Pennsylvania State University

Annual Progress Report: 2010 Formula Grant

Reporting Period

July 1, 2013 – June 30, 2014

Formula Grant Overview

The Pennsylvania State University received $7,001,127 in formula funds for the grant award period January 1, 2011 through December 31, 2014. Accomplishments for the reporting period are described below.

Research Infrastructure Project 1: Project Title and Purpose

Research Infrastructure Renovation to Create the Penn State Center for Translational Informatics - The purpose of this project is to create a Center for Translational Informatics that will provide the necessary software tools for researchers to collect and manage patient data, analyze genomic information, and mine the electronic medical record for data to support research into comparative effectiveness and personalized medicine. To create the Center, approximately 600 sq ft of existing space will be renovated to establish a collaborative workspace to house informaticists, IT analysts, and data analysts who will consult with basic and clinical researchers, develop software, and manage databases. Software and hardware will be purchased to allow the collection, management, analysis and transfer of large datasets.

Anticipated Duration of Project

1/1/2011 – 12/31/2014

Project Overview

With advances in understanding the genetic basis of disease, translational research has relied more heavily on biomedical informatics tools, and new technologies are being exploited to make clinical research and healthcare delivery more efficient and effective. In this project we are creating a new Center for Translational Informatics (CTI) that will transcend departmental, college, and campus boundaries to implement a systematic, multi-thematic effort toward sustainable pathways for translational research. The specific aims of this project are to (1) Renovate existing space to create the bioinformatics support infrastructure for the CTI; (2) Support clinical research by assisting researchers in the management and analysis of complex biological and clinical data.

The CTI will house the Director and support staff who will consult with faculty to identify informatics needs and develop solutions, and will have space for training faculty, staff, and students on the use of software. The CTI will purchase software that will allow researchers to

conduct large-clinical trials more efficiently, analyze genomic data, keep clinical data secure and private, move large amounts of data more easily, and access the electronic medical records of patients to promote research in comparative effectiveness and personalized medicine.

Principal Investigator

Richard Rauscher, MS
Director of Research Informatics and Computing
Penn State University College of Medicine
500 University Drive
Hershey, PA 17033

Other Participating Researchers

Istvan Albert, PhD, Arthur S. Berg, PhD, William DiStefano, Benjamin Tshudy, Daijang Liu, PhD, Anna Salzberg - employed by Penn State University College of Medicine

Expected Research Outcomes and Benefits

Penn State has been engaged in a decade-long mission of removing barriers to translational research by creating targeted pathways to unify traditional information technology (IT) and biomedical informatics through systems and services that merge scholarly activity with production delivery, and by forming interdisciplinary centers that transcend departmental, college, and campus boundaries to seed new collaborative research directions and provide shared research infrastructure and instrumentation. The proposed Center for Translational Informatics (CTI) will forge a systematic, multi-thematic effort toward sustainable pathways for translational research by leveraging existing strengths and latent synergies between basic science and translational research to reach new milestones at the confluence of (1) genomics and biomedical research; and (2) informatics and clinical best practices. Targeted collaborations and additional investment and development of human capital and shared research infrastructure by Penn State will lead to new breakthroughs in translational research. Innovative projects supported through the CTI will use informatics to answer biological questions and to accelerate how laboratory findings and best practices are best implemented in the clinical setting, for example, using genomics data to support personalized medicine. Electronic medical records (EMRs) are one of the most exciting potential resources for research data. The CTI will allow researchers to access patients' EMRs to conduct complex genomic analyses of disease susceptibility and therapeutic outcomes across diverse patient populations and to identify best practices for Institutional Review Board interactions, participant consent, data sharing, returning research results, etc.

Summary of Research Completed

The renovation of the space to create the CTI is complete. The Clinical Trials Management System design and implementation project is nearing completion and is set to go live in September 2014.

Specific Aim 1: Renovate existing space to create the CTI and create the computational

infrastructure to support genomics and clinical research at Penn State.

The institution has recruited Dajiang Liu, PhD, a Biostatistics Professor, to work with Dr. Arthur Berg and Ms. Anna Salzberg in analyzing genetic data for our researchers. He is located in the CTI consulting space as designed and he and his colleagues have assisted numerous researchers in the genomic analyses of the various projects taking place here.

Scientific discussions and collaborations with University Park are also being enhanced by the use of the Polycom RealPrescence system. The use of the system has significantly increased over the past year. Some of the uses include monthly 2 hour meetings (September through June) with the Health Sciences Council at University Park, Clinical and Translational Sciences Institute update meetings with key function leaders (about 7 meetings a quarter), CTSI Staff meetings monthly with University Park, about 10 review meetings annually, and approximately 30 education meetings per annum. These meetings have improved the communications and collaborations in science between both campuses.

Specific Aim 2: Support clinical research by assisting researchers in the management and analysis of complex biological and clinical data by purchasing new software and hardware.

To support clinical research, the CTI has purchase software and hardware to advance research in three areas: (1) the conduct of clinical research; (2) research comparing the comparative effectiveness of clinical interventions; and (3) genomics research to support personalized medicine.

1. <u>Conduct of Clinical Research</u>

The implementation of the Clinical Trials Management System (CTMS) is almost complete with the completion percentage at over 95%. The CTMS allows for the monitoring and tracking of multiple clinical studies. Coordinators and managers can log on to the system to check enrollment, budget, timeline and other information about a specific study. Significant progress has been made over the last year with the 2 major phases – Study creation and Participant Tracking – being fully completed. The integration between Penn State's Institutional Review Board and CTMS is complete and they have been tested to make sure that the cross communication and data transfer is seamless. This allows for a one-stop shop for the study coordinators to verify if the clinical study has IRB approval to consent patients. A release from the vendor that proved more complicated in implementing, delayed the schedule by a little. The slower, meticulous implementation of this vendor release was preferred over a hasty one to prevent problems downstream after the full roll out of the system.

2. <u>Comparative Effectiveness Research</u>

Comparative effectiveness research is the research that is aimed at identifying the benefits and harms of the different treatment regimens as designed in different conditions. This is done by querying de-identified data that is kept centrally on different diseases and conditions. The number of data requests to mine our institution's EMR has increased. We have had over 110 requests from departments varying from anesthesiology to pediatric surgery. The information mined has also been diverse in nature and has included topics like cardiac arrest information, umbilical artery dopplers, rib fractures, etc. With this capability, our researchers are able to find

highly specific data in a short amount of time and this speeds up the research at this institution significantly.

3. Analysis of Genetic Data

There are multiple completed and ongoing projects for which genomic data has been extracted from various patients. The researchers are working with the CTI and other staff from the Institute for Personalized Medicine to obtain this information and analyze it. One of the exciting projects involves the analysis of genetic information from identical twins, one who has Parkinson's disease and one who does not. This would allow researchers to identify genetic variations; some that could help explain how Parkinson's disease develops. Some of the other genetic project analyses include identification of candidate genes for diabetes, genome-wide mapping of large organized heterochromatin domains, and analyzing RNA-seq datasets for gene differential expression. This definitely proves that the location of the CTI immediately adjacent to the IPM is allowing for improved collaborations.

Research Project 2: Project Title and Purpose

A Trial to Evaluate the Safety and Tolerability of a Novel Oral Iron Supplement for the Treatment of Iron Deficiency Anemia - The purpose of this study is to conduct a Phase 1 clinical trial designed to evaluate the safety of a novel oral iron supplement utilizing a nutritional strain of yeast. Iron deficiency anemia (IDA) is a problem of world-wide importance that affects approximately 1.3 billion people. IDA has been identified by the World Health Organization as the number one nutritional disorder in the world. The elderly, infants, children, and women of reproductive age are particularly affected by iron deficiency and iron deficiency anemia. In the US, IDA is prevalent in lower socioeconomic groups. The impact on health, mother-child interactions, and days lost from the workplace is widely published. In many countries and among the poor in the US, the primary driver for iron deficiency is the lack of available iron rich foods. Current treatment strategies for IDA are inadequate because of poor absorption rates for iron in oral tablets and significant side-effects leading to non-compliance. The primary objective of this project is to demonstrate the safety and tolerability of a novel oral iron supplement. The secondary objective is to begin to evaluate efficacy and identify a potential therapeutic dose of this novel oral iron supplement.

Anticipated Duration of Project

3/1/2012 – 12/31/2014

Project Overview

Iron deficiency is the most prevalent nutritional disorder in the world and constitutes a public health epidemic that affects more than 3.5 billion people in developing as well as developed countries. Iron deficiency impairs overall health and work productivity, delays motor and mental development in infants, increases maternal hemorrhaging, increases the risk of premature delivery, and is associated with postpartum fatigue and depression. Iron deficiency affects more than a quarter of women following child delivery and in the most severe situation is associated

with maternal or fetal death.

Current approaches to the treatment of iron deficiency are not universally available and contain basic flaws in concept and utilization. The primary objective of this project is to demonstrate the safety and tolerability of a novel oral iron supplement. The secondary objective is to begin to evaluate efficacy and identify a potential therapeutic dose of this novel oral iron supplement.

Aim: Perform a pilot study on iron deficient human subjects to determine safety, tolerability, and iron absorption rates using our oral iron supplement.
The study will begin with 3 patients meeting inclusion criteria. These individuals will receive 300 mg tablets of the novel oral iron supplement. The iron content of these tablets is 30 mg. This dose was chosen as the initial starting dose because it contains the RDA for iron for pregnant women.

