor College of Medicine, Houston, TX USA

Cranial neural crest cells are specified from ectoderm during gastrulation and possess a remarkable ability to generate diverse cell types. The mechanisms controlling the timing of neural crest formation are critical for proper development of craniofacial structures. Here, we identify two groups of miRNAs that counterpoise the transition from pluripotent ectoderm to multipotent migratory neural crest. Further, we find that miR-302 expression is maintained in migratory neural crest in mouse and is critical for regulating developmental timing of neural crest formation. Genetic rescue showed that miR-302 coordinates proper timing of neural crest formation by regulating expression of Sox9, a conserved neural crest specifier. Combined miRNA and mRNA profiling from gastrulation to the end of neurulation in mice revealed a set of genes co-targeted by multiple prevalent miRNA families. We used an unbiased screening to show that these co-targeted genes promote the formation of neural crest in chicken. Our findings reveal a post-transcriptional regulatory network that controls the balance between stemness and differentiation in neural crest cells.

Funding Sources: CPRIT, V-Foundation, B+ Foundation

miR-302 Regulates Neural Tube Closure Via Inhibition of Sonic Hedgehog Signaling

Keuls R[1], Finnell R[2,3], Parchem R[4,5,6]

[1] Baylor College of Medicine Graduate School for Biomedical Sciences, DDMT Program, Houston, TX Rachel.Keuls@bcm.edu
[2]Center for Precision Environmental Health, Baylor College of Medicine, Houston, TX Finnell@bcm.edu
[3]Departments of Molecular & Cellular Biology and Medicine, Baylor College of Medicine, Houston, TX
[4]The Center for Cell and Gene Therapy, Baylor College of Medicine, Houston, TX USA Parchem@bcm.edu
[5]Department of Neuroscience, Baylor College of Medicine, Houston, TX USA
[6]The Stem Cells and Regenerative Medicine Center, Baylor College of Medicine, Houston, TX USA

Formation of the neural tube is a complex and carefully regulated series of events involving precise orchestration of signaling, such that even slight alterations lead to catastrophic embryonic malformations, many of which result in lethality. Sonic hedgehog (Shh) signaling in the middle of the developing neural tube is well known to promote bending of the neural plate to form the neural folds, whereas inhibition of Shh along the edges of the neural plate is necessary for the convergence of the neural folds to form the neural tube. Mutants that result in increased Shh signaling exhibit neural tube closure defects (NTDs) such as exencephaly and spina bifida, which are among the most common birth defects. Here we study NTDs using the miR-302 knockout mouse which displays a completely penetrant failure of the neural tube to close. To better understand the regulation of Shh signaling during neural tube closure, we used single cell sequencing to compare transcriptomes of wildtype and miR-302 knockout embryos. We identified the Shh ligand as a predicted target of miR-302 during neural tube closure and find that cells at the middle of the neural tube ectopically express Shh, suggesting a mechanism preventing formation of dorsolateral bending. We identified misregulated Shh effector proteins, such as Msx1 that transcriptionally regulate a number of downstream target genes. Furthermore, single cell ATAC sequencing, which interrogates open chromatin regions, revealed global reorganization of enhancers and an overall increase in chromatin accessibility of genes that promote Shh signaling upon miR-302 deletion. Our findings suggest that miR-302 post-transcriptionally regulates the Shh pathway for precise patterning and closure of the neural tube.

Funding Sources: CPRIT, V-Foundation, B+ Foundation

Signal Amplification of Nonspecific DNA Hybridization Events for Compressed Sensing of Pathogens

Kota PK[1], Drezek RA[1,2], Baraniuk RG[2]
1. Department of Bioengineering, Rice University
2. Department of Electrical and Computer Engineering, Rice University

Corresponding author: Richard G. Baraniuk, Department of Electrical and Computer Engineering, Rice University, 6100 Main St., Houston, TX, richb@rice.edu

We present a novel approach for the scalable, rapid diagnostics of pathogenic infections with compressed sensing. Methods to date have either relied on highly specific sensors with limited multiplexing capacity or on complex sample amplification and sequencing. Instead, we use a small set of DNA probes bind with partial complementarity at many locations across whole bacterial genomes. Each species has a signature response based on the number of binding events of each probe which we call an *affinity vector*. Contrary to standard fingerprinting methods, we incorporate compressed sensing and sparse recovery to unmix signals composing of multiple pathogens. With our signal processing strategy, each additional probe confers exponential multiplexing capacity. To detect clinically relevant pathogen loads, we are currently developing a strategy for signal amplification of each nonspecific hybridization event through *hybridization chain reaction* (HCR). HCR uses special DNA probes that fold into a hairpin structure and can colocalize at a binding site by forming long DNA polymers. We present *in silico* results of the design of hairpin probes that synergize with our compressed sensing approach. By strategically positioning fluorophores and quenchers on the DNA probes, we will significantly increase both the separation and magnitude of pathogen's affinity vectors. These efforts will maximize our diagnostic accuracy and demonstrate potential for clinical translation. This work is supported by the NLM Training Program in Biomedical Informatics and Data Science T15LM007093, Program Director Dr. Lydia Kavraki.

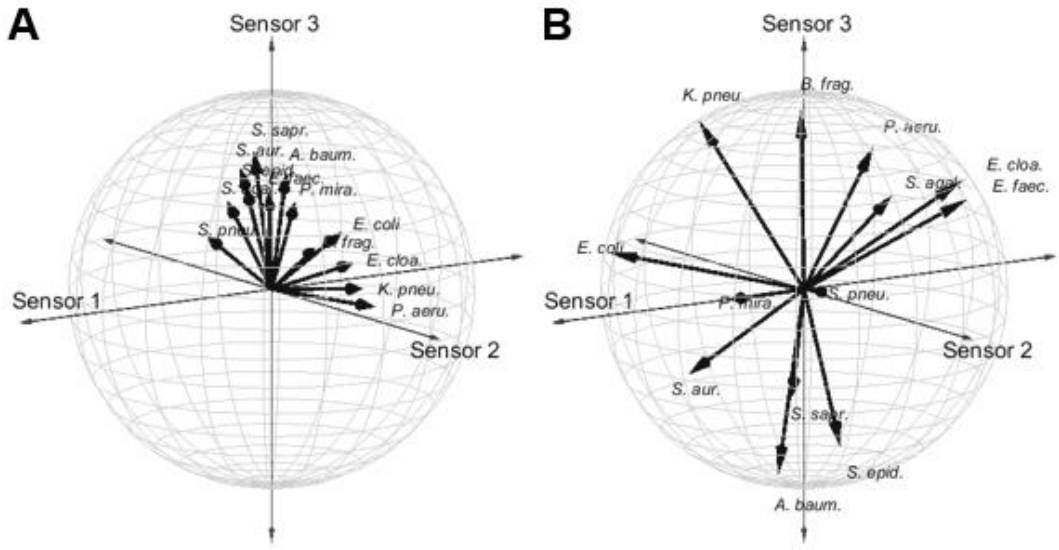


Figure: Separation of unit-normalized bacterial affinity vectors with our original (A) and proposed systems (B). We can acquire both positive and negative signals with HCR, allowing fingerprint signals to spread and enabling greater diagnostic accuracy.

Applying Graph Convolutional Neural Networks for Drug Metabolism Prediction

Litsa EE, Mitchell NE, Kavraki LE
Department of Computer Science, Rice University, Houston, TX, United States

Corresponding author: Eleni Litsa, Department of Computer Science, Rice University, 6100 Main St., Houston, TX 77005-1827, Email: el24@rice.edu

A pharmaceutical drug is a chemical compound of a specific structure that induces certain biological activity when it enters the human body. Once a chemical compound is ingested, it may interact with enzymes which can alter its structure through metabolic reactions. This process may result in reduced therapeutic action of the drug or even toxicity through the production of harmful metabolites. In order to ensure efficacy and safety, drug design studies must account for the metabolic activity of the drug, which results in a very large number of laboratory experiments. Over the years, various computational tools have been developed in order to assist this process. At the same time, intensive efforts have been made for the development of chemical databases specifically for drug-related data. Such datasets provide information regarding the structure of drugs, as well as the enzymes that interact with them and the metabolites that are formed. Such extensive chemical datasets have enabled the application of machine learning algorithms for predicting drug metabolism. Most of these efforts focus on predicting which atoms in the molecule will participate in the chemical transformation when the molecule is metabolized by a specific family of enzymes. This information gives pharmacologists insights on how to optimize the structure of the drug in order to manipulate its metabolism without compromising its therapeutic action. These existing efforts for drug metabolism prediction, though, have been hindered by the fact that chemical molecules do not have a straightforward numerical representation, which traditional machine learning algorithms require. Instead, chemical molecules are represented as graphs, where the graph nodes correspond to the atoms in the molecule and the graph edges correspond to the bonds formed between atoms. Most of the existing machine learning-based approaches rely on expert knowledge to engineer features that represent either the individual atoms within the molecule or the entire molecule, depending on the prediction task.

In this project, we are exploring Graph Convolutional Neural Networks (GCNNs) for predicting drug metabolism, in order to automatically derive task-specific atom and molecule representations. We divide the drug metabolism prediction problem into two prediction subtasks: First, we are predicting which enzymes are more likely to metabolize a given molecule. Second, we are predicting which atoms are more likely to be involved in an enzymatic reaction. For each prediction task, we are using GCNNs to learn task-specific graph embeddings. A GCNN is a neural network-based architecture that learns vector representations either at a graph or at a node level based on the structure of the graph and the prediction task. We are utilizing the graph level embeddings to predict which enzymes will more likely metabolize a given molecule. We are also learning node level embeddings to predict whether a given atom can be involved in a metabolic transformation by a given enzyme. These embeddings have been obtained by training the GCNN on two separate datasets. For the enzyme prediction task, we have constructed a dataset by collecting information about enzymatic interactions from publicly available datasets. For the prediction of reacting atoms, we are using a standard dataset that has been used in the literature for developing and testing computational methods that predict reacting atoms in metabolizing drugs. Preliminary results show that our method for predicting the atoms involved in metabolic reactions can fairly compete against existing approaches without the need for expert-engineered features. Coupled with the model that predicts which enzymes can metabolize a chemical molecule, we aim to develop a computational tool that provides a more comprehensive study of drug metabolism in order to ensure safety and efficacy of drugs.

Acknowledgements: This work has been supported by funds from Rice University and CPRIT RP170508.

3D Label-free, Real-time, Chemistry-Sensitive Imaging to Identify Parathyroid Adenoma and Hyperplasia and Classify Parathyroid Glands and Recurrent Laryngeal Nerve During Surgery

Li J[1], Wang Y[1,2], He Y[1], Farach-Carson MC[3] and Wong S[1]
1. Translational Biophotonics Laboratory, Department of Systems Medicine and Bioengineering, Houston Methodist Cancer Center, Weill Cornell Medicine
2. Department of Breast-thyroid-vascular Surgery, Shanghai General Hospital, Shanghai Jiao Tong University
3. Diagnostic and Biomedical Sciences, UT Health Science Center at Houston
Corresponding Author: Stephen Wong, Translational Biophotonics Laboratory, Department of Systems Medicine and Bioengineering, Houston Methodist Cancer Center, Weill Cornell Medicine. 6670 Bertner Ave. R6-211, Houston, TX, 77030. Email:stwong@houstonmethodist.org.

Background: Accidental excision or damage to a healthy parathyroid gland (PG) during thyroidectomy and the inability to identify diseased PGs during parathyroidectomy can result in postsurgical complications. Meanwhile, hyperparathyroidism includes primary and secondary hyperparathyroidism, which can cause damage in skeleton, kidney, and heart, with bone pain and renal colic. About 85% of primary hyperparathyroidism patients have only a single adenoma, which is critical to preserve healthy parathyroid during surgery. Therefore, it is critical to distinguish healthy parathyroid when the operation is being processed since the extent of the parathyroidectomy depends on the accuracy of the diagnosis.
Objective: To evaluate the capability of combining 3-dimentional label-free Coherent Anti-Stokes Raman Scattering (CARS) imaging with a deep-learning algorithm to diagnose thyroid cancer and to distinguish a parathyroid gland and nearby recurrent laryngeal nerve (RLN) from surrounding tissues.
Methods: Human cancerous and healthy thyroid tissue were obtained from Shanghai General Hospital. Porcine thyroid gland, RLN, adipose tissue, muscle tissue, lymph nodes, and parathyroid glands were harvested in fresh. Each sample was divided into two sections, one for CARS imaging and the other for hematoxylin and eosin staining (H&E). The H&E images were then classified by an experienced pathologist. PAX-8 immunohistochemistry and parathyroid hormone (PTH) tests were performed to further confirm the identification of parathyroid tissue. Some imaging analysis methods were performed to quantify some parameters, such as overall normalized intensity, regional normalized intensity, etc., to quantitatively differentiate the six types of tissues. Deep-learning algorithms are developed for auto-diagnosing thyroid cancer and for differentiating parathyroid, thyroid, lymph node, fat, RLN, and muscle.
Results: Human thyroid cancer tissue can be identified immediately and automatically through CARS imaging combined with deep-learning algorithms. Parathyroid adenoma/hyperplasia presented as enlarged chief cells and clear cells with reduced fat droplets compared with normal parathyroid. The CARS identification results were confirmed by an experienced pathologist. All six different porcine tissues can be identified by the same system, and the correspondence to human equivalents was confirmed by the pathologist co-investigator. The accuracy of both identification is above 98%. Structurally, multiple pool-like structures surrounded or partially filled by chief cells with characteristics of round, smaller size, and evenly distributed nuclei can be seen in the CARS image of fresh porcine parathyroid gland. Bright and wide straps are presented in CARS image of a fresh porcine nerve because nerve bundles contain rich lipid membranes. Quantitative analysis was performed to differentiate parathyroid, thyroid, lymph node and fat, and differentiate RLN from muscle.
Significance: The label free imaging system combining CARS imaging and deep-learning algorithm can reliably identify thyroid cancer tissue and differentiate it from thyroid gland, lymph node, and adipose tissue morphologically, without staining, accurately, and immediately. Therefore, this system is a potential tool for quantitatively differentiating parathyroid glands and RLN from surrounding tissues and identifying thyroid cancer tissue in real-time for diagnosis and during parathyroid surgery. Ultimately, the imaging technique can be incorporated into handheld probe for use in clinical diagnosis, such as cancer margin detection during surgery for minimal dissection and reducing nerve impairment during prostate cancer surgery for better outcome.

Integrating Spatial Features into FTIR Histological Image Classification

Lotfollahi M, Berisha S, Daeinejad D, Mayerich D

University of Houston

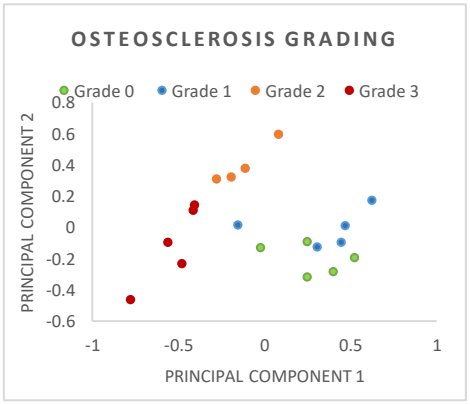
Current methods for cancer detection rely on tissue biopsy, chemical labeling/staining, and examination of the tissue by a pathologist. Though these methods remain the gold standard, they are non-quantitative, susceptible to human error, and can only capture highly specific molecular content from the sample. Fourier transform infrared (FTIR) spectroscopic imaging has shown some potential as a quantitative alternative to traditional histology. However, identification of histological components requires reliable classification based on molecular spectra, which are susceptible to artifacts introduced by noise and scattering. In addition, infrared spectroscopy provides significantly lower resolution than what is available with traditional bright-field or fluorescent imaging. As a result, several tissue types - particularly in heterogeneous tissue regions - confound traditional classification methods.

By taking advantage of the spatial context of an infrared image, classification results can be significantly improved. Convolutional neural networks (CNNs) are the current state-of-the-art in image classification, providing the ability to learn spatial characteristics by training a series of correlational filters. In this work, we demonstrate that CNNs with architectures designed to process both spectral and spatial information can significantly improve classifier performance over per-pixel spectral classification using the best models. We report classification results after applying CNNs to data from tissue microarrays (TMAs) to identify histologically relevant tissue components, including those that have limited chemical signal, such as adipocytes. Experimental results show that the use of spatial information in addition to the spectral information brings significant improvements in the classifier performance. The work demonstrates the application and efficiency of deep learning algorithms in improving the diagnostic techniques in clinical and research activities related to cancer.

Comprehensive Evaluation of Myelofibrosis in Bone Marrow Using Infrared Imaging

Mankar R[1], Bueso-Ramos CE[2], Yin CC[2], Mayerich D[1], Reddy R[1]
1. Department of Electrical and Computer Engineering, University of Houston, Houston, USA
2. Department of Hematopathology, University of Texas MD Anderson Cancer Center, Houston, USA

Osteosclerosis and myefibrosis (MF) are complications of myeloproliferative neoplasms (MPN). These disorders result in excess growth of trabecular bone and collagen fibers that replace hematopoietic cells, resulting in abnormal bone marrow function. Treatments using imatinib and JAK2 pathway inhibitors can be effective on osteosclerosis and fibrosis, therefore accurate grading is critical for tracking treatment effectiveness. Current grading standards use a four-class system based on analysis of biopsies stained with three histological stains: hematoxylin and eosin (H&E), Masson's trichrome, and reticulin. However, conventional grading can be subjective and imprecise, impacting the effectiveness of treatment. In this paper, we demonstrate that mid-infrared spectroscopic imaging may serve as a quantitative diagnostic tool for quantitatively tracking disease progression and response to treatment. The proposed approach is label-free and provides automated quantitative analysis of osteosclerosis and collagen fibrosis.



For accurately characterizing of osteosclerosis and collagen deposition, two adjacent tissue sections from a biopsy are obtained from each of the 20 patients in the study. One of the section from each patient is stained with chemical stains, with H&E stain for osteosclerosis and with Masson's trichrome stain for collagen deposition to generate ground truth for spectral classifier training. The other section is imaged using Fourier transform IR imaging and then this hyperspectral image is preprocessed to compensate for scattering artifacts in the imaging before data analysis using machine learning algorithms. For both types of grading, Spectral classifier (RandomForest) is used for classification of histology classes using spectral information from hyperspectral images of tissue biopsies. Spectral classification is performed at the pixel level where each pixel is a vector which represents mid-IR spectrum. With this trained classifier, hyperspectral images of clinical biopsies were classified into histology classes. Results from spectral classifier are then used for morphological analysis using linear discriminant analysis (LDA) for grading osteosclerosis and collagen deposition in BM biopsies. This classifier uses morphological/spatial features which are extracted from 2-dimensional images obtained by performing classification for histology classes

This work shows the potential for quantitative bone biopsy grading using infrared spectroscopic imaging. Concurrence between pathologists is between 89.4% - 94.9\% for osteosclerosis grading and between 84.6% - 91.3% for collagen deposition grading. The proposed method achieves accuracies of 84.4% for osteosclerosis grading using 20 samples and accuracy for collagen deposition grading is 50%, which increases to 89.6% when considering only grades 0/1 and 2/3. We believe that higher resolution would enable a better differentiation because pathologists rely on high-resolution collagen features to differentiate between these grades. Spatial resolution introduces a critical limitation of infrared spectroscopic imaging for histological applications. However, newly emerging optimal photothermal IR (O-PTIR) imaging technique can be used to overcome spatial resolution limitation and achieve better results for MF grading.

Corresponding author: Rupali Mankar, Email: rupali.mankar@gmail.com
This work is supported by the NLM Training Program in Biomedical Informatics and Data Science T15LM007093, Program Director Dr. Lydia Kavraki

Identifying a Cohort of Patients Living with HIV from a Large Electronic Health Record Database

May SB[1], Giordano TP [2], Gottlieb A[1]
[1] School of Biomedical Informatics, University of Texas Health Science Center – Houston
[2] Department of Medicine – Infectious Disease, Baylor College of Medicine

Corresponding Author: Sarah B. May, School of Biomedical Informatics, University of Texas Health Science Center – Houston, 7000 Fannin, Suite 600, Houston, TX, sarah.b.may@uth.tmc.edu

Introduction: Currently, there are an estimated 1.1 million people living with HIV (PLWH) in the US, 15% of whom are not aware of their HIV status. Identifying these individuals and linking them to HIV care is an important step for reducing the number of new transmissions. Another important tool for HIV prevention is pre-exposure prophylaxis (PrEP), antiretroviral medication taken by uninfected individuals. PrEP significantly reduces HIV incidence in high risk populations, however, of the 1.1 million people with indications for PrEP according to CDC guidelines, only 7% have been prescribed the medications. Data from electronic health records (EHRs) provides a rich resource for building predictive models identifying patients at high risk for health outcomes, such as HIV infection. The first step to developing these tools is to accurately and reliably identify cohorts of patients with the outcome of interest from the data. The goal of this study was to develop an algorithm for identifying a cohort of PLWH from a large EHR database that can then be used to develop models to predict which patients are at highest risk for HIV infection and are good candidates for prevention strategies such as PrEP.

Methods: For this study, we used the Cerner HealthFacts database, a de-identified EHR database containing records for 68 million unique patients from over 600 participating hospitals and clinics in the United States. Information on ICD-9 and ICD-10 codes for HIV and HIV-related comorbidities were extracted, along with HIV screening and confirmatory tests, HIV viral load (VL) measurements, and antiretroviral medications. Patients were required to have at least one instance of an ICD-9/10 code for HIV *and* a positive confirmatory lab test *or* a prescription for HIV antiretroviral medication to be confirmed as diagnosed with HIV. Once the cohort was identified, descriptive analyses were performed to determine the demographic and clinical characteristics of the patients and provide insight into the validity of the cohort.