This study will utilize the standard "3+3" rule-based dose-escalation scheme, which uses predetermined dose levels and cohorts of three patients. In this study, the following dosages will be used: 300 mg per day (Level 1), 600 mg once a day (Level 2); 600 mg 2X per day (Level 3), 600 mg 3X per day (Level 4-target therapeutic dose equal to current standard of care at 180 mg iron per day). The final dose (Level 5) is 600 mg 4X per day and would be used in extreme cases of iron deficiency. Given that the iron content of the tablets is 10%, individuals receiving the maximum dose in our study are ingesting 240 mg of iron; well below the 1000 mg of iron given therapeutically for conditions of iron deficiency that require intravenous intervention. Since there are 5 different doses, the minimum number of subjects is 3 and the maximum number is 5 times 6 for a total of 30 subjects, each will be involved with the study for 3 months.

Principal Investigator

James R. Connor, PhD
University Distinguished Professor
Penn State University
Department of Neurosurgery
Hershey Medical Center
Hershey PA, 17033

Other Participating Researchers

Michael Evans, MD; Jeffery Pu, MD - employed by Penn State University

Expected Research Outcomes and Benefits

Essential Data Generated:
We will have safety, tolerability and hematologic data on 30 subjects. Specifically we will collect blood samples to obtain the following measurements:

- CBC including hemoglobin and hematocrit,
- Plasma iron, TIBC, and serum ferritin.

- Liver Tests: AST (SGOT), ALT (SGPT), ALP, GGT (GGTP).
- Kidney Tests: Analytes Albumin, BUN/creatinine, chloride, CO2 content, creatinine clearance, glucose, K+, total protein, Na+, protein.

At the end of the funding period we expect to have data that demonstrate we have a safe oral iron supplement. Within the proposed doses that will be examined, we expect to find a therapeutic dose from which we will develop a Phase II/III clinical trial.

The long term benefits to the population are expected to be a safe, efficient and economical method to treat iron deficiency anemia. Furthermore, our mode of iron delivery is a nutritional staple in most cultures of the world so there should be little cultural resistance to consuming our product.

Summary of Research Completed

We are excited to report that we have enrolled our first subject in the clinical study. There are currently 46 candidates who meet criteria that will be contacted at their next clinical visit. We have had 37 screen failures in the past few months. Of these individuals 36 do not meet eligibility criteria. We enrolled the first subject on 6/3/14. This individual has returned for the first evaluation on June 14. In compliance with the conflict of interest data management plan I am not given any information on the enrolled subjects.

The reasons for screen failures are:
29/37 (78.37%) - Hgb / Fer / Fe Sat. levels too high
9/37 (24.32%) – Inflammatory condition
6/37 (16.21%) – Significant cardiovascular disease
4/37 (10.81%) – Asthma and actively using anti-asthmatic tx
3/37 (8.11%) – Anemia not related to IDA
3/37 (8.11%) – Active malignancy within one year
2/37 (5.41%) – Chronic kidney disease
2/37 (5.41%) – Smoker
2/37 (5.41%) – irritable bowel syndrome
2/37 (5.41%) – Psychiatric disease that would compromise dx management
1/37 (2.70%) – Pregnant1/37 (2.70%) – HIV positive

To address the screen failures and increase the number of subjects available to us, we have taken a number of steps:

1) IRB approved the amendment to the protocol's inclusion / exclusion criteria

The amended protocol broadens our potential pool of eligible patients and there are now 46 eligible patients who meet our modified criteria. These candidates are found in the following clinics:

11 or 24% = Family and Community Medicine
16 or 35% = Hematology/Oncology

9 or 20% = Internal Medicine

The remaining 10 are scattered throughout the system: Gynecology, Orthopaedics, Gastrointestinal, Emergency Department, Vascular Surgery, etc.

2) The coordinator for this trial met with the leadership of Family and Community Medicine and reported that "the meeting went very well. They seemed receptive to the study and recruitment strategy proposed". The faculty of F&CM will vote regarding participation in clinical studies that are proposed from outside department investigators and we should know the results of that vote by August 15.

Admittedly, we had anticipated being much further along with the study at this time. The IRB initially approved the original protocol for this study over 2 years ago. However, further full board review was required for changes that we proposed in the dosing schedule (based on new data we obtained in a non-human primate study) and clinical leadership (Dr. Al-Mondhiry who retired was replaced as medical director by Dr. Jeffrey Pu). We submitted the documentation requested to the clinical trials office handling our project in August, 2013 and provided clarification and supplemental information to the IRB following requests in October, November, and January. The last request (in January) required an additional certification of the product post-pilling. We do not perform these analyses in my laboratory to avoid any perception of conflict of interest so it took almost one month to obtain these data and then submit the results to the IRB. We provided the information to the IRB and were advised in April that the study had been approved. Immediately on obtaining this information, we shipped the product with labels to our pharmacy. I was informed on April 21st that the pharmacy was requesting a change to the IRB approved label and thus we had to reprint labels and change bottles. Thus, although we are disappointed that we have not enrolled more subjects, we are very excited that the recruitment of subjects is now underway and that enrollment has started.

Research Infrastructure Project 3: Project Title and Purpose

Research Infrastructure Renovation to Create a Biomedical Research Imaging Core Facility - The overall goal of the project is to establish an imaging core to meet existing and rapidly growing needs for an equipped biomedical research imaging facility dedicated to experimental approaches in the life sciences. We will do this by renovating 2300 sq ft of space to optimally place existing and/or acquired instruments that include (i) confocal, (ii) deconvolution (iii) electron microscopes with support equipment and (iv) a facility director (to be named) to provide expert guidance in optimal sample preparation, supervision of instrument use, data collection and training for researchers and students in imaging and image processing.

Duration of Project

6/1/2011 - 6/30/2014

Project Overview

The objective of this project is to create a Core facility dedicated to biomedical research imaging in the Penn State Hershey College of Medicine. The project aims to: 1) create a newly renovated

2300 sq ft research infrastructure to house existing imaging equipment; 2) centralize the existing Core Facilities shared equipment, which includes confocal and transmission electron microscopes; 3) create a desirable workspace to recruit an experienced Imaging Core Director able to aid investigators in a wide variety of complex imaging projects. The confocal microscopes are configured specifically for live-cell and regular histology studies to meet the rapidly growing need for a variety of sophisticated imaging techniques at the tissue, cell and sub-cellular levels, including time lapse studies, fluorescence resonance energy transmission (FRET) and fluorescent recovery after photobleaching (FRAP). The electron microscopes collect images at the sub-cellular and molecular level, and include a cryo transmission electron microscope (cryoEM) - a JEOL model JEM-2100 that is a 200 kV Transmission Electron Microscope (TEM), with a Lanthanum Boride (LaB6) emitter, a full Cryo package, CCD camera and a stage-tilting capability for tomography. This particular cryo-TEM is one of the easiest to operate and maintain, making it the choice of many cryo TEM facilities as a first instrument to acquire. It functions better than comparable instruments to fulfill the basic needs of a general research environment and is often retained as a training instrument when facilities expand because it has a reputation for versatility in a multi-user environment and is relatively inexpensive to repair and maintain.

Broader Impacts: A wide variety of structural and functional studies will be made possible in a centralized location which can be supervised by an expert Core Director. The instrumentation placed in the optimized space will offer the powerful tools of direct visualization that can guide and complement other research approaches focusing on a variety of health issues including viral infections (including HIV and herpes), cancer, diabetes and its complications (primarily diabetic retinopathy, muscle protein synthesis, kidney and brain complications), heart disease, and degenerative diseases of the central and peripheral nervous system (including Parkinson's disease, Alzheimer's disease and restless leg syndrome).

Principal Investigator

Susan Hafenstein, PhD
Assistant Professor
Pennsylvania State University
Hershey Medical Center
Hershey PA 17033-0850

Other Participating Researchers

Alistair Barber, PhD, Leslie Parent, MD, Bruce Stanley, PhD, Robert Ashley, AAS, Roland Myers, BA, Joshua Yoder, PhD - employed by Pennsylvania State University College of Medicine

Expected Research Outcomes and Benefits

Core Facilities shared equipment has become a major strength of the Penn State Hershey College of Medicine since its inception in1987. Since then the facilities, services and associated expertise have grown dramatically. Expanding and optimizing the Imaging Core will enable the

continuation and enhancement of ongoing research. Equipment will be placed in an adequate facility and infrastructure to function optimally. Microscopes will image rapid events in live cells or explant tissue transfected with fluorescent protein markers, or labeled with metabolic dyes, to determine changes in protein expression, trafficking and redistribution in response to pharmacological, environmental or genetic manipulations. Transmission electron microscopy (TEM) is suitable for simple negatively stained biological sample viewing and data collection for reconstruction as well as thin sectioning, trimming and semi-thin sectioning, polymer and materials science applications. Cryo electron microscopy (cryoEM) provides two additional resources not currently available to Penn State COM: cryoEM single particle reconstruction techniques and tomography.

Research enhanced by live imaging will have broad relevance to human health because it will enable us to study rapid interactions between proteins and other molecules in living cells, in real time. Determining the location and movement of any protein is essential to understanding its function in health and disease. TEM and CryoEM allow three-dimensional (3D) imaging at the subcellular level, filling a gap between the atomic resolution provided by NMR and X-ray, and visualization by light microscopy. Imaging instrumentation will enhance a wide variety of public health projects being conducted at the Milton S. Hershey Medical Center, including cancer, diabetes, heart disease, viral infection and neurodegenerative diseases.

Summary of Research Completed

A) Supported Research Programs by High End Light, Laser Scanning Microscopes and Image Processing Station Located in Microscopy Imaging Core. There are a number of research programs, which are actively using these research methods to resolve research issues in their area of interests. These include, but are not limited to: a) Macrophage-Tumor Cell Fusions from Peripheral Blood of Melanoma Patients (Gary Clawson Lab); Mechanisms of influenza infection in macrophages (Z. Chroneos Lab); Accurate colocalization measurements of HSV proteins expressed in transfected mammalian cells (John Wills Lab); Function and mechanisms of tegument protein UL16 of herpes simplex virus (John Wills Lab); Localization of transfected viral proteins (John Wills Lab); Localization specific protein targets (e.g. MHCI, Klra2, PirB) to specific cell types in different regions of the brain or retina (Bill Freeman Lab); Morphological and quantitative microscopy studies of synaptophysin reactive synapses (Bill Freeman Lab);Location of SREBP in cellular environment (Diane Thiboutot Lab);H2S toxicity in live cellular environment (Phillipe Haouzi Lab); Live changes in mitochondrial morphology (HG Wang Lab); Role of BIF-1 in autophagy and tumorigenesis (HG Wang Lab); Visualizing the intracellular trafficking of retroviral Gag protein and RNA in fixed and living cells (Leslie Parent Lab); Events of rous sarcoma virus assembly in real-time (Leslie Parent Lab); Expression of eYFP in the rat brain neurons/grafted neurons (Thyagarajan Subramanian Lab); Visualization of plasma membrane lipid micro-domains (Mark Kester Lab); Quantifying mu-opioid receptors on cell membranes in the paraventricular nucleus of the hypothalamus (Robert Bonneau Lab); Measure intracellular calcium in renal epithelial cells after exposure to hydrogen peroxide (Brian Reeves Lab); Human papillomavirus (HPV) infection (Craig Meyers); Tissue morphogenesis in Drosophila embryos using various subcellular markers (Afshan Ismat, F&M College); Imaging to look at mammary ducts of the mouse to determine if the transgene Ub-CreER; Confetti +/- is functioning properly in the mammary gland (Ed Gunter Lab); Quantifying eGFP cellular

presence in bones and bone mineral fluorescent labelling 2D lengths (Donahue Lab); Quantifying DNA amount in tumor cells to distinguish diploid vs. tetraploid cells (Gary Clawson Lab);Visualizing and quantifying phenotypes of alveolar macrophages (Joanna Flores Lab); Imaging Armadillo and Engrailed protein levels in fly embryos expressing transgenic constructs (David Roberts, F&M College); Role of the Human Ion Channel TRPM2-L in Mitochondrial Bioenergetics and Tumor Growth (Barbara Miller Lab); Novel Nanoparticle Therapy for Pancreatic Cancer (Gail Matters Lab); Targeting Pancreatic Cancer with Aptamers to the CCK-B Receptor (Clawson and Matters Labs).

There are already several peer reviewed publications reporting images generated from our light microscopy facility. In addition, there are several other manuscripts in preparation or submitted for publication. For instance:

Cleary et al. (2014) Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers *Nature* 508, 113-117;

Bann, D. V., et al. (2014). "A murine retrovirus co-Opts YB-1, a translational regulator and stress granule-associated protein, to facilitate virus assembly." *J Virol* 88(8): 4434-4450;

Lochmann, T. L., et al. (2013). "NC-mediated nucleolar localization of retroviral gag proteins." *Virus Res* 171(2): 304-318;

Masser, D. R., et al. (2014). "Insulin treatment normalizes retinal neuroinflammation but not markers of synapse loss in diabetic rats." *Exp Eye Res* 125C: 95-106;

Nadaraia-Hoke, S., et al. (2013). "Alterations in the MA and NC domains modulate phosphoinositide-dependent plasma membrane localization of the Rous sarcoma virus Gag protein." *J Virol* 87(6): 3609-3615;

Rye-McCurdy, T. D., et al. (2014). "Mechanistic Differences between Nucleic Acid Chaperone Activities of the Gag Proteins of Rous Sarcoma Virus and Human Immunodeficiency Virus Type 1 Are Attributed to the MA Domain." *J Virol* 88(14): 7852-7861;

Stake, M. S., et al. (2013). "Nuclear trafficking of retroviral RNAs and Gag proteins during late steps of replication." *Viruses* 5(11): 2767-2795;

Clawson et al. (2014) Macrophage-Tumor Cell Fusions from Peripheral Blood of Melanoma Patients: Initial Characterizations, *PLOS ONE* (Under review);

Hoffman, N. et al. (2014) A Splice Variant of the Human Ion Channel TRPM2-L Inhibits Mitochondrial Bioenergetics and Tumor Growth, *J. Biol. Chem.* (Under Revision)].

<u>B) Summary of Research Completed with TEM JEOL 1400 and JEOL 2100 (cryoEM).</u>
Annual use of the JEOL1400 + sample preparation lab for thin sectioning and immune-labeling assists 22 Principal Investigators from 15 different departments, at 20 hours per week
The JEOL2100 cryo-electron microscope became operational April 30, 2013. Since then the microscope has been used by 12 different departments and 13 different PIs at the hourly rate shown below.

JEOL2100	Number of Experiments	4-hour cryo Sessions	8-hour cryo Sessions	Labs assisted	Downtime
2013 (12 months)	83	30	27	14	2 months
2014 (7 months)	61	8	10	9	2 months

CryoEM Outreach Program currently involves four independent schools:

1) Berkshire School: Advanced Math/Science Research Program Director Dr. April Burch
2) Gettysburg College**:** Gettysburg Phage Hunter Program directors are Biology Department Chair Veronique Dellesalle and Professor Greg Krukonis.
3) Lehigh University: Professor Vassie Ware heads the SEA, Phage Hunters program at Lehigh
4) Bucknell University: The Sea Phage Hunter program at Bucknell is directed by Department Chair, Marie Pizzorno and Professor Emily Stowe.

Resulting publications:

Shingler KL, Cifuente JO, Ashley RE, Conway JF, Makhov AM, Hafenstein S. The Enterovirus 71 procapsid can serve as an immunological decoy to sequester neutralizing antibodies and rescue virus infection. Under Review PLoS Pathogens, May 10, 2014 (Corresponding Author)

Shingler K, Organtini L, Hafenstein S Enterovirus 71 Virus Propagation and Purification. Bio-Protocols (on line) 2014. (Corresponding Author)

Organtini LJ, Makhov AM, Conway JF, Hafenstein S*, Carson SD*. Kinetic and structural analysis of coxsackievirus B3 receptor interactions and formation of the A-particle. J Virol. 2014 Mar 12. [Epub ahead of print] PMID: 24623425 *last two authors share corresponding authorship.

Scientific Progress:

a) Guan J, Bywaters SM, Brendle SA, Lee H, Ashley RE, Yoder JD, Makhov AM, Conway JF, Christensen ND, Hafenstein S. Epitope analysis of Two Specific Monoclonal Antibodies for Human Papillomavirus 16.

The structures of human papillomavirus type 16 (HPV16) complexed with fragments of antibody (Fab) from two neutralizing monoclonal antibodies (MAbs), H16.V5 and H16.1A, were studied by three-dimensional cryo-electron microscopy. We determined that loops from two adjacent copies of the L1 capsid protein on the capsid surface make up the conformational epitopes; however, additional loops were also found to contribute to the epitope of H16.V5. H16.1A and H16.V5 MAbs and Fabs showed strong neutralization ability in pre- and post-cell attachment neutralization assays. Thus, inhibition is possible via the interaction of the variable domain with the capsid epitope, as well as with a whole antibody related mechanism. The structural analysis combined with biological assays suggest that the mechanism of neutralization used by both Fab/MAb is likely due to the stabilization of capsids that prevents conformational changes essential during the virus infectious pathway. The orientation of the Fab relative to the capsid surface suggests that MAb also may neutralize through crosslinking.

b) Lee H, Brendle SA, Bywaters SM, Guan J, Ashley RE, Yoder JD, Makhov AM, Conway JF, Christensen ND, Hafenstein S. Cryo-EM study reveals the mechanism of HPV capsid stabilization used by neutralizing antibody H16.V5.