Results: We identified 109,960 patients with an ICD-9/10 code for HIV in the Cerner HealthFacts database, of which 45,027 met the criteria for a confirmed HIV diagnosis. Of those with confirmed HIV, 20,399 (45%) were confirmed with a positive confirmatory lab test (western blot or HIV VL >1000 copies/mL), 15,745 were confirmed because they had a prescription for HIV antiretroviral medications, and 8,879 (20%) met both requirements. The final HIV-positive cohort was 69% male and 53% African American. Sixty-eight percent attended a clinic in an urban area, and 42% had some form of public health insurance (i.e. Medicaid/Medicare). They had a median of 21 (IQR 7-47) encounters per patient, and an average of 4.3 ± 3.25 years of follow-up time in the database.

Conclusion: We were able to develop and implement an algorithm to identify a cohort of PLWH from a large, national EHR database. Our next steps are to further refine the algorithm by modifying it to begin with positive HIV screening tests rather than ICD codes, as well as developing methods for separating the cohort into incident and prevalent HIV cases. This will be necessary to achieve our overarching goal of developing risk prediction models for HIV infection. These models will provide guidance to clinicians when ordering repeat HIV screening tests, as well as, identify patients who are candidates for PrEP and facilitate initiation on these medications.

This work is supported by the NLM Training Program in Biomedical Informatics and Data Science T15LM007093, Program Director Dr. Lydia Kavraki.

Targeting the *Leishmania* Hsp100 N Domain to Prevent Infective Parasite Stage Differentiation

Mercado JM[1,2], Lee J[2], Nuri Sung[2], The-Poot CF[3], Villanueva-Lizama LE[3,4,5] Bottazzi ME[4,5,6], Kathryn Jones[4,6], Lee S[2], Cruz-Chan VJ[3,4,5], Tsai FTF[2,6*]

1. Department of Molecular and Cellular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA.
2. Department of Biochemistry and Molecular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA.
3. Dr. Hideyo Noguchi Regional Research Center, Autonomous University of the Yucatan, Mérida, Yuc 97000, MX.
4. Texas Children's Hospital Center for Vaccine Development, Houston, TX 77030, USA
5. Department of Pediatrics, National School of Tropical Medicine, Baylor College of Medicine, Houston, TX 77030, USA
6. Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston, TX 77030, USA.

*Corresponding author: Francis T.F. Tsai, Department of Biochemistry, Baylor College of Medicine, 1 Baylor Plz, Houston, Texas 77030, ftsai@bcm.edu.

Leishmaniasis is a vector-borne lesion forming disease endemic in many developing countries caused by the *Leishmania* parasite. *Leishmania spp.* are dimorphic, protozoan parasites that spread to mammalian hosts during the blood meals of infected female sandflies. *Leishmania* exist as either non-infective promastigotes in the sandfly gut or infectious amastigotes that are responsible for the clinical manifestation of Leishmaniasis. Promastigote-to-amastigote differentiation is triggered during mammalian infection and is accompanied by *Leishmania* Hsp100 chaperone expression. *Leishmania* Hsp100 is a member of the Hsp100 family of protein foldases and is particularly similar to the ClpB/Hsp104 subfamily of disaggregases. Hsp104 chaperones are absent in animal cells making *Leishmania* Hsp100 a potential target for the treatment of Leishmaniasis. *Leishmania* Hsp100 is a multi-domain protein consisting of an N domain, an M domain, and two AAA+ domains. We recently showed the N domain of yeast Hsp100 contributes towards the recovery of proteins from an aggregated state where mutations in the N domain abolish chaperone activity. Therefore, I hypothesize that *Leishmania* Hsp100 activity can be depleted by small molecule inhibitors targeting the N domain that block *Leishmania* stage differentiation. To exploit *Leishmania* Hsp100 for drug design, I determined the crystal structure of the *L. mexicana* Hsp100 N domain at 1.4 Å resolution and developed an assay to assess for N domain function. The outcome of my work may be used to combat Leishmaniasis and related human infections caused by trypanosome parasites.

Evolutionary Action Through Machine Learning Methods to Discover Driver Genes in Cancer Somatic Mutations

Parvandeh S[1], Katsonis P[1], Downhower LA[1], Lichtarge O[1]

1.Department of Molecular & Human Genetics, Baylor College of Medicine

Corresponding Author: Olivier Lichtarge, Department of Molecular & Human Genetics, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas, E-mail: lichtarg@bcm.edu

The genotype-phenotype relationship determines how mutations affect health in the short term and evolution in the long term. Understanding this relationship is an important challenge in biology and biomedical research. Clinicians typically aim to study the genotype-phenotype relationship with simple statistics, which limits their analyses. These studies will be substantially improved by advanced statistics and machine learning in broad, integrative association studies that mine high-throughput data to link traits, phenotypes, or diseases with specific candidate driver genes that drive the character. Here, we propose an approach to uncover the genotype-phenotype relationship using a machine learning method, EPIstasis MUTations ESTimatoR (EPIMUTESTR) and the evolutionary action (EA). EA uses evolutionary information to score the functional impact of any somatic mutation. Briefly, EPIMUTESTR applies the k nearest neighbor (k-NN) algorithm to identify which set of genes best distinguish cases from controls. It selects a random sample from a mixture of cases and controls and finds their k nearest neighbors based on the Euclidean distance adapted to EA. Considering whether these neighbors are cases or controls samples, and the difference of the EA scores, the weight of each gene is adjusted accordingly. In the context of cancer, we were able to consider, in multiple steps, multiple genes. The input data were obtained from the Cancer Genome Atlas (TCGA) whole exome sequencing of somatic mutations for 33 cancer types. We simulated controls by generating a set of random nucleotide mutations in each gene. We identified 387 cancer candidate genes with high precision compared to other available methods. We assessed the output by the recovery of known cancer genes, according to other methods, and comparison to gold standards such as COSMIC and experimental screens for oncogenes using DepMap. The outcome suggests new cancer genes and pathways for patient risk and therapy stratification.

This project is under the CPRIT grant number RP170593.

Study of a Heteromeric Kainate Receptor Gluk2/k5 By Probing Single-Molecule FRET

Paudyal N, Litwin DB, Berka V, Flores EC, Jayaraman V

Center for Membrane Biology
University of Texas Health Science Center at Houston, 6431 Fannin St, Houston, TX,
Nabina.Paudyal@uth.tmc.edu

Kainate receptors are ligand-gated ion channels that belong to the ionotropic glutamate receptor family. They are unique among glutamate receptors in being involved in both excitatory and inhibitory synaptic transmission. These receptors are homomeric and heteromeric assemblies of subunits GluK1-5 arranged as dimer of dimers. Each subunit consists of an extracellular amino-terminal domain, a ligand binding domain, a transmembrane domain and an intracellular C-terminal domain. Despite heteromeric GluK2/K5 receptors being the most predominantly expressed kainate receptor in the brain, there are currently no structures of the full length GluK2/K5. We have used smFRET measurements to determine the specific organization of the GluK2/GluK5 within the tetramer and additionally studied the structural changes associated with agonist binding, activation, and desensitization of the receptor. These results of the heteromeric receptors are compared to the homomeric receptors for which there are full length structures available.

HAMBP Fellowship (*National Institutes of Health grant T32GM008280 [NP])*

Using Geographic Information Systems (GIS) and Data Science to Uncover Important Factors for Health-driven Resiliency Plans: Preparing Houston for Future Disasters

Pecha J[1], Main A[2], Rogers S[2], Oluyomi A[1]
1. Environmental Health Service, Department of Medicine, Baylor College of Medicine
2. Gerald D. Hines College of Architecture, University of Houston

Corresponding author: Joseph Pecha, MS2, c/o Environmental Health Service, Baylor College of Medicine, One Baylor Plaza, Houston, TX, 77030, E-mail: joseph.pecha@bcm.edu

Analysis of spatial data through the use of geographic information systems (GIS) offers a powerful way to quantify damage and target resources efficiently after a natural disaster.[1] In late August 2017, Hurricane Harvey poured a record-setting 60.5 inches of rain on the Houston area causing 88 deaths, $125 billion in damage, and the destruction of 204,000 homes.[2] Northeast Houston was particularly hard-hit and much of the area ranks in the top fifth for social vulnerability according to the city's Housing and Community Development Department.[3] Preventing future damage requires thinking ahead. Resiliency plans offer a solution by appraising the resources and vulnerabilities endemic to each area.[4] Understanding community networks, communication, health, and preparedness are necessary for communities to quickly recover from future disasters.[5, 6]

This study focused on several neighborhoods in northeast Houston (Aldine, East Aldine, Trinity/Houston Gardens, Kashmere Gardens, Greater Greenspoint, Eastex-Jensen, and East Houston). The aim of the study was to use only publicly available data to visually characterize these areas with respect to social determinants of health and conduct geospatial analysis to help allocate resources effectively. Using GIS, basic demographic variables (e.g. income status, population density); social determinant variables (e.g. number of Federally Qualified Health Centers (FQCHs), percentage uninsured); and health-specific variables (e.g. mortality, birthweight) were mapped for visual comparison across Harris County. A secondary aim was to compare the prevalence of these measures (sociodemographic, health determinants, and health outcomes) between the neighborhoods and the rest of Harris County using the statistical protocols appropriate for the data type (continuous versus categorical).

Results showed that these neighborhoods had among the highest poverty rates in the county, the lowest high school graduation rates, uninsurance rates over 25%, and, in neighborhoods west of I-69, very limited English proficiency. East of I-69, the neighborhoods had a higher number of FQHCs, but statistically less people—including fewer uninsured people—were served by them. Mortality rates were high in neighborhoods east of the highway, but average in those further west. Overall, one-third to one-half the population of these neighborhoods had no usual source of healthcare and more people than elsewhere in Harris County delayed medical care due to cost. Geospatial analysis showed that much of Greater Greenspoint, Aldine, and some of East Aldine could not access a healthcare center within a 15-minute driving distance.

This study provoked many questions: Why is accessing care in this area so difficult? Why is mortality lower where there is a greater deficiency in healthcare access? Why do many people near an FQHC report "no usual source of care"? Compared to the rest of Harris County, these neighborhoods are particularly disadvantaged and suffer from a higher disease burden. GIS methodology allowed us to integrate environmental and data science to advance our understanding of issues related to natural disasters and disparities in health determinants, health care access, and health outcomes. These findings may help community stakeholders (e.g., political and business leaders, researchers, and local residents) create realistic health-focused resiliency plans for these neighborhoods.

Works Cited:

1. Holt, James B et al. “Use of BRFSS data and GIS technology for rapid public health response during natural disasters.” *Preventing chronic disease* vol. 5,3 (2008): A97.
2. Amadeo, Kimberly. “Hurricane Harvey Shows How Climate Change Can Impact the Economy.” The Balance, The Balance, 25 June 2019, www.thebalance.com/hurricane-harvey-facts-damage-costs-4150087.
3. “City of Houston - Recovery from Hurricane Harvey.” *City of Houston HCDD*, recovery.houstontx.gov/.
4. Kafle, Shesh.. “Measuring disaster-resilient communities: a case study of coastal communities in Indonesia.” *Journal of Business Bontinuity & Emergency Planning*. 5. 316-26, 2012.
5. Norris, Fran H., et al. “Community Resilience as a Metaphor, Theory, Set of Capacities, and Strategy for Disaster Readiness.” *American Journal of Community Psychology*, vol. 41, no. 1-2, 2007, pp. 127–150., doi:10.1007/s10464-007-9156-6.
6. Sandifer, Paul A, and Ann Hayward Walker. “Enhancing Disaster Resilience by Reducing Stress-Associated Health Impacts.” *Frontiers in public health* vol. 6 373. 21 Dec. 2018, doi:10.3389/fpubh.2018.00373

Structural *In Silico* Analysis of Post-Translational Modifications Present on Dnak Protein

Rigo MM[1], Borges TJ[2], Murshid A[3], Nitika[4], Wolfgeher D[5], Calderwood S[3], Truman A[4], Bonorino C[6]
1. School of Medicine, Pontificia Universidade Catolica do Rio Grande do Sul
2. Schuster Family Transplantation Research Center, Harvard Medical School
3. Department of Radiation Oncology, Harvard Medical School
4. Department of Biological Sciences, The University of North Carolina at Charlotte,
5 Department of Molecular Genetics and Cell Biology, The University of Chicago
6 Department of Surgery, University of California at San Diego

Corresponding author: Mauricio Menegatti Rigo, School of Medicine, Pontificia Universidade Catolica do Rio Grande do Sul, Av. Ipiranga, 6681, Rio Grande do Sul, Porto Alegre, Brazil, E-mail: mauricio.rigo@pucrs.br

Molecular chaperones known as Heat Shock Proteins (HSPs) are key proteins involved in a wide range of functionalities such as protein synthesis, folding, disaggregation, and degradation. One of the first chaperones to be described was the Hsp70 protein. The Hsp70 has an important role assisting substrate folding and regulating the activities of other proteins. It has been shown that post-translational modifications (PTMs) are often present in Hsp70 of several organisms, which can influence its behavior in respect to the environment and the interaction with molecular partners. Until now, a wide range of PTMs were already described for these chaperones, such as acetylation, phosphorylation, glycosilation, sumoylation and others. In this way, a PTM works as a key mechanism to add variability to what is generated from genomic information. Data on the extracellular effects of heat shock proteins are still controversial. **Objective:** We hypothesized that at least some of the inconsistencies observed in different studies on extracellular functions of Hsp70 could be due to PTMs. Here, we aimed to model DnaK protein, the Hsp70 chaperone from *Mycobacterium tuberculosis* (UniProt Entry: P9WMJ9), and analyze a set of PTMs present on the protein surface through an *in silico* approach. **Methods:** We produced DnaK protein in two different expression systems (*E. coli* and *P. pastoris*), retrieving the data from common PTMs (acetylation and phosphorylation) through mass spectrometry analysis. Since the 3D structure of DnaK was not resolved, we performed a homology modeling of the protein using Modeller software. The C-terminal region of the protein, also known as "lid" region, was modeled using an *ab initio* approach using the Robetta server. The final structure was energy minimized with GROMACS and validated though several programs (Verify 3D, ERRAT, ModEval and ProQ). The final model was modified with the information of PTMs using PyTMs plugin from PyMol program. The models with or without PTMs were qualitatively compared regarding topography and electrostatic potential. **Results:** The PTMs occurred on different residues such as lysine, tyrosine, serine and threonine. These residues were mainly exposed on the protein surface and the electrostatic map shows important charge differences in the PTMs added to the protein when produced in different expression systems. The introduction of acetylation promoted a more negative state of the DnaK, and this negative effect was more pronounced on *P. pastoris* expression system. It is interesting to note that some of the PTMs were conserved between both *E. coli* and *P. pastoris* expression system, mainly on nucleotide (NBD) and substrate binding domains (SBD), suggesting a possible role for these PTMs on the properly DnaK function. **Conclusion:** Overall, each PTM could influence the dynamics of the protein, as well as the interaction affinity with other molecules. Depending on the expression system, the "chaperone code" can change.
Funding: CAPES

Structure and Activation Mechanisms in a Group III Metabotropic Glutamate Receptor

Rosario-S JI[1], Wang Z[1,2], Wensel TG[1,2]
1. Department of Quantitative and Computational Biosciences, Baylor College of Medicine
2. The Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine

The metabotropic glutamate receptor type 6 (mGluR6) expresses exclusively in ON-bipolar cells (ON-BCs) in the retina and transmits light information from photoreceptors to ON-BCs. Dysfunction of this receptor causes permanent congenital stationary night blindness (CSNB) in patients, and can present other retinal and refractive phenotypes. mGluR6 belongs to the group III mGluRs, and until now there is no full-length structure yet from any group III mGluR, and no structures even of mGluR6 fragments. The mechanism of action by which mGluR6 modulates the light transduction pathway remains to be defined. The central hypothesis of this proposal is that glutamate binding to the mGluR6 extracellular ligand-binding domain induces, through long-range allosteric mechanisms, conformational changes at the cytoplasmic face that rapidly expose a structural interface required for G protein action. The research proposed will allow to complete the following Specific Aims:

1. Determination of mGluR6 specificity, allosteric properties and mechanisms of action: To test the hypothesis that the allosteric interactions by which mGluR6 relays information across its domains are specific for mGluR6, based on its unique function and localization, high-resolution structures of mGluR6 in apo and ligand-bound states will be determined using electron cryo-microscopy single particle reconstruction (cryo-EM SPR). mGluR6 will be purified and reconstituted in lipid nanodiscs and confirmed to be embedded in a functional state before electron imaging.

Specific Aim 2. Elucidate mechanisms of G protein recruitment and activation: mGluR6 interacts with the $G_{i/o}$ protein family to initiate signaling pathways, but the conformational changes that translate agonist binding extracellularly into rearrangements that enable G protein coupling and activation remain uncharacterized. The complex with mGluR6 and the heterotrimeric G protein will be solved to sub-nanometer resolution using cryo-EM SPR. Ligand binding and G protein specificity assays will help determine the parameters that confer most plausibility and structural stability to this protein complex. Single molecule Förster resonance energy transfer (smFRET) will measure the distribution of a range of conformational states and the kinetics of transitions between them, for mGluR6 and other group III mGluRs.

Accomplishing these aims and delivering a comprehensive model of mGluR6 mechanism of action and dynamics is essential for understanding light transduction physiology and pathology mechanisms, and designing therapies.

Funding: Houston Area Molecular Biophysics Program (T32-GM008280-31)

Baylor College of Medicine Medical Scientist Training Program

Deep Quantitative Analysis of Multiplexed ImmunoHistoChemstry (IHC) Images of Alzheimer's Rat Brain Tissue

Saadatifard L[1], Jahanipour J[1], Roysam B[1]

1.Department of Electrical and Computer Engineering, University of Houston

Corresponding author: Leila Saadatifard, Electrical and Computer Engineering, University of Houston, 4800 Calhoun Rd, Houston, TX 77004, E-mail: lsaadati@central.uh.edu

Alzheimer's disease (AD) is a neurodegenerative brain disease primarily affecting elderly individuals. Currently, more than 5 million Americans have been diagnosed with AD and this number is expected to reach 20 million by 2050. AD manifests as a progressive decline in cognitive function leading to significant morbidity and eventual death. Although the cognitive decline is associated with degeneration of neurons, it is important to recognize that a neuron's health is defined by the health of its cellular neighborhood composed of glia and vasculature. The neuronal neighborhood and vascular network are main keys for studying neuronal degeneration and brain functions that are important to understand neurodegenerative diseases such as Alzheimer's.

In this work, we developed a pipeline for comprehensive profiling of progressive alterations to different brain cell types over time, with a particular focus on quantifying alterations to the cellular neighborhoods of neurons. A multiplexed fluorescence imaging method generates high resolution images of the whole brain tissue for different cell types (Table 1).

Channel number	Fluorophore	Biomarker
1	DAPI	DAPI
2	AF488	S100beta astrocytes
3	AF594	Abeta, plaques
4	AF647	NeuN neurons
5	AF790/800	Iba1 microclimate

Table 1. Antigens and Fluorophores used for multiplex imaging

We integrated an automated pipeline for extracting the specific fluorescent signals of interest while rejecting non-specific intra-channel autofluorescence and cross-channel bleed-through, and compensating for non-uniform illumination and photobleaching. They will then be aligned with single-pixel accuracy to enable montaging and image data accumulation across rounds, while masking out the occasional tissue folds and tears that may arise during the imaging.

We trained an automated deep convolutional neural network for reliable and large-scale cell detection, segmentation, and phenotypic classification. This network will be trained to detect and delineate cell nuclei with >95% accuracy, and it will analyze the association of molecular markers of cell type and cell status for each cell, to precisely identify the type and phenotypic status of each cell. For identifying the phenotypic status of cells, we will train an efficient variant of Hinton's capsule network which is capable of analyzing object-part relationships (cells and their component parts), and able to outperform state-of-the-art deep networks when the number of training samples is small. All classifiers will be validated against manually established reference datasets.

This work is supported by the NLM Training Program in Biomedical Informatics and Data Science T15LM007093, Program Director Dr. Lydia Kavraki

Mechanism of Neutralization of Human GII Noroviruses by Cross-reactive Human IgA

Salmen W[1], Hu L[2], Alvarado G[3], Atmar R[1], Estes M[1], Crowe J[3], Prasad B[1,2]

1. Department of Molecular Virology and Microbiology, Baylor College of Medicine
2. Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine
3. Department of Pathology, Microbiology, and Immunology, Vanderbilt University Medical Center

Corresponding Author: B.V.V. Prasad, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, 1 Baylor Plaza, Houston, Tx Email: vprasad@bcm.edu

Human noroviruses cause approximately 685 million cases of acute gastroenteritis and are responsible for an estimated 50,000 deaths worldwide in children under the age of five. Since 2002, the predominating circulating strain of norovirus is the rapidly evolving GII.4 strain. Until recently, there was little understanding of the mechanism by which the host mounts antibody-mediated neutralization against norovirus. Recent studies have identified a panel of human monoclonal antibodies (mAbs) from patients previously infected with a GII.4 strain. Within this panel, the neutralizing IgA antibody, NORO-320 was identified, which cross-reacts with viruses from several GII genotypes including GII.3, GII.4, GII.6, GII.12, and GII.17. To understand the mechanism by which NORO-320 mediates neutralization, we determined the crystal structure of the protruding domain (P-domain) of the GII.4 capsid protein, VP1, in complex with NORO-320 antigen-binding fragment (Fab). This crystal structure revealed that the Fab binds nearly perpendicular to the P-domain dimer near a region that is close the shell domain, which is responsible for capsid assembly and integrity. Sequence analysis indicates that the NORO-320 binding site in the P-domain is conserved in viruses of the GII.3, GII.4, GII.6, GII.12, and GII.17 genotypes, thereby providing a structural basis for the observed cross-reactivity of NORO-320. As this epitope is distant from the characterized histo-blood group antigen (HBGA) binding site, the mechanism of neutralization is not by directly blocking the receptor interaction. Our crystallography data suggests that the close proximity of the antibody binding site to the shell domain may facilitate neutralization by disruption of capsid integrity. To test this hypothesis, we are performing negative-stain of NORO-320 Fab in complex with virus-like particles of GII.4 to observe particle integrity. This study will provide insight into the mechanism by which the human adaptive immune system can elicit broadly neutralizing antibodies against a rapidly evolving virus.