Human papillomavirus 16 (HPV16) is a worldwide health threat and an etiologic agent of cervical cancer. Here we present a structural study of HPV capsids and antibody interactions. The cryo-EM structures of HPV16 immature-like and mature capsid particles were solved individually and as complexes with fragment of antibody (Fab) from the neutralizing antibody H16.V5. The immature capsid-Fab complex map showed that binding of the Fab induced

significant conformational changes that imposed more ordered surface loops, rigid invading arm structures, and tighter inter-capsomeric connections at the capsid floor than were seen in the immature particle structure alone. The H16.V5 Fab preferentially bound hexavalent capsomers and the stabilizing effect was directly related with the number of bound Fabs. The structural analysis also showed that the Fab distinguishes subtle differences on antigenic sites. The current study describes the mechanism of the post-attachment neutralization activity of H16V5. Fab.

c) Spadafora, V., Unrath, W.C., Swenson, A. M., Campbell, K.S., and Yengo, C.M., Characterization of the Cardiac Myosin Motor Function in Heart Failure.
Myosins are motor proteins that cyclically interact with actin filaments using the energy from ATP hydrolysis to produce force and function in muscle contraction, cell division, and organelle transport. Mutations within the subdomains of myosin can interfere with the communication necessary for proper motor function. These mutations in cardiac myosin can result in cardiomyopathies, the leading cause of heart failure in young people and are implicated in two main phenotypes: hypertrophic and dilated. Cryo-electron microscopy was used to reconstruct a three dimensional image of the actin and myosin interaction as an approach to characterize the effects of the mutations on actin binding. In order to further examine the impact of heart failure on cardiac myosin, we performed a blind study on samples obtained from heart transplant patients.

d) Shingler KL, Ashley RE, Makhov A, Conway J, *Hafenstein S.* The structure of a picornavirus entry intermediate captured in situ.
Using nanodics we have isolated the first biologically accurate picornavirus entry intermediate. We have successfully used an approach for preparing complexes of membrane proteins for TEM analysis by using of lipid bilayer nanodiscs (100-150Å in diameter) to immobilize membrane proteins. This approach is well suited to the analysis of PTLP and viruses interacting with receptor proteins as we form virus complexes in a biologically significant presentation. The resulting *in situ* approach visualizes the interaction between receptor and virus, and also the resulting dynamic complexes that form in order to eject the viral genome. The nanodiscs can be formed in the presence of full-length receptor, which will insert into the disc as the detergent, *octyl glucoside* (n-octyl-β-D-glucoside), is removed from the solution by dialysis. The composition of the nanodisc lipids can also be changed to mimic specific membranes.

e) Shingler KL, Hegarty J, Sansger W, Ashley R, Stewart D, Hafenstein S. Structural and functional analysis of phage tail-like particles that kill *Clostridium difficile*.
Phage Tail like Particles (PTLP): With David Stewart, a collaborator in the Department of Surgery at Penn State College of Medicine we are studying Phage Tail-Like Particles (PTLP) of C. *difficile* bacteria. The overall research focus is to use a structural approach to complement the ongoing genetic and molecular biology studies to understand the mechanism of selective recognition and penetration of the host membrane by PTLP.

Figure 1**.** Basal mammary tumour cells from a donor mouse (red) expressing a lineage marker (green) comingle with host-derived epithelial cells. The elongated mammary duct on the left retains a normal bilayered architecture in the absence of infiltration by the donor-derived tumour subclone. Tumours often display a complex subclonal organization. In a mouse model of breast cancer initiated by Wnt signaling, *Allison Cleary et al.* show that some tumours are biclonal — composed of basal and luminal clones with distinct genetic alterations. These clones cooperate to maintain tumour growth that is dependent on secretion of Wnt by the luminal cells. On the cover, *Cover: Thomas Abraham/Penn State Imaging Core* [Adapted from ***Nature*** Issue April 03, 2014]

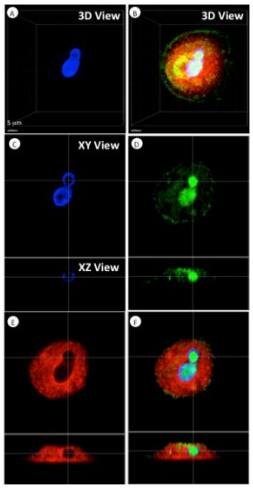

Figure 2**.** Representative 3D confocal images of Circulating Tumor Cells (CTC) stained for nuclei with DAPI (Blue) and fluorescent markers specific for the pro-carcinogenic cytokine MIF (Green) and the pan-macrophage marker CD68 (Red). Nuclei appeared to have "holes" or "tunnels" through them, and these holes/tunnels actually stained strongly for MIF. Panel [A-B]: 3D projections. Panel [C-F]: Sectional views showing holes in the nuclei where pro-carcinogenic cytokine MIF accumulated [Adapted from Clawson et al. ***PLOS ONE*** (Under Review)].

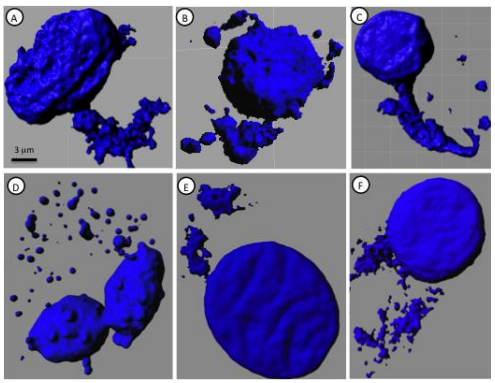

Figure 3. Representative 3D surface rendered confocal images of Circulating Tumor Cells (CTC) stained for nuclei with DAPI or TO-PRO (Blue) featuring depolyploidization of nuclei with the extrusion of DNA from nuclei. There was apparent "Shedding" of DNA from the nuclei into the cytoplasm, which was evident in essentially all of the CTCs. In many cases, this appeared as tubes of chromatin being extruded. DNA Ploidy graph is also shown. The control normal Human Pancreatic Duct Epithelial Cell (HPDE) showed the expected diploid DNA content. The populations of cultured Circulating Tumor Cells (CTC) (from 3 separate patients) showed cells with DNA distribution peaks corresponding to diploid and tetraploid, but with many aneuploidy cells distributed through the range, and some with DNA contents ranging up octaploid or even higher [Adapted from Clawson et al. ***PLOS ONE*** (Under Review)].

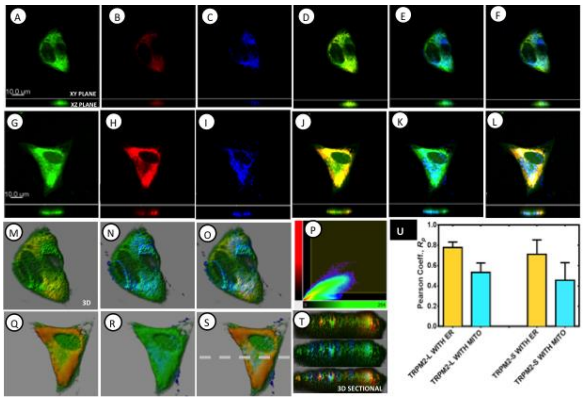

Figure 4**.** Subcellular Localization of TRPM2 was determined with 3D confocal microscopy. Confocal microscopy was performed with SH-SY5Y cells transduced with lentivirus expressing GFP-tagged TRPM2-L or TRPM2-S isoforms (Green), and DS-Red ER to identify endoplasmic reticulum (Red). Mitochondria were detected with antibody to Tom20 and appropriate secondary (Blue). Panel [A-F]: Sectional views (both XY and XZ planes) of TRPM2-L (A), ER (B), mitochondria (C) and the overlays of TRPM2-L with ER (D), TRPM2-L with mitochondria (E) and TRPM2-L with ER and mitochondria (F). Panel [G-L]: Sectional views (both XY and XZ planes) of TRPM2-S (G), ER (H), mitochondria (I) and the overlays of TRPM2-S with ER (J), TRPM2-S with mitochondria (K) and TRPM2-S with ER and mitochondria (L). Panel [M-P]: 3D rendered views of the overlays of TRPM2-L with ER (M), TRPM2-L with mitochondria (N), TRPM2-L with ER and mitochondria (O) and a colocalization plot of the overlay of TRPM2-L with ER (P). [Adapted from Hoffman et al. ***J. Biol. Chem.*** (Under Revision)].

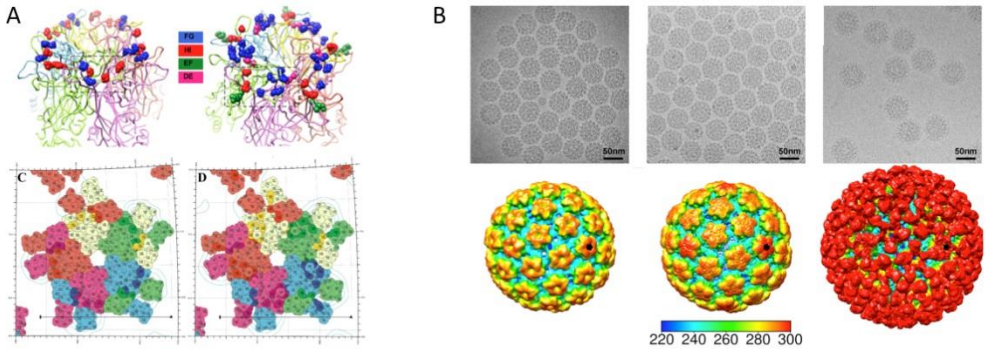

*FIGURE 5(*A) The H16.1A and H16.V5 epitopes mapped to a single HPV16 capsomer (PDB ID 3OAE) (32). Five L1 capsid proteins (hot pink, light salmon, light yellow, pale green, and light sky blue) comprise one capsomer. The inset shows the color code for the virus contact residues (shown as spheres) corresponding to different loops on the virus surface: FG (royal blue), HI (red), EF (forest green), and DE loop (deep pink). The black dotted box shows the epitope of one Fab. The roadmap shows the stereographic projection of a capsomer where the polar angles φ and θ represent the latitude and longitude of a point on the viral surface, respectively. The epitopes of H16.1A (C) and H16.V5 (D) are shown in the darker colors for each corresponding L1 protein (violet red, orange red, gold, lime green, and royal blue). The radial projection of the whole Fab molecule onto the surface of the virion is shown with gray contours representing the height of the projected density. B) 2D data and the 3D maps of HPV and HPV-Fab complex.

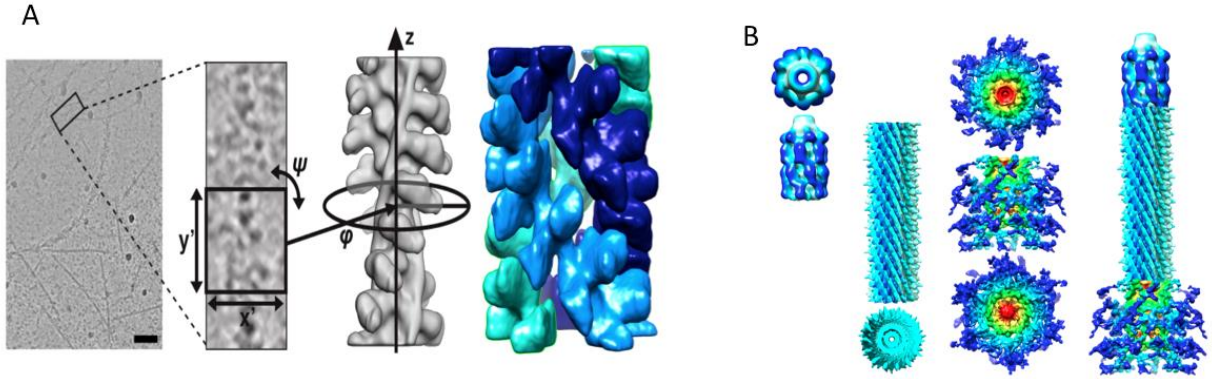

Figure 6 A) From Behrmann et al, 2011: The process used to calculate a helical reconstruction from micrograph to 3-D image and the most recent image of our reconstruction of recombinant myosin in complex with actin. This study investigates a mutant cardiac myosin implicated in several cardiomyopathies. The cardiac myosin mutant was bound to actin and the complex was verified by negative stain transmission electron microscopy. Subsequently, the sample was vitrified and used for cryo-electron microscopy data collection. 19 micrographs were acquired. Segments of actin-myosin complex were selected using helixboxer in EMAN. The selected helicies were reconstructed using IHRSR++ and SPIDER. The resulting preliminary density map shows the motor domain of myosin bound to actin filaments. B) The cap, tail, and baseplate structures of the non-contractile PTLP.

Research Project 4: Project Title and Purpose

Sphingolipid-based Therapeutics for AML - The long term goal is to develop new therapies for the treatment of acute myeloid leukemia (AML). These funds will support three specific aims

that provide additional/ new/ critical data that are essential for submission of an integrated interdisciplinary NIH P01 application in 2011 that will investigate the role of dysfunctional sphingolipid metabolism in acute myeloid leukemia. The funds will support three team projects that will benefit each of the eventual P01 projects.

Duration of Project

5/1/2011 – 6/30/2012

Summary of Research Completed

This project ended during a prior state fiscal year. For additional information, please refer to the Commonwealth Universal Research Enhancement C.U.R.E. Annual Reports on the Department's Tobacco Settlement/Act 77 web page at http://www.health.state.pa.us/cure.

Research Project 5: Project Title and Purpose

Genome-wide Mapping of Transcriptional Factors Involved in Pathogenesis of LGL Leukemia – Large Granular Lymphocyte (LGL) leukemia is characterized by a clonal proliferation of antigen-primed terminally differentiated effector memory cytotoxic T cells (CTL). The overall objective of this project is to understand the pathogenesis of LGL leukemia by identifying the key genes which promote leukemic LGL survival. This project capitalizes on strengths of Penn State University by investigating the pathogenesis of a leukemia discovered by the PI utilizing a genomic approach in collaboration with colleagues at Penn State-University Park. T-bet is considered the master switch for the generation of antigen primed effector CTL. We previously showed that leukemic LGL are characterized by T-bet upregulation. Thus, LGL leukemia provides us with a unique model to study the functional role of T-bet in generalizing and maintaining CTL response.

Duration of Project

7/1/2011 – 6/30/2013

Summary of Research Completed

This project ended during a prior state fiscal year. For additional information, please refer to the Commonwealth Universal Research Enhancement C.U.R.E. Annual Reports on the Department's Tobacco Settlement/Act 77 web page at http://www.health.state.pa.us/cure.

Research Project 6: Project Title and Purpose

Gut Microbial Metabolism and Receptor-Mediated Regulation of the Host Environment - As a whole, the gut microbiota contains nearly 100 times more genes than that in the human genome of which some are used to tightly control and regulate the host environment. In spite of this, little is known about specific gut microbial metabolism, their positive or negative effects on the host, or how microbes use specific diet-derived substrates to modulate the host environment (e.g.,

receptor-mediated regulation of intestinal inflammation). This project seeks to understand and characterize—using advanced metabolomic approaches, mouse models, and high throughput sequencing—the gut microbiota metabolome and how specific diet-derived substrates (e.g., tryptophan) are used to influence and modulate the host environment.

Anticipated Duration of Project

7/1/2011 - 12/31/2014

Project Overview

Broad objective: Metabolomics has generated overwhelming amounts of evidence that the gut microbiota contribute significantly to the host's metabolic phenotype. The studies described herein would generate a detailed view of the host-gut microbiota interaction (through the aryl hydrocarbon receptor [AHR]) and lay groundwork for future studies.

Specific Aim 1: Optimize anaerobic microbial culture and develop high throughput culture systems, along with biofluid isolation, and metabolomics analyses.

Experimental approach: Microbes will be cultured with and without various dietary substrates such as tryptophan. Advanced chromatography-mass spectrometry (LC-MS, GC-MS) coupled with chemometrics and multivariate data analysis will be used to metabolically phenotype each microbial strain.

Specific Aim 2: Generate and identify probable associations between metabolites and gut microbiota changes, utilizing dextran sodium sulfate (DSS) and *C. rodentium* to modulate the gut microbiota.

Experimental approach: Urine metabolomics of control wild-type, control *AHR*-null, DSS-treated (or *C. rodentium*) wild-type, and DSS-treated (or *C. rodentium*) *AHR*-null mice will be conducted to assess any changes in the urinary metabolome that are indicative of changes in the gut microbiota. The cecal contents will be typed for microbes and the inflammation assessed.

Specific Aim 3: Develop analytical and data analysis pipelines that provide a metabolomics platform for the biomedical research community at Penn State University.

Experimental approach: Build a robust pipeline for sample preparation, data generation, and data analysis built around three robust analytical platforms that together allow for the comprehensive analysis of the metabolome of simple and higher organisms. A combination of GC-MS and LC-MS methodologies will be developed that allow for automated data reduction and compound annotation.

Principal Investigator

Andrew Patterson, PhD
Assistant Professor, Veterinary and Biomedical Sciences
The Pennsylvania State University
201 Life Sciences Building
University Park, PA 16802

Other Participating Researchers

Philip B. Smith, PhD, James G. Ferry, PhD, Mark W. Signs, MS, Jeffrey M. Peters, PhD - employed by Penn State University

Expected Research Outcomes and Benefits

The studies outlined in this project will impact human health and disease. These studies may open new insights into antibiotic design and development, and provide unique views of inflammatory bowel disease, particularly from the vantage point of the gut microbiota.

This project will focus specifically on the AHR, given that microbial-derived metabolites of tryptophan, including kynurenic acid, xanthurenic acid, indoxyl sulfate, tryptamine, and indoleacetic acid, are known ligands for this receptor. Furthermore, since the gut has the highest concentration of T-cells and AHR activates potent inflammatory and anti-inflammatory pathways, it seems logical to conclude that AHR may be a primary target of the gut microbiota to modulate the host environment. Interestingly, AHR recently has been shown to be important for the generation of regulatory T cells (Tregs) as well as a specific subset of CD4+ T cells known as T helper 17 (Th17) cells. While Tregs decrease immune reactivity and Th17 cells promote the immune response, it is thought that AHR is important for altering the balance between these two T cell populations. These observations may help to explain the immuno-suppressive effect of the AHR ligand TCDD. Interestingly, the indoleamine 2,3-dioxygenase (IDO) pathway, the rate-limiting step in tryptophan degradation, may be essential for the development of human Tregs at least in culture.

Additionally, the methodologies and pipelines produced in the course of the present study will be instantly available and applicable to numerous other existing and planned biomedical research projects that would benefit from the generation of metabolomic data.

Summary of Research Completed

Over the last year, we have continued our work on understanding the impact of modulating the microbiome for the prevention and/or amelioration of diet-induced obesity and complications such as non-alcoholic fatty liver disease (NAFLD). We expanded our current understanding of the interaction of the farnesoid X receptor (FXR) with the gut microbiome by using combinations of ultra-high pressure liquid chromatography (UHPLC) coupled with mass spectrometry (LC-MS) and ^{1}H nuclear magnetic resonance spectroscopy (^{1}H NMR) to further elucidate the signaling mechanisms that exist between the gut and the liver. We have identified

pronounced changes in bacterial fermentation (reduced with tempol treatment in control and high fat diet-fed mice) that may help to reduce and/or eliminate the development of NAFLD. This work is currently under revision for publication. We have recently purchased a new Astrios Flow Cytometer Cell Sorter to gauge populations of inflammatory cells following exposure to treatments that modulate the host microbiome.

We continue to focus on understanding how tempol either directly or indirectly influences the *Lactobacillus spp* in the guts of mice. Both in vivo and in vitro, tempol causes a pronounced reduction in *Lactobacillus johnsonii* growth and this effect is entirely dependent on anaerobic conditions. We are currently working to develop mutant strains of *Lactobacillus spp* that lack bile salt hydrolase activity, an enzyme involved in bile salt deconjugation, as we suspect this enzyme is key to the tempol story. These studies are currently ongoing and we are feeding mice *Lactobacillus johnsonii* and recording its influence on bile acid levels as well as weight modulation.