Poster #48

Tissue Cross-Talk In Longevity Regulation

Savini M[1], Wang MC[1,2]

1. Developmental Biology Program, Baylor College of Medicine
2. Howard Hughes Medical Institute, Baylor College of Medicine

Corresponding author: Meng C. Wang, Huffington Center on Aging, Baylor College of Medicine,1 Baylor pl., Houston, TX, E-mail: wmeng@bcm.edu

The rate of aging can be modulated by nutritional, environmental and metabolic cues. How those signals are coordinated among different tissues is a fundamental question for understanding aging regulation. The fat storage tissue plays an important role in regulating the aging process. Indeed, intestine-specific alterations in metabolic genes are sufficient to alter lifespans in both *C. elegans* and *Drosophila*. Recently, lipids and neuropeptide signals have been shown to be crucial endocrine signals linking peripheral metabolic tissues and the central nervous system.

In our previous work, we established a transgenic model (*lipl-4 Tg*) where a lysosomal lipase *lipl-4* is constitutively expressed in the intestine, the fat storage tissue of *C. elegans.* The *lipl-4* transgenic (*Tg*) animals have increases lifespan by more than 40% and transcriptional induction of neuropeptide processing genes compared to wild-type (wt) animals. We also identified a specific neuropeptide that is both required for the *lipl-4*-induced longevity and sufficient to prolong lifespan upon pan-neuronal overexpression.

Based on these findings, we hypothesize that lipid signals derived from *lipl-4*-induced lipolysis mediate the inter-tissue communication between the intestine and the nervous system. Most lipid messengers need to be transported into the target tissues via binding to lipid chaperones like Fatty Acid Binding Proteins (FABPs). FABPs, known as LBPs (Lipid Binding Proteins) in nematodes, are conserved from *C. elegans* to humans and consist of a class of nine LBP genes, *lbp-1* to *lbp-9*. We discovered two interesting candidates for mediating the intestine-neuron communication.

Future works will further investigate the tissue-specific effects of the candidate *lbps*, and the more downstream neuropeptide-mediated longevity mechanism. This project is at the interface of genetics, bioinformatics and nutritional regulation, which well integrate multidisciplinary training combining genetics, environmental health and data analysis. I expect my project will reveal novel regulatory mechanisms of longevity and diet-fat-neuron crosstalk, and have a broad impact on human health.

Funding sources: Training in Precision Environmental Health Sciences T32 ES027801

MicroRNA-122 and Poly-C Binding Protein 2 Promote Hepatitis C Virus Replication by Binding to Overlapping Targets on the 5' UTR.

Scott SD[1], Li Y[2], Lemon SM[2], Choi KH[1]
[1] Department of Biochemistry and Molecular Biology, Sealy Center for Structural Biology and Molecular Biophysics, University of Texas Medical Branch at Galveston
[2] Department of Medicine, The University of North Carolina at Chapel Hill

Corresponding author: Seth D Scott, Biology and Molecular Biology, University of Texas Medical Branch at Galveston. 301 University Boulevard, Galveston, TX, 77550. sedscott@utmb.edu .

Hepatitis C virus (HCV) is an important human pathogen that is capable of persistent infections in hepatocytes. Maintenance of the persistent infection requires precise control over protein translation and RNA synthesis. Successful HCV replication is dependent on many host factors; of importance to this study is microRNA-122 (miR-122) and Poly-C Binding Protein-2 (PCBP2), both of which have been shown to promote HCV replication. PCBP2 promotes viral protein translation while miR-122 promotes both translation and RNA synthesis in a PCBP2 dependent manner. The HCV 5' UTR contains two miR-122 binding sites near the first stem-loop (SLI) and a PCBP2 binding region that overlaps the second miR-122 binding site. PCBP2 also binds to the 3' UTR and has been shown to circularize the HCV genome. The goal of this project is to better characterize the binding of these factors to the HCV genome and understand whether their competition causes significant changes in the structure of the 5' UTR.

Both miR-122 and PCBP2 were characterized through their binding to the 5' UTR of HCV through isothermal calorimetry (ITC) and electrophoretic mobility shift assays (EMSAs). The first 45 nucleotides of the HCV 5' UTR (5'-HCV$_{45}$) and miR-122 were synthesized. PCBP2 was expressed in BL21 (DE3) cells and purified. ITC was conducted by titrating either miR-122 or PCBP2 into 5'-HCV$_{45}$. The results of the ITC showed the two binding sites for miR-122 have unequal affinities with K_d's of 979±84 nM and 11.1±1.5 nM. The binding was found to be very exothermic with changes in enthalpy of -139.6±72 kJ/mol and -425.5±90 kJ/mol respectively. PCBP2 on the other hand has been found to have an affinity to 5'-HCV$_{45}$ similar to the lower affinity miR-122 binding site. In addition to these affinity values, EMSA was used to analyze the interactions among 5'-HCV$_{45}$, miR-122 and PCBP2, and to determine the molar ratios required for saturation. Our result shows that an excess of miR-122 blocks binding of PCBP2 to the 5'-HCV$_{45}$. Together this data suggests that these two factors are competing for overlapping binding sites on the HCV genome. Experiments to characterize this competition are ongoing.

We suggest that the competition of these two factors may in part regulate whether an individual HCV genome is used for RNA synthesis or protein translation in the cell. Future experiments will better characterize the competition between these two factors. Additionally, we have started initial screens to solve the structures of both miR-122 and PCBP2 bound to the 5' UTR of HCV. This work will give a better understanding of how HCV maintains precise viral loads during a persistent infection through controlled interactions with host factors

Founding Sources:
Supported by a training fellowship from the Gulf Coast Consortia, on the Houston Area Molecular Biophysics Program (Grant No. T32GM008280).

Modeling Mutational Processes in Cancer

Shih DJH[1], McGrail D[1], Müller P[5], Lee N[3], Peng G[4], Chen J[3], Do K-A[2], Lin S-Y[1]

[1] Department of Systems Biology, University of Texas M.D. Cancer Center
[2] Department of Biostatistics, University of Texas M.D. Cancer Center
[3] Department of Experimental Radiation Oncology, University of Texas M.D. Cancer Center
[4] Department of Clinical Cancer Prevention, University of Texas M.D. Cancer Center
[5] Department of Statistics and Data Science, University of Texas at Austin

Corresponding author: Shiaw-Yih Lin, Department of Systems Biology, University of Texas M.D. Cancer Center, 1300 Moursund Street Unit 1058, Houston, TX 77030. Email: sylin@mdanderson.org

Background:
Cancer cells depend on a precarious balance between DNA damage and repair. Defects in DNA repair allow cancer cells to acquire mutations in oncogenes and tumor suppressors, in order to progress or evolve resistance to treatment. Conversely, excessive DNA damage can cause cellular death or produce new cancer-specific antigens that induce immune response against cancer cells. Recent sequencing studies have begun to characterize patterns of mutations by computationally decomposing mutational profiles across numerous cancer types into mutational signatures. However, several of these discovered signatures are highly similar and thus redundant, and most signatures have unknown etiology, revealing the gap that the catalogue of mutational processes is incomplete and poorly understood.

Objectives:
We aim to characterize mutational patterns in cellular models of DNA repair deficiency and to establish an improved statistical model for the discovery of parsimonious signatures of mutational processes.

Methods:
Inactivating mutation in *MSH2* was introduced by CRISPR-Cas9 genome editing in the non-malignant epithelial cell line MCF10A, and abrogation of MSH2 protein expression was confirmed by western blotting. The *MSH2* mutant cells were cultured for > 100 population doublings and single clones were derived. Mutations in *MSH2* mutant and parental cells were determined by whole-exome sequencing. Moreover, a preliminary Bayesian model for factor analysis was developed using a determinantal point process prior; model learning was performed using reversible jump Markov chain Monte Carlo.

Results:
The *MSH2* mutants accumulated a limited spectrum of single nucleotide variants, with high frequencies of C>T and T>C substitutions, and this mutational spectrum is modestly similar to the COSMIC mutational signature 6, which has been associated with mismatch repair deficiency. In contrast to the HAP1 near-haploid cancer cell line that was previously used for studying mutational processes, the MCF10A cell line did not exhibit high prevalence of C>A mutations (which are suggestive of oxidative damage) under our culture conditions. Additionally, simulation results confirmed that the fitting of our model can reach convergence quickly and that the determinantal point process prior promotes the discovery of a dissimilar, non-redundant set of latent factors.

Conclusions:
MCF10A is a suitable baseline cellular model for studying mutational processes due to its intact DNA repair pathways. Further, determinantal point process is an appropriate prior for discovering interpretable mutational signatures. Therefore, the interim results demonstrate the feasibility of our experimental and computational approaches for modeling mutational processes.

High-Throughput Optical Coherence Elastography for Colorectal Cancer Detection

Singh M[1], Mayerich D[2], Bhutani M[3], Larin K V[1,4,]*
[1]Deparment of Biomedical Engineering, University of Houston
[2]Department of Electrical and Computer Engineering, University of Houston
[3]Department of Gasteroenterology and Hepatology, University of Texas MD Anderson Cancer Center
[4]Department of Molecular Physiology and Biophysics, Baylor College of Medicine
*Kirill V. Larin, Department of Biomedical Engineering, University of Houston, 3517 Cullen Blvd, Room 2027, Houston, TX; klarin@uh.edu

There are approximately 140,000 new cases of colorectal cancer (CRC) each year in the USA, and approximately 50,000 CRC-related deaths (2014 CDC Cancer Statistics). Early stage colorectal cancer survival rates are high (>90%), but late stage colorectal cancer survival rates are very low (~13%). Moreover, only ~40% of CRCs are detected at early stages. Therefore, there is a need for early and accurate detection of colorectal cancer, as well as techniques for evaluating therapeutic outcomes.

My project will utilize optical coherence elastography (OCE) to quantify the micrometer-scale biomechanical properties of healthy and cancerous colon tissue biopsy samples. OCE is the elastographic functional extension of optical coherence tomography (OCT), which is an interferometric optical imaging technique capable of μm-scale spatial resolution and nm-scale motion sensitivity. OCE will enable high-resolution imaging of smaller areas of interest as compared to traditional elastographic techniques, such as magnetic resonance elastography or ultrasound elastography. For example, aberrant crypt foci (ACF) have been implicated in colon cancer etiology, but their micrometer-scale dimensions mean that standard clinical imaging modalities cannot detect ACFs. Moreover, it is unknown if ACF biomechanical properties are different from surrounding tissues and whether biomechanical properties can provide crucial information for malignancy classification. My project aims to identify CRC precursors such as ACFs with high-resolution multimodal optical and biomechanical imaging.

In addition to elastographic imaging, OCE provides a free OCT image, which can image tissues in 3D with micrometer-scale spatial resolution. OCT also can image up to 2 mm in tissue, which other optical imaging modalities, such as confocal microscopy, cannot. Therefore, the optical and structural properties of the colon tissue can also be assessed in 3D. Thus, my project will utilize the tissue optical properties, such as scattering and heterogeneity, and local structural features in additional to biomechanical properties when developing computational models to assess tissue diseased state and severity. Previous work in our lab on other tissues, such as kidneys and skin, showed that combining OCT and OCE with computational models can improve disease detection accuracy, sensitivity, and specificity by ~20%. The OCT+OCE system will be developed into a clinic-friendly setup that can be integrated into the current CRC detection and management pipeline. Tissue biopsy samples will be imaged by the OCT+OCE system, and computational models will link the high-resolution and high-dimension optical and biomechanical data to pathology reports, histopathology analyses, and known CRC biomarkers. Preliminary testing has shown the ability to detect hard inclusions simulating tumors in phantoms at sub-clinical scales with both optical and mechanical parameters.