We have shifted our focus from the DSS model to focus more on the impact of environmental contaminants on the gut microbiome. We have chosen this model as it is much more timely and relevant to the current obesity research we have been conducting with tempol. Alteration of the gut microbiota through diet manipulation and environmental contaminants may disturb physiological homeostasis, leading to various diseases such as obesity and type 2 diabetes. Since most exposure to environmentally-persistent organic chemicals occurs through the diet, the host gastrointestinal tract and commensal gut microbiome are likely to be exposed, in some cases chronically, to low doses of persistent organic pollutants (POPs). We reported that 2,3,7,8-tetrachlorodibenzofuran (TCDF), a persistent environmental contaminant, profoundly impacts the gut microbiome and the host metabolism in an aryl hydrocarbon receptor (AHR)-dependent manner. Wild type and *Ahr*$^{-/-}$ male (6-week-old) C57BL/6J mice (5-6 per group) were treated with and without TCDF through the diet for five days at 24 µg/kg body weight. Combined 16S rRNA metagenomics, ^{1}H NMR metabolomics, targeted UHPLC-LC-MS and biochemical assays were used. Dietary TCDF altered the gut microbiome by preferentially shifting the ratio of *Firmicutes* to *Bacteroidetes* (Fig. 1) in an Ahr-dependent manner. TCDF-treated mouse cecal contents were enriched with genus *Butyrivibrio* and *Flavobacteria*, but depleted in the genus *Clostridia* and *Oscillobacter* in comparison with the vehicle-treated mice. Changes in the gut microbiome were associated with altered bile acid metabolism by TCDF through enterohepatic circulation (Fig. 2). Further, dietary TCDF inhibited the FXR signaling pathway and triggered significant inflammation and host metabolic disorders involving activation of bacterial fermentation, hepatic lipogenesis, gluconeogenesis and glycogenolysis, in an AHR-dependent manner (Fig 3). These findings provide new insights into the biochemical consequences of TCDF exposure involving the alteration of the gut microbiome, modulation of nuclear receptor signaling, and disruption of host metabolism.

For Aim 3 we have established a fully functional, state-of-the-art metabolomics facility that is widely-used by undergraduates, graduates, post-doctoral fellows, and principal investigators. While data analysis is not yet fully automated, we have established numerous data analysis workflows for various biofluids and tissue extracts as well as analysis pipelines specific for lipidomics. These efforts are ongoing and constantly evolving to best suit the needs of the facility users.

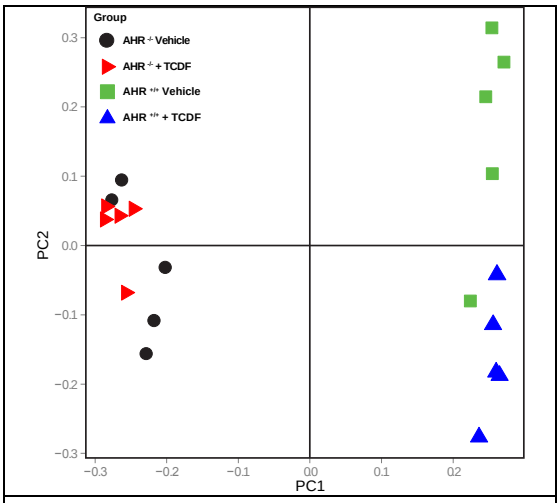

Fig. 1. Mice were treated for 5 days with dietary TCDF (24 µg kg^{-1}). (A) Principal coordinate analysis of the total population of the gut microbiome shows the TCDF effects are Ahr-dependent.

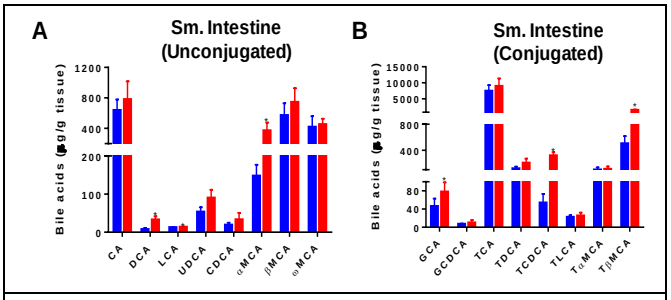

Fig. 2. Quantification of specific bile acids (conjugated shown in panel A, unconjugated shown in panel B) levels in intestinal tissue and feces throughout the enterohepatic circulation of vehicle and TCDF-treated (24 µg kg^{-1}).

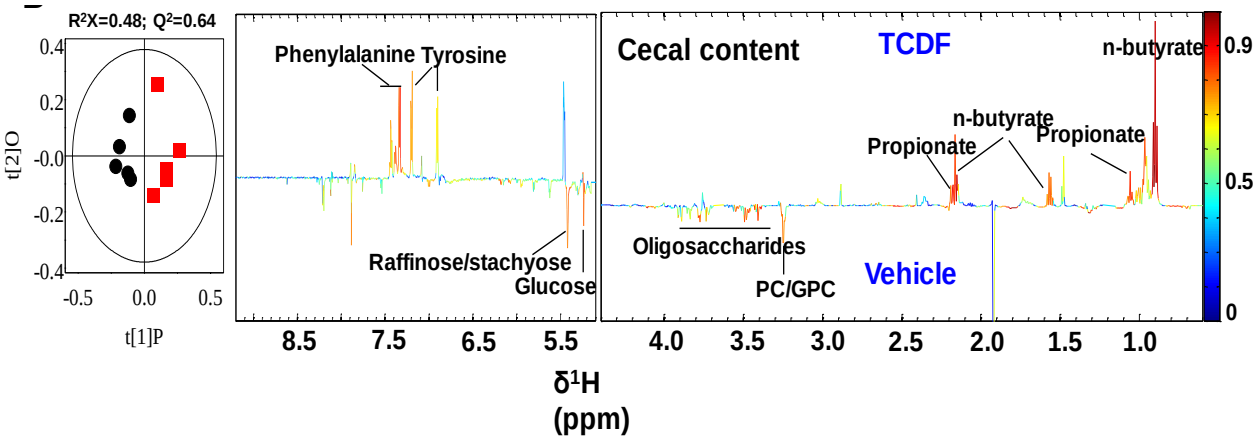

Fig. 3. Dietary TCDF Impacts host metabolism and Activates the Bacterial Fermentation Process. OPLS-DA scores (left) and coefficient-coded loadings plots for the models (right) from NMR spectra of aqueous cecal content.

Research Project 7: Project Title and Purpose

Neurogenic Actions of Dietary Salt on Sympathetic-Cardiovascular Function: The purpose of this project is to identify how dietary salt acts within the central nervous system to adversely affect cardiovascular function and morbidity. Compelling evidence from clinical studies indicates excess dietary salt intake contributes to the pathogenesis of several cardiovascular

diseases and increases morbidity. Recent observations from our research team indicate the central nervous system contributes to the adverse cardiovascular effects of dietary salt. Therefore, this project will (1) identify how the central nervous system detects changes in dietary salt intake, (2) determine how dietary salt affects sympathetic-cardiovascular circuits within the central nervous system, and (3) determine whether dietary salt intake affects sympathetic-cardiovascular regulation differently in males versus females.

Duration of Project

1/1/2011 – 12/31/2012

Summary of Research Completed

This project ended during a prior state fiscal year. For additional information, please refer to the Commonwealth Universal Research Enhancement C.U.R.E. Annual Reports on the Department's Tobacco Settlement/Act 77 web page at http://www.health.state.pa.us/cure.

Research Project 8: Project Title and Purpose

Resolving the Structure and Function of Proteomes - The project proposes to: (i) develop/refine computational methods capable of accurate assessment of domain boundaries, conservation profiles for predicting mutants, and fold-recognition and atomic-resolution structural models for proteins from large screens, (ii) integrate/synergize these computational endeavors with the updated mass-spectrometry (MS) core-center for high-throughput identification and computational characterization of proteins obtained from wet-lab experiments and (iii) build the infrastructure for interdisciplinary training for investigators/students involved in these endeavors. This is the next step towards an interactive, high-throughput, collaborative, and interdisciplinary Center for Computational Proteomics.

Anticipated Duration of Project

1/1/2011 – 12/31/2014

Project Overview

The overall goal of this project is to develop a world-class proteomics center at Penn State. By bridging the current strengths at PSU in computational biology (The Center for Computational Proteomics (CCP)) and the existing mass-spectrometry (MS) core-facility, an interdisciplinary center will be formed capable of high-throughput computational proteomics.

Specific Aim 1: Develop/Refine high-throughput computational approaches capable of accurate assessment of protein characteristics. Results obtained from the recent Critical Assessment of Techniques for Protein Structure Prediction (CASP) competition by CCP indicate that the Adaptive-BLAST algorithm developed by CCP is capable of performing high-throughput structural homology modeling; overcoming a major hurdle to proteomics research. However, proteins that have less than 30% identity with known protein folds are still extremely difficult to

accurately predict and model. Our preliminary data demonstrates that with sufficiently large and divergent position specific scoring matrix (PSSM) libraries, that protein folds can be predicted with >90% accuracy. To obtain acceptable accuracy levels for all structural folds, structural libraries comprised of divergent fold-PSSMs will be generated and benchmarked. Computational and wet lab experiments will be performed to determine the accuracy of the libraries. CCP will also interact with the MS Center to develop all the necessary tools and infrastructure (i.e. website, wiki-base, etc.) such that the resources developed during the course of this project will be open access.