Manmohan Singh is supported by a fellowship from the Gulf Coast Consortia, on the NLM Training Program in Biomedical Informatics and Data Science (T15LM007093).

Lineage-specific Epigenetic Reprogramming by Inorganic Arsenic in Primordial Germ Cells

Smith EA[1], Hodges HC[1,2,3,4]
1. Dept. of Molecular & Cellular Biology, Baylor College of Medicine, Houston, TX 77030
2. Center for Precision Environmental Health & 3. Dan L. Duncan Cancer Center; Baylor College of Medicine, Houston, TX 77030
4. Center for Cancer Epigenetics; MD Anderson Cancer Center, Houston, TX 77030

Corresponding author: H. Courtney Hodges, Dept. of Molecular & Cellular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX, chodges@bcm.edu.

Inorganic arsenic is a major environmental toxicant that significantly impacts early development and has well-established roles as an epigenetic reprogramming agent. Millions worldwide are exposed to naturally occurring arsenic through ingestion of contaminated food and water. Arsenic exposure not only has a broad impact disrupting several body systems, but it has also been shown to disrupt normal developmental processes. Epidemiologic data suggests early-life exposure to arsenic predisposes children to adverse future health outcomes. Hence, primordial germ cell development is a key window where the developing embryo is vulnerable to environment-induced epigenetic reprogramming; however, relatively little is known about how prenatal exposures influence predisposition to pathology later in life. As a result, an unbiased accounting of the dynamics of these exposures and mitigating factors during early embryogenesis is currently lacking.

I hypothesize that epigenetic reprogramming by inorganic arsenic arises in a lineage-specific manner that varies across distinct germ layers. Single-cell sequencing technologies offer a data-rich experimental approach to test this hypothesis and reveal the effects and dynamics of arsenic exposure on the developing epigenetic landscape. Such approaches allow for detailed dissection of cellular features of subpopulations that are often obscured by traditional population-based bulk methods. Embryoid bodies (EBs) are an ideal *in vitro* 3D model of early embryogenesis, because they possess a mixture of the three primordial germ layers and multipotent developmental potential. I will first examine the effects of physiologic levels of inorganic arsenic exposure to EBs using single-cell RNA-seq, providing the first high-resolution cellular transcriptome map of arsenic-induced changes during early development. Separately, I will identify the specific changes to chromatin architecture in each germ layer upon arsenic exposure using single-cell ATAC-seq. By overlapping motif profiling from scATAC-seq with transcriptome data from scRNA-seq, I will reveal specific transcription factors whose activities are altered in response to arsenic exposure. Furthermore, following exposure to arsenic, I will use FACS to isolate cells from each of the three germ layers, then use ChIP-seq to measure histone modification patterns across each layer to identify altered heritable histone marks such as H3K9me3. Similar isolation techniques will be coupled with reduced representation bisulfite sequencing (RRBS) to define changes in DNA methylation patterns that contribute to epigenetic changes that are inherited to daughter cells. Finally, our studies will characterize the persistence and reversibility of these exposures following withdrawal or chelation therapy. We aim to test the hypothesis that a subset of the epigenomic changes are reversible following treatment.

Overall, our high-resolution study will make direct connections between arsenic-induced epigenetic changes and health outcomes during early embryogenesis. Our *in vitro* dissection of the *in vivo* outcomes upon arsenic exposure will define exposure-induced epigenetic reprogramming across cell lineages during embryogenesis. Combining an *in vitro* model of embryogenesis with single-cell and lineage-targeting approaches will reveal the cell-type and cell-state specific epigenetic patterns and their dysregulation during arsenic exposure and treatment, significantly aiding the interpretation of human health studies.

This study is supported by NIEHS T32 ES027801 (E.A.S.), NIH grant R00 CA187565 (H.C.H.), CPRIT grant RR170036 (H.C.H.), and the V Foundation (H.C.H.).

Prenatal Exposure to Polycyclic Aromatic Hydrocarbons (PAHs) Augments Neonatal Hyperoxic Lung Injury and Alters the Gut Microbiome in Mice: Mechanistic Role of Cytochrome P450 (CYP)1A1, 1A2, and 1B1

Swanson L[1], Jiang W[1], Veith A[1], Hoffman K[1], Lingappan K[1], and Moorthy B[1]

[1]Section of Neonatology, Department of Pediatrics, Baylor College of Medicine, Texas Children's Hospital, Houston, TX 77030

Corresponding Author: Lauren Swanson, Section of Neonatology, Department of Pediatrics, Baylor College of Medicine, Feigin Center, 1102 Bates Ave, Houston, Texas, Email: Lauren.Swanson@bcm.edu

Pregnant women living in the vicinity of superfund sites or smokers are at higher risk for preterm delivery, in part due to exposure to polycyclic aromatic hydrocarbons (PAHs). Preterm infants often require treatment with supplemental oxygen (hyperoxia) that in turn could lead to a chronic lung disease of prematurity called bronchopulmonary dysplasia (BPD). The molecular mechanisms by which hyperoxia causes lung injury are not understood, but cytochrome P450 (CYP) enzymes have been implicated. Studies from our laboratory have demonstrated that CYP1A1 induction is protective against hyperoxic lung injury in mice while CYP1B1 acts as a pro-oxidant. The central hypothesis of this project is that prenatal administration of PAHs [i.e. benzo[a]pyrene (BP), or a mixture of BP and benzo(b)fluoranthene (BbF)] differentially exacerbates lung injury, alveolar simplification, and causes dysbiosis of the gut microbiome in neonatal mice following postnatal hyperoxia, and that this effect is altered in mice lacking the gene for cytochrome P450 *(Cyp)1a1, 1a2,* or *1b1.* Experiments were performed to determine the dose-response of prenatal PAH administration on postnatal hyperoxic lung injury testing PAH doses of 7.5mg/kg, 15mg/kg, and 30mg/kg. Results showed that hyperoxic lung injury is augmented in a dose-dependent manner with higher doses of prenatal PAH exposure causing greater lung injury in neonatal mice. Timed pregnant WT (C57BL/6J), *Cyp1a1*-null, *Cyp1a2*-null and *Cyp1b1*-null mice were treated orally with the vehicle corn oil (CO) or mixture of PAHs BP and BbF (7.5 mg/kg each) on gestational days 16-19. The newborn mice obtained from these mothers (term, day 21) were exposed to hyperoxia or room air for 14 days. On PND 14, the mice were sacrificed, and lung injury and alveolar simplification were assessed by quantification of the radial alveolar count (RAC). Results showed that PAH exposure differentially exacerbates lung injury in WT, *Cyp1a1*-null, *Cyp1a2*-null and *Cyp1b1*-null mice and that hyperoxia caused reduction in RAC across all genotypes. PAH treatment resulted in significant induction of *Cyp1a1* gene expression in room air, with suppression of *Cyp1a1* expression following hyperoxia. Expression of pro-inflammatory cytokine genes (TNF and IL-6) were examined in pulmonary specimens to determine the degree of inflammation across samples. In room air, *Cyp1a1*-null and *Cyp1b1*-null mice expressed significantly more TNF compared with WT mice. In hyperoxia, TNF and IL-6 expression were enhanced in all genotypes. 16S rRNA sequencing of gut microbiome samples at the PND14 timepoint revealed differences in Bray-Curtis beta diversity observed between PAH and CO groups in WT mice, and this was not seen in *Cyp1a1*-null mice, suggesting that Cyp1a-mediated metabolism of PAH plays a role in altering the intestinal microbiome. Future studies could lead to the development of novel strategies against BPD in premature infants exposed prenatally to PAHs.

Research reported in this publication was supported in part by the National Institute of Environmental Health Sciences of the National Institutes of Health under Award Number T32ES027801. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Additional funding sources include grants 1R56 ES030221-01 and 5R01HL129794-04 to Bhagavatula Moorthy.

Defining Transbilayer Lipidomic Profiles and Determining Consequences of Biophysical Asymmetry in Mammalian Plasma Membranes

Symons JL[1], Levental KL[1], Levental I[1]

[1]Department of Integrative Biology and Pharmacology, McGovern Medical School at UTHealth
[2]The University of Texas MD Anderson UTHealth Graduate School of Biomedical Sciences

Corresponding author: Ilya Levental, Department of Integrative Biology and Pharmacology, The University of Texas Health Science Center at Houston, McGovern Medical School, 6431 Fannin St, Houston, TX, E-mail: ilya.levental@uth.tmc.edu

A fundamental and broadly conserved feature of eukaryotic cells is an unequal distribution of lipids between the two leaflets of the plasma membrane bilayer. Maintaining lipid asymmetry is energetically costly, implying an essential, though as yet poorly understood, physiological role. While the broad features of phospholipid distribution between plasma membrane leaflets have been defined for decades, the asymmetric distribution of cholesterol, the most abundant component of the plasma membrane, remains a major open question. Similarly, the asymmetry of plasma membrane phospholipids can be perturbed via lipid scrambling, though it is largely unknown whether specific phospholipids are redistributed during this process and whether cholesterol transbilayer distribution is also altered. Finally, the extent to which this lipid scrambling changes the biophysical properties of the plasma membrane in live cells is unknown. Through novel FRET and enzyme-assisted phospholipidomic techniques, we will rigorously define transbilayer lipidomic profiles in steady-state and scrambled mammalian plasma membranes. Using fluorescent spectroscopy of environment-sensitive probes, we will examine changes in plasma membrane biophysical properties - including lipid packing, permeability, and diffusivity – induced by phospholipid scrambling. The mechanisms underlying effects in biomembranes will be defined via biomimetic membrane model systems and computational approaches.

This work is supported by:
Gulf Coast Consortia, Houston Area Molecular Biophysics Training (HAMBP) Program (NIH Grant No. T32GM008280)
UT MD Anderson UTHealth Graduate School of Biomedical Sciences, American Legion Auxiliary Fellowship in Cancer Research

Connect the Dots: A Method for Interpreting Multi-Metabolite Perturbations for the Diagnosis of Metabolic Disease Using Disease-Specific Metabolomics Networks

Thistlethwaite, L.R.[1,2], Petrosyan, V.[1], Riehle, K.[1], Burrage, L.C.[1], Scaglia, F.[1], Miller, M.J.[3], Elsea, S.H.[1], Milosavljevic, A.[1,2]

1. Department of Molecular & Human Genetics, Baylor College of Medicine, Houston, TX
2. Program in Quantitative & Computational Biosciences, Baylor College of Medicine, Houston, TX
3. Department of Medical & Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN

Corresponding author:
Lillian Thistlethwaite
Program in Quantitative & Computational Biosciences
Baylor College of Medicine
1 Baylor Plaza, Houston, TX
lillian.thistlethwaite@bcm.edu

Untargeted metabolomics profiling captures a global view of the metabolic state of an individual, and diagnosis using untargeted metabolomics is data-driven and unbiased. One prominent practice in diagnosing Inborn Errors of Metabolism (IEM) is to follow a targeted approach, where a diagnosis is nominated based on clinical symptoms observed, and targeted biochemical testing is performed to either confirm or rule out the nominated diagnosis. Problems with this targeted approach to diagnosis occur when clinical symptoms are undifferentiated across multiple IEM. In these circumstances, untargeted metabolomics profiling has shown to be useful. The challenge with diagnosis using untargeted profiling, however, is quantifying the significance of multi-metabolite perturbations observed in individual profiles using a standardized, automatic and transparent methodology. Currently, interpretation of untargeted metabolomics data is performed manually, where clinicians use their biochemical knowledge to reason about the patterns observed. In effect, metabolite perturbations can be examined by superimposing those perturbations onto biochemical pathway maps, and searching for bottlenecks where metabolite perturbations around an affected enzyme show opposite directionality.

To improve accuracy and transparency and to automate this subjective and manual diagnostic process, we develop CTD, a computational diagnostic method that “connects the dots” between metabolite perturbations observed in patient’s metabolomics profiles with patterns observed in disease-specific networks. In this work, we show that disease-specific metabolomics perturbation networks lead to more accurate detection of disease-specific metabolite perturbation patterns when compared to general biochemical pathways. Furthermore, using disease networks for 16 different IEM (e.g., 4 organic acidemia disorders, 2 Zellweger spectrum disorders, 4 urea cycle disorders, 3 nucleotide metabolism disorders, and 3 others), we show that CTD reproduces expert performance in diagnosing IEM disorders.

Lillian Thistlethwaite is supported by a training fellowship from the Gulf Coast Consortia, on the NLM Biomedical Informatics Training Program (Grant No. T15 LM007093).

Characterization Of Relationships Between Pathogenesis And Antimicrobial Resistance-Encoding Mobile Genetic Elements In Group A *Streptococcus*

Vega LA[1], Sansón-Iglesias M[1], Flores AR[1,2]

1. Division of Infectious Diseases, Department of Pediatrics, McGovern Medical School
2. Center for Antimicrobial Resistance and Microbial Genomics at the University of Texas Health Sciences Center at Houston

Corresponding author: Anthony R. Flores, Division of Infectious Diseases, Department of Pediatrics, McGovern Medical School. 6431 Fannin St. MSE R.234. Houston, TX 77030. Anthony.R.Flores@uth.tmc.edu

Mobile genetic elements (MGE) frequently contain antimicrobial resistance genes (AMR) genes and may be rapidly transferred between bacterial strains. MGE presence can alter global gene regulation patterns that may in turn enhance the pathogen's virulence and/or transmissibility, providing a selective advantage to the maintenance and propagation of AMR that is independent of antimicrobial selective pressure. Surveillance of pediatric Group A *Streptococcus* disease in Houston, TX identified *emm92* serotype strains with high-frequency AMR to multiple second-line antimicrobials (e.g. macrolides, tetracycline, and aminoglycosides). The genes conferring aminoglycoside and tetracycline resistance reside on a MGE dubbed *ICESpyM92*. Moreover, *emm92* GAS isolates were recovered almost exclusively from invasive or skin and soft tissue infections. Genomic analyses of Houston and nationally representative GAS encoding *ICESpyM92* revealed a high degree of relatedness, suggesting recent emergence of the AMR isolates. The high frequency of AMR in *emm92*, compared to AMR in other circulating GAS strains, combined with an association with severe disease suggests a possible selective advantage, apart from AMR, driving maintenance of the resistance-encoding MGE. We hypothesize that *ICESpyM92* alters the streptococcal transcriptome, enhancing virulence and/or transmissibility of GAS, thus contributing to high-frequency resistance. To test this hypothesis, we are 1) determining the correlation of *emm92* serotype emergence in severe/invasive GAS disease with acquisition of *ICESpyM92*, 2) characterizing the GAS virulence phenotypes and transcriptome associated with the presence of *ICESpyM92*, and 3) examining regulatory elements in *ICESpyM92* that potentially influence global gene regulation in GAS *emm92*. Whole genome sequencing (WGS) of *emm92* isolates from local and nationally representative strains combined with phylogenetic analyses will characterize evolution of *emm92* GAS, including the emergence of AMR, and determine associations with GAS disease phenotypes (e.g. invasive infections). Comparison of *emm92* GAS and an isogenic mutant lacking *ICESpyM92* using *in vivo* infection models, and of their respective transcriptomes through RNAseq analysis, will measure changes in the GAS transcriptome related to *ICESpyM92* and the contribution of this MGE to disease caused by *emm92* GAS. Completion of the proposed studies will serve as a foundation for further investigation into the role of MGE and AMR in GAS disease. Moreover, the skills and expertise gained through these studies will enable examination of similar effects in related Gram-positive pathogens. Understanding the influence of resistance-encoding MGE on pathogen virulence and transmissibility will assist in effectively combating the emergence of AMR pathogens.

Supported by NIAID grant 1T32AI141349-01A1

The Subclonal Architecture of Triple Negative Breast Cancer

Authors: Vera Recio CC[1], Wang W[1], Lozano G
1. University of Texas MD Anderson Cancer Center

Corresponding Author: Carlos C Vera Recio, Department of Quantitative Sciences, University of Texas MD Anderson Cancer Center, 1400 Pressler St, Houston, TX Email: cvera@mdanderson.org

Triple negative breast cancer (TNBC) studies have identified tumor heterogeneity and subclonal mutations as mechanisms of resistance to therapy. The ARTEMIS trial: A Randomized, TNBC Enrolling trial to confirm Molecular profiling Improves Survival; is a CPRIT funded clinical trial to identify and characterize chemo-sensitive TNBC. ARTEMIS trial generates germline and tumor whole exome sequencing (WES) data from patients at each of their treatment steps (n = 360). However, tumor samples tend to under-sample cancer cell populations, making subclonal reconstruction of individual tumors very challenging. We propose a novel computational method which leverages time information and data from multiple samples to improve subclonal reconstruction of TNBC tumors.

We use clinical outcome to define TNBC cases as resistant or sensitive. WES data generated at various stages of each patient's treatment is processed using a state-of-the-art bioinformatics pipeline to estimate somatic point mutations, allele-specific copy number aberrations, tumor purity and single time-point cluster assignments for point mutations. We leverage co-clustering relationships of point mutations, time information, statistical and machine learning techniques to remove noise from our mutation information while extracting evolutionary relationships.

We're currently processing 118 tumor samples coming from 59 patients, with data from an additional 300 patients still being generated. At the end of our research project, we'll have developed a novel method for subclonal reconstruction of tumor applicable to a prevalent and aggressive disease. As part of our ongoing work, we are establishing an analysis pipeline which facilitates finding patterns of chemo-resistance and sensitivity in TNBC.

This work is supported by the NLM Training Program in Biomedical Informatics and Data Science T15LM007093, Program Director Dr. Lydia Kavraki.

Poster #64

Small-Molecule Inhibitor of OGG1 Suppresses Proinflammatory Gene Expression and Inflammation

Wang K[1], Hao W[1,2], Zheng X[1], Pan L[1,3], Ba X[1,2], and Boldogh I[1]
[1]Department of Microbiology and Immunology, University of Texas Medical Branch at Galveston, Galveston, Texas, USA.
[2]Institute of Genetics and Cytology, Northeast Normal University, Changchun, RP China
[3]Department of Physiology, Xiangya Medical School, Central South University, Changsha, RP China

Corresponding author: Istvan Boldogh, Department of Microbiology and Immunology, University of Texas Medical Branch at Galveston, Galveston, Texas, USA. Email: sboldogh@utmb.edu

8-oxo-7,8-dihydroguanine (8-oxoG) is generated by one electron oxidation of guanine in nucleic acids by environmental exposures (physical, chemical, biological agents)-induced as wells endogenously generated reactive oxygen species (ROS). It is primarily repaired by 8-oxoguanine DNA glycosylase1 (OGG1) during the DNA base excision repair pathway to maintain genome fidelity and to prevent its potential mutagenicity. Its accumulation in the genome has been correlated with cellular dysfunction, and various inflammatory diseases as well as aging. However, recent studies show that it may serves as epigenetic-like mark before its excision and OGG1 is a regulator of gene expression.
Because *Ogg1*-deficient mice are resistant to acute and systemic as well as allergic inflammation, we proposed that OGG1 inhibition may represent a strategy for the prevention and treatment of inflammatory processes. To do so, a selective active-site inhibitor of OGG1 was developed, which inhibits OGG1 substrate binding and initiation of DNA base excision repair and showed no toxicity in cultures cells or in animals. Impact of small molecule was tested using various molecular approaches (excision assays, qRT-PCR, EMSA) on pro-inflammatory gene expression, and mouse model of lung inflammation (histology and clinical symptoms). Administration of OGG1 inhibitor(s) to mice prevents tumor necrosis factor–alpha– and lipopolysaccharides (LPS)- as well as viral infection induced OGG1-DNA interactions at regulatory regions of proinflammatory genes. This, in turn, decreases DNA occupancy of nuclear factor kappaB and proinflammatory gene expression, resulting in decreased recruitment of immune cells and histological changes in lungs. Thus, targeting OGG1-driven DNA base repair may have clinical utility to lessen inflammatory conditions induced by environmental exposures.
Funding: National Institute of Allergic and Infectious diseases (NIAID AI062885).

Role of Ataxia Telangiectasia and Rad3 Related Cell Cycle Regulation in Occlusive Vascular Disease Pathogenesis in Majewski Osteodysplastic Primordial Dwarfism Type II

Wright JM[1], Niu A[1], Zhou Z[1], Buja LM[2], Milewicz DM[1]
1. Department of Internal Medicine, Division of Medical Genetics, UTHealth McGovern Medical School
2. Department of Pathology and Laboratory Medicine, UTHealth McGovern Medical School
Corresponding author: Jamie Wright, Department of Internal Medicine, Division of Medical Genetics, UTHealth McGovern Medical School, 6431 Fannin, Houston, TX, E-mail: jamie.m.wright@uth.tmc.edu.

Majewski Osteodysplastic Primordial Dwarfism Type II (MOPDII) is a rare autosomal recessive disorder caused by biallelic loss of function mutations in, *PCNT*, the gene encoding the centrosome protein pericentrin. Patients with MOPDII have a striking clinical phenotype, including severe pre- and post-natal growth restriction, microcephaly, and bony dysplasia. Early-onset vascular disease, including Moyamoya disease, cerebral aneurysms, and coronary artery disease, is a major source of morbidity and mortality in this population. Vascular diseases driven by gene mutations tend to be early-onset and occur in the absence of traditional risk factors like atherosclerosis. Previous studies in our lab suggested that such genetically triggered vascular disease is most likely due to aberrant proliferation of vascular smooth muscle cells (vSMCs). As a centrosome scaffold protein, pericentrin is involved in a wide range of cell activities, including mitotic spindle formation, microtubule nucleation, and ataxia telangiectasia and rad3 related (ATR)-mediated cell cycle regulation. We hypothesize that **loss of function mutations of pericentrin drive vascular disease in patients with MODPII through dysregulation of the ATR pathway in vSMCs leading to hyperproliferation in response to cell stress, and through altered vSMC migration due to abnormal microtubule elongation**.