Specific Aim 2: Outfit the current MS core-facility with cutting-edge instrumentation/improved data collection services, and synergize the pipeline with our computational protein models. To excel at proteomics research, the technology within the current MS center must be updated. We plan to purchase a Thermo LTQ Orbitrap Velos (ETD) mass spectrometer, a Dionex Ultimate 3000 nanoflow LC, and a Proxeon ion source; all of which are currently top-of-the-line instruments that are capable of identifying thousands of proteins from a single sample. *Most importantly, we will work closely with the experimental biologists using the MS facility to translate our predictions to the benchtop. In these experiments we will critically evaluate our assessment of soluble protein fragments, ligand-binding pockets (e.g. protein-protein, protein-lipid, and/or protein-pharmacophore binding), and putative loss-of-function or gain-of-misfunction mutants. These combined aims will result in a pipeline which provides researchers with protein identification and cutting edge computational models*.

Principal Investigator

Damian B. van Rossum, PhD
Research Associate Professor
Pennsylvania State University
518 Wartik Labs
University Park, PA 16802

Other Participating Researchers

Andriy Anishkin, PhD, Jagadish Babu, MS, Nigel Deighton, PhD, Tatiana N. Laremore, PhD - employed by The Pennsylvania State University
Kit S. Lam, MD, PhD, Randen L. Patterson, PhD, Sree V. Chintapalli, PhD, Gaurav Bhardwaj, PhD - employed by University of California-Davis, Sacramento CA

Expected Research Outcomes and Benefits

It is expected that developing a high-throughput MS facility and coupling it to high-throughput computational analysis of proteins at the sequence level, will result in a world-class center capable of asking research questions of the highest quality. Typical questions asked of MS include: (i) What protein(s) bind to my ligand of interest, and/or (ii) What changes in expression/localization result from a particular cellular condition? *The output of MS is the identification of a set of protein coding sequences. If the sequences are not annotated sufficiently, the biochemical characterization of these proteins can be a slow arduous process.*

Typical outputs of our computational algorithm include: (i) domain boundaries for protein sequences, (ii) conservation profiles for predictions of putative mutants, (iii) quantitative statistics for putative functions and structural fold(s) and isolation of templates for structural modeling. Together, these approaches give added value to each other and help inform the experimental scientist. During the course of this project we will: (i) determine the best methods for MS sample preparation, collection, and measurements; (ii) refine, evaluate, and improve the accuracy of our method for protein homology modeling; (iii) create transparent protocols for all processes and computational databases/interfaces to inter-relate these data. The potential benefits from this project are broad and have immediate impact. From pharmacophore development to gene therapy and beyond, the tools and initiatives being developed during this project will assuredly promote Penn State as a leader in computational proteomics. At the same time, this project will provide our faculty and students with state-of-the art tools and computational support. We expect that this will also translate to both Center grants and increased R01s for our faculty. Importantly, without these improvements, obtaining Proteomic Center grants are simply outside our reach. Upon completion of this project, we expect to be highly competitive for NSF Training and Center grants, as well as Keck grants and other awards.

Summary of Research Completed

Specific Aim 1- Over the last reporting cycle we have made advances in evolutionary, structural, and functional inferences of ion channels. Taken together, these computational models are important visualization and analytical tools which helped guide the basic science and speed the discovery process. Specifically, we completed and published a collaborative project, described below and highlighted in the Huck Institutes News Archives:
http://www.huck.psu.edu/about/news-archive/human-heart-beats-using-nearly-billion-year-old-molecular-mechanism

Martinson AS, van Rossum DB, Diatta FH, Layden MJ, Rhodes SA, Martindale MQ, Jegla T. Functional evolution of Erg potassium channel gating reveals an ancient origin for IKr. Proc Natl Acad Sci U S A (2014) Apr 15;111(15):5712-7. PMID: 24706772 (TSF supported)

Mammalian Ether-a-go-go related gene (Erg) family voltage-gated K(+) channels possess an unusual gating phenotype that specializes them for a role in delayed repolarization. Mammalian Erg currents rectify during depolarization due to rapid, voltage-dependent inactivation, but rebound during repolarization due to a combination of rapid recovery from inactivation and slow deactivation. This is exemplified by the mammalian Erg1 channel, which is responsible for IKr, a current that repolarizes cardiac action potential plateaus. The *Drosophila* Erg channel does not inactivate and closes rapidly upon repolarization. The dramatically different properties observed in mammalian and *Drosophila* Erg homologs bring into question the evolutionary origins of distinct Erg K(+) channel functions. Erg channels are highly conserved in eumetazoans and first evolved in a common ancestor of the placozoans, cnidarians, and bilaterians. To address the ancestral function of Erg channels, we identified and characterized Erg channel paralogs in the sea anemone *Nematostella vectensis*. *N. vectensis* Erg1 (NvErg1) is highly conserved with respect to bilaterian homologs and shares the IKr-like gating phenotype with mammalian Erg channels. Thus, the IKr phenotype predates the divergence of cnidarians and bilaterians. NvErg4 and *Caenorhabditis elegans* Erg (unc-103) share the divergent *Drosophila* Erg gating phenotype.

Phylogenetic and sequence analysis surprisingly indicates that this alternate gating phenotype arose independently in protosomes and cnidarians. Conversion from an ancestral IKr-like gating phenotype to a *Drosophila* Erg-like phenotype correlates with loss of the cytoplasmic Ether-a-go-go domain. This domain is required for slow deactivation in mammalian Erg1 channels, and thus its loss may partially explain the change in gating phenotype.

The subunit structure of typical Ether-a-go-go superfamily channels consists of a transmembrane channel core sandwiched between two intra-cellular gating domains: an N-terminal eag domain containing a Per-Arndt-Sim (PAS) protein interaction domain and a cyclic nucleotide binding homology domain (CNBHD) in the C-terminus. Structural analysis of Eag1 has revealed that the eag domain, which is required for slow deactivation in human and mouse Erg1, docks on the CNBHD. An eag domain has not been identified in DmErg and CeErg, suggesting that loss of eag domain function might explain the fast-deactivating DmErg-like gating phenotype. We found no evidence for eag domains nematode Ergs or NvErg2-5 in verified or predicted coding sequence or genome sequence. However, DmErg possesses a domain resembling the eag domain, but with reduced sequence homology. Surprisingly, two other arthropod species, *Anopheles gambiae* (mosquito) and *Daphnia pulex* possess Erg orthologs with highly conserved eag domains. Secondary structure analysis predicts a consistent pattern of alpha helices and beta sheets in eag domains that is selectively disrupted in the DmErg eag domain, suggesting that the domain has degenerated. We reasoned that functional loss of the eag domain would reduce selection pressure on CNBHD residues proposed to form a docking site for the eag domain. To test this hypothesis, we aligned the CNBHD of various Ether-a-go-go superfamily channels with Eag1 to examine the relative conservation of predicted interface residues and non-interface residues. Briefly, the multiple sequence alignments of the eag domain (22 taxa from Eag, Elk, and Erg subfamilies) and CNBHD (26 taxa from Eag, Elk, and Erg subfamilies) were generated in Jalview v2.8 by T-Coffee and Muscle, respectively, under default settings and manually edited. Pairwise estimates of evolutionary divergence between sequences were calculated in MEGA5 using number of amino acid differences/site. Ambiguous positions were removed for each sequence pair. A distance matrix was constructed for the entire eag domain using 192 positions. Separate distance matrices were constructed for 27 predicted eag-interacting CNBHD residues and non-interacting CNBHD residues (118 positions), as defined in a co-crystal structure of the Eag1 eag and CNBHD domains. Based on these sequence distance matrices, we quantified mean ± SEM for multiple pairwise comparisons between or within various sequence groups. In-group or between-group comparisons were statistically evaluated with an unpaired two-tailed T-Test.

In mammals, interface residues were significantly more conserved between the Erg, Elk and Eag families. Both interface and non-interface residues are highly conserved in *Daphnia* and *Anopheles*, two protostome species with conserved eag domains. However, CNBHD interface residues in nematode Ergs (no eag domain) and DmErg (de-generate eag domain) are significantly less conserved than non-interface residues. These results support the hypothesis that eag domain presence has a positive influence on the conservation of CNBHD interface residues, and that the eag domain of DmErg is vestigial. Comparisons of *Nematostella* and mammalian Erg sequences show a similar pattern. Conservation of the entire CNBHD is higher for NvErg1 (which has an eag domain) than for NvErg2-5 (no eag domain). Overall conservation is reduced between *Nematostella* and mammals due to early evolutionary divergence. Fast deactivation

therefore appears exclusively in Erg channels that lack the eag domain: DmErg, CeErg and NvErg4.

Specific Aim 2- Currently, there are more than 50 active Orbitrap users from Penn State and beyond, and this number is steadily growing. Orbitrap is being used around the clock with short breaks for routine maintenance and calibrations. In addition to the ongoing collaborations and routine analyses, the following collaborations have been initiated during the past year between the Core facility that hosts the Orbitrap and Penn State faculty, described below, with several manuscripts in preparation.

PI: Bernhard Luscher, The role of GODZ mediated palmitoylation in controlling neuronal inhibition. This project investigates the role of the post translational modification palmitoylation in modulating neuronal signaling in the mouse brain. Dr. Luscher's doctoral student Casey Kilpatrick uses Orbitrap for comparing the abundance of palmitoylated proteins in wild type mice and mice in which one of the enzymes that is responsible for palmitoylation (GODZ/DHHC3) is knocked out. Using TMT labeling of palmitoylated subset of mouse brain proteins, Luscher and co-workers have identified several candidate substrates of GODZ by analyzing which proteins were less abundant in the knock-out samples.

PI: Kenneth Keiler, Antibiotic development. Orbitrap's high resolving power and high mass accuracy was critical in determining the sequence of cyclic peptide with antimicrobial properties. Cyclic peptides can be produced in bacteria using a technology developed in the Benkovic lab at Penn State called Split Intein Circular Ligation of Proteins and Peptides (SICLOPPS). From a library of randomized cyclic peptides, Keiler and his doctoral student Shaima El-Mowafi selected molecules that inhibit tagging by tmRNA or proteolysis of the tagged proteins. These cyclic peptides inhibit the reaction *in vitro* and kill *C. crescentus* and *Bacillus anthracis* cells when added to growing cultures. An antimicrobial nonapeptide from a library with a variable amino acid at 6 position was analyzed by Orbitrap FTMS and MS/MS and found to contain lysine at 6 position.

PI: Manuel Llinas, Center for Malaria Research (CMaR) at Penn State. Joana Santos, a postdoctoral researcher in Llinas group is using Orbitrap for identifying protein-protein interactions between members of the ApiAP2 family of DNA binding proteins and other nuclear proteins in *Plasmodium Falciparum*, the causative agent of malaria. Joana has determined the protein interaction partner of one of the APiAP2 proteins and has found a complex of proteins involved in transcription regulation of genes required for host cell invasion based on data obtained on Orbitrap. Valerie Crowley, another postdoc in Llinas group is trying to identify proteins that bind a DNA-binding *P. Falciparum* protein, PFL1085w which is involved in the sexual conversion of the parasite. This information will be used to regulate PFL1085w and/or its binding partners.

PI: Jeffrey Sundstrom, Hershey Medical Center. A mouse model of diabetic retinopathy is used to determine the mechanisms of this disease at the protein level. Orbitrap is used for differential proteomics analysis of healthy and diseased mouse retinas.

PI: Squire Booker, Antibiotic resistance. Antibiotic resistance remains a major health threat.

Recently, a widespread bacterial protein called Cfr was discovered that conferred resistance to seven classes of antibiotics including linezolid, considered an antibiotic of last resort. Cfr employs a unique mechanism to methylate the electrophillic C8 position of adenosine 2503 of the 23 rRNA that is located in the PTC of the ribosome. The Booker laboratory has trapped and characterized several intermediate states of Cfr during catalysis. Using the Orbitrap MS, Booker and his team were able to prove and characterize the existence of several states, including a unique disulfide species that is proposed to form during turnover.

PI: Kathleen Postle, Identification of trapped interaction sites between *Escherichia coli* TonB and FepA. Michael Gresock, a Ph.D. student in Postle lab is using Orbitrap to study the *Escherichia coli* TonB system which is important for the uptake of important, scarce nutrients such as iron siderophores across the outer membrane of Gram-negative bacteria. TonB utilizes energy derived from cytoplasmic membrane proton motive force to interact with outer membrane transporters such as FepA to allow passage of iron-siderophore complexes into the periplasmic space. The TonB system is a virulence factor for Gram-negative bacteria. To study how TonB and FepA interact *in vivo*, photo-crosslinked TonB-FepA complexes are purified and analyzed using Orbitrap to determine sites of interaction. This will allow identification of *in vivo* interactions between the two proteins that are not apparent in the crystal structures, but which are known to exist.

Research Project 9: Project Title and Purpose

Genomics of Amyotrophic Lateral Sclerosis – Amyotrophic lateral sclerosis (ALS), or Lou Gehrig's disease, is a neurodegenerative disease in which the nerves that control voluntary muscle movement die, causing weakness, paralysis, and death, usually from respiratory failure, within 3-5 years of symptom onset. The disease can be inherited in a small number of cases. About a half dozen genes have been identified that can cause ALS in these families, but about half of the patients with familial ALS and the vast majority of patients with sporadic ALS do not have known genetic mutations. Our goals are to recruit patients with sporadic and familial ALS from the Penn State ALS clinic and to analyze their DNA for known and new genetic mutations associated with ALS that may aid in diagnosis and may serve as new targets for drug discovery.

Anticipated Duration of Project

5/29/2013 – 12/31/2014

Project Overview

Amyotrophic lateral sclerosis (ALS), or Lou Gehrig's disease, is a devastating neurological disease for which there is no effective therapy. Approximately 5-10% of cases are inherited, while a genetic basis for the remaining cases has not yet been determined. The primary objective of this project is to further define the genetic architecture of ALS. Our goal is identify new mutations and new genes associated with ALS. *Our specific aims are to 1) Establish processes for consenting ALS patients and collecting biological samples (e.g., saliva, blood, etc.) for sequence analysis; 2) Conduct targeted sequence analysis of known familial ALS (FALS) genes in patients recruited in aim 1; 3) Conduct whole genome/exome sequencing on selected patients*

who do not have mutations in known FALS genes. The Penn State Institute for Personalized Medicine (IPM) will facilitate recruiting strategies and will work closely with the ALS clinic in consenting participants, data capture, and sample acquisition. DNA will be isolated from biological samples and subject to next generation sequencing in the IPM Genomics Core facility. Analysis of DNA sequence data will be performed in the Bioinformatics core of the IPM. The discovery of new mutations in known genes will be analyzed in the context of detailed phenotype data collected by the PI and his clinic staff to identify possible clinical heterogeneity. New genes identified will be targets for future drug discovery efforts.

Principal Investigator

Zachary Simmons, MD
Professor of Neurology and Humanities
The Pennsylvania State University
500 University Drive
Hershey, PA 17033

Other Participating Researchers

Glenn S. Gerhard, MD; James Connor, PhD, James R. Broach, PhD; Lisa Schneper, PhD; Kathryn Sheldon, PhD; employed by Penn State University College of Medicine

Expected Research Outcomes and Benefits

This project aims to identify mutations in both known genes and new genes related to Amyotrophic Lateral Sclerosis (ALS), or Lou Gehrig's disease, a neurodegenerative disease characterized by progressive degeneration of motor neurons of the brain, brainstem, and spinal cord. The identification of known mutations will allow for fine-grained analysis of the clinical spectrum of ALS in relation to specific mutations in order to determine whether subclinical heterogeneity exists. The identification of new genes will enable studies into disease mechanism and may provide new insight into the pathogenesis of ALS. New genes will also serve as targets for the development of pharmaceuticals, permitting targeting of subtypes of ALS based on genetics and pathogenesis, in place of the current crude therapeutic trials, which are aimed broadly at the ALS phenotype, and have generally been unsuccessful. To date, there is only one FDA-approved drug for ALS (riluzole) which extends lifespan, on average, 60-90 days.

Summary of Research Completed

The primary focus of this research project is to identify the genetic underpinnings of amyotrophic lateral sclerosis (ALS), particularly in those patients who present with no family history of the disease. The work is conducted as a collaborative effort between the ALS clinic within the Penn State Hershey Neuroscience Institute ("Clinic") and the Institute for Personalized Medicine ("IPM"), in which the Clinic identifies, characterizes and treats ALS patients and the IPM consents patients to participate in research addressing the causes and treatment of ALS. The IPM also arranges for blood and saliva collection from patients and, where appropriate, patients' relatives and then conducts genomic analyses of those samples.

Investigators in the Clinic and the IPM together analyze the genomic information on patients and correlate that information with disease presentation. Results from this study should yield insights into the diverse mechanistic bases of the disease and thereby pinpoint the therapeutic approaches that might prove efficacious and identify which patients might respond to which treatment.

To achieve this end, we proposed three specific aims to be completed under support of this grant and have made significant progress in all aims:

Aim 1. Establish processes for consenting ALS patients and collecting biological samples (e.g., saliva, blood, etc.) for sequence analysis.

We have established a seamless process for consenting patients and collecting samples. We have embedded consenters in the Clinic to meet with and consent patients under an approved IRB protocol. We then either add a request for an extra blood sample to be drawn in the course of their normal clinical blood work or provide patients or other family members with saliva kits to collect saliva samples. Our consent rate in the clinic has been remarkably high, with greater that 95% of patients agreeing to participate in this study. As of August 18, 2014, we had consented and collected samples from 235 patients along with 60 family members comprising 40 different families.

Aim 2. Conduct targeted sequence analysis of known familial ALS (FALS) genes in patients recruited in aim 1.

The number of genes identified as causative of familial ALS has recently grown to >20 for those rigorously associated by linkage analysis and >100 suspected to be associated by indirect methods. Consequently, the cost of specifically sequencing the known genes in a patient's genomic DNA sample is greater than the cost of sequencing the entire exome. Accordingly, we have eschewed this approach to studying ALS genetics in our cohort and have rather gone directly to exome sequencing for all patients. Thus, this aim has been subsumed into Aim 3.

Aim 3. Conduct whole genome/exome sequencing on selected patients who do not have mutations in known FALS genes.

Previous observations have established that only 5-10% of ALS patients have a family history of the disease, which means that >90% of patients have no near relatives who have presented with the disease. However, twin studies have estimated that ALS heritability to be 60-70%, suggesting that many patients who present with sporadic ALS (sALS) may have an underlying genetic cause. Accordingly, we have focused on patients with apparently sporadic forms of the disease to determine whether unidentified genetic factors are contributing to the disease.

<u>fALS</u>. As a prelude to conducting studies of sALS patients, we conducted exome sequencing of two fALS patients. We sequenced the complete exomes of both patients and then examined the 20 or so genes previously