To test this hypothesis, we are utilizing a novel smooth muscle cell specific *Pcnt* knock-out (*Pcnt*$^{SM\text{-}KO}$) mouse model developed in our laboratory to examine the effects of pericentrin knock-out on vSMC proliferation, migration, and ATR-mediated cell cycle regulation *in vitro* in explanted vSMCs. In addition, we are assessing the vascular phenotype *in vivo* using echocardiography, tail cuff blood pressure measurement, and histologic analysis.

Findings from these experiments will increase our understanding of vascular disease pathogenesis in patients with MOPDII and of ATR regulation in occlusive vascular disease more generally. These findings will open new avenues of investigation for prevention and management of occlusive vascular disease.

Funded by: Training Interdisciplinary Pharmacology Scientists (TIPS), National Institutes of Health grant T32GM120011 to The University of Texas Health Science Center at Houston and support from the Potentials Foundation.

Poster #59

Metabolic Fate of Saturated and Unsaturated Dietary Lipids

Yavuz A[1], Wang MC[1,2,3]

1. Huffington Center on Aging, Baylor College of Medicine
2. Molecular and Human Genetics, Baylor College of Medicine
3. Howard Hughes Medical Institute, Baylor College of Medicine

Corresponding author: Meng C. Wang, Huffington Center on Aging, Baylor College of Medicine, 1 Baylor Pl., Houston, TX, E-mail: wmeng@bcm.edu

Dietary fat is a key environmental factor that impacts the risk of obesity. Two major dietary fat types are saturated fatty acids and unsaturated fatty acids. Excess accumulation of saturated fatty acids in non-adipose tissues are toxic and may also lead to other diseases including nonalcoholic fatty liver diseases and cardiomyopathy. Unsaturated fatty acids, on the other hand, are less toxic and may even exert protective effects. Recent studies suggest that difference of the fatty acids in lipotoxicity may be associated with their distinct efficiency during the incorporation into lipid droplets (LDs). However, the underlying molecular mechanisms governing this LD incorporation heterogeneity between the different fatty acids remain poorly understood.

The purpose of this project is to dissect the molecular mechanisms that regulate LD incorporation heterogeneity between different fatty acid molecules using *C. elegans* as a model organism. Previously, our lab reported a clear distinction between saturated and unsaturated fatty acids in their incorporation into LDs using isotope-labeling coupled stimulated Raman scattering (iSRS). iSRS is a chemical imaging strategy for tracking the spatiotemporal dynamics of specific lipid molecules at the subcellular resolution in living cells and organisms. In our current work, we are applying this method in an unbiased genetic screen to identify mutants causing the different rates in LD incorporation of Palmitic acid-D31 and Oleic acid-D34, which represent saturated and unsaturated fatty acids, respectively. In our pilot screen, we have identified three mutants that displayed significantly increased LD incorporation of saturated fatty acids. We have sequenced these candidates and utilized computational tools to identify the variants. We are currently using RNAi to knockdown our candidates for further verification of the causative mutations.

Funding sources: Training in Precision Environmental Health Sciences T32 ES027801

Inhibition of OGG1' Interactions with its Genomic Substrate Decreases Proinflammatory Gene Expression

Zheng X[1] , Wang K[1], Hao W[1,2], Pan L[1,3], Ba X[1,2], and Boldogh I[1]
[1]Department of Microbiology and Immunology, University of Texas Medical Branch at Galveston, Galveston, Texas, USA
[2]Institute of Genetics and Cytology, Northeast Normal University, Changchun, RP China
[3]Department of Physiology, Xiangya Medical School, Central South University, Changsha, RP China

Corresponding author: Istvan Boldogh, Department of Microbiology and Immunology, University of Texas Medical Branch at Galveston, Galveston, Texas, USA. Email: sboldogh@utmb.edu

Reactive oxygen species (ROS) generated by environmental chemical, physical, biological exposures generates oxidatively damaged macromolecules including proteins lipids and DNA. In DNA and RNA 7,8-dihydro-8-oxoguanine (8-oxoG) is generated by one electron oxidation of guanine. 8-OxoG is repaired primarily by the 8-oxoguanine DNA glycosylase 1 (OGG1)-initiated DNA base excision repair pathway. Although it considered as a mutagenic base lesion, recent studies show that it may serve as epigenetic-like mark before its excision and OGG1 is a modulator of gene expression. Because *Ogg1*-kockout mice are increasingly resistant to acute, systemic and allergic inflammation, we hypothesized that OGG1 inhibition may be a strategy to prevent and treatment of inflammatory gene expression. Therefore, specific OGG1 inhibitor(s) was tested using various molecular methods (qRT-PCR, electrophoretic mobility shift assays, chromatin immunoprecipitation, immunoblotting) and cell culture models to examine its inhibitory effect on 8-oxoG excision, DNA occupancy of transcription factors and expression of pro-inflammatory genes. Results showed that only the active-site inhibitor of OGG1, prevented OGG1's binding to its genomic substrate, base excision, decreased binding of nuclear factor kappaB (NFkappaB) its consensus sequences, while did not show inhibitory effect on NFkappaB signaling pathways. In turn, it decreased proinflammatory gene expression in various cell types induced by tumor necrosis factor alpha, lipopolysaccharides or viral infections. In controls, inhibitor of OGG1' glycosylase activity or knock-down of NEIL1 or 2 had no significant effect on proinflammatory gene expression. Thus, targeting OGG1 binding to and repair of oxidatively modified guanine is an effective way to prevent/alleviate inflammatory gene expression.
Funding: National Institute of Allergic and Infectious diseases (NIAID AI062885).

Elucidating the Mechanism of Developmental Toxicity of the Anti-retroviral Drug Dolutegravir

Yeo L[1], Thierer JH[2], Gorelick DA[3], Finnell RH[3] and Cabrera RM[3]
1. Department of Biochemistry and Molecular Biology, Baylor College of Medicine
2. Department of Biology, Johns Hopkins University
3. Department of Molecular and Cellular Biology, Baylor College of Medicine

Corresponding author: Lythou Yeo, Department of Biochemistry and Molecular Biology, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030, E-mail: Lythou.Yeo@bcm.edu

Dolutegravir (DTG) is an integrase inhibitor that is commonly administered in combination with other medications as antiretroviral therapy (ART) or as a post-exposure prophylaxis for human immunodeficiency virus (HIV) infection. Recent findings show correlation of DTG with up to a 9-fold increase in neural tube defect (NTD) risk. We hypothesize that DTG-induced developmental toxicity results from its interaction with some metabolic pathway that is essential for proper neurulation, such as the folate cycle and one-carbon metabolism pathways. It is well-established that folate-deficiency is an NTD risk factor, and maternal folate supplementation is currently the most effective prevention strategy for the majority of NTD cases. Preliminary data supports that DTG is a partial antagonist of folate receptor 1 (FOLR1) and causes developmental toxicity that coincides with pre-gastrulation embryonic development. On-going studies include determining the critical window for the onset of DTG teratogenic activity and using the CRISPR/ Cas9 system to generate mutant zebrafish models of various folate receptors and transporters for future studies to identify the specific folate uptake mechanism and downstream interactions that mediate DTG toxicity. Results from a zebrafish time course experiment suggest that embryos must be exposed to 100μM DTG prior to gastrulation to produce developmental toxicity. With embryos that were exposed to 100μM DTG 2 hours post fertilization (hpf), a teratogenic phenotype where the posterior end of the embryo develops as a black mass appears between 8 and 12 hpf. Affected embryos do not survive thereafter. Completion of this study will help further our understanding of how DTG works as a teratogen and ultimately, these findings may lead to potential countermeasures for DTG-induced NTDs.

Research reported in this publication was supported by the National Institute of Environmental Health Sciences of the National Institutes of Health under Award Number T32ES027801. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Expression and Characterization of the Human Astrovirus (HAstV) VP90 Capsid Protein for Investigating the Structural Features of the Proteolysis-Mediated Viral Maturation Process

Ykema MR[1], Xun M[1], Tao YJ[1]

1. Department of Biosciences, Rice University

Corresponding author: Matthew Ykema, Department of Biosciences, Rice University, 6100 Main Street, Houston, TX, Email: mry3@rice.edu

The human astrovirus (HAstV) is a non-enveloped virus with a positive-sense RNA genome and causes gastroenteritis in infants, the elderly and immunocompromised individuals. The virus has no vaccine, a high mutation rate, and emerging clinical symptoms like lethal encephalitis. Further structural and biophysical analysis must be done on the capsid protein to determine how it allows the virus to enter the host cell and transport its genome. The HAstV capsid is coded by the viral genome as a 90 kD polyprotein labeled as VP90, which forms a non-infectious particle. The capsid undergoes posttranslational modification through sequential proteolytic cleavages to form the mature infectious state. Little is understood of the infectious domains, cleavage sites, or capsid structural dynamics. We plan to express the immature virus-like particle using an E. coli expression system, simulate its maturation in vitro, then investigate these isolated intermediate particles with structural studies and cell culture testing.

Multiple VP90 capsid protein constructs have been expressed in E. coli and purified using HisTrap and size exclusion chromatography. Constructs with short N-terminal truncations had better expression and solubility when compared to full length constructs. Size-exclusion chromatography and transmission electron microscopy have indicated the capsid protein is forming two major states, a ~180 kD dimer and a rod-shaped virus-like particle. The VP90 dimer was screened for the formation of protein crystals for use in X-ray crystallography, but no crystals have been observed. The formation of the rod-like oligomerization was unexpected, and is being investigated to determining the assembly pathways and the antigenic properties of these particles. Once icosahedral virus-like particle formation has been confirmed, the particles will be used for structural reconstruction using cryo-electron microscopy, in vitro maturation assays, and liposome infiltration assays. This research is funded by the Welch Foundation (C-1565 to YJT) and (HAMBP T32GM008280 to MY)

Poster #62

Computer Simulations Show Key Role Stochasticity of Replication Fork Speed Plays in the Dynamics of DNA Replication

Yousefi R, Rowicka M
Department of Biochemistry & Molecular Biology, University of Texas Medical Branch

Eukaryotic DNA replication is elaborately regulated to ensure that the genome is faithfully replicated in a timely manner. Replication initiates at multiple origins, from which replication forks emanate and travel bi-directionally. Activation of replication origins and fork speed are stochastic in individual cells, but reproducible population-wide. To study the complex spatio-temporal regulation of replication, models of DNA replication in S. cerevisiae have been developed, but none have considered stochastic replication fork speed. Here, we present Repli-Sim, the first model of DNA replication, which includes stochastic speed of the replication fork. Utilizing data from both wild-type and hydroxyurea-treated yeast cells, we show that Repli-Sim achieves more accurate results than models assuming constant fork speed. Due to the stochastic nature of replication, its completion in a timely fashion is a challenge. Previously proposed solutions promoted finishing replication by modifying replication initiation and origin activation, while we propose empirically-derived modification in replication speed based on distance to the approaching fork, which promotes completion of replication.

This research was supported by the NIH grant R01GM112131 to M.R. (R.Y. and M.R.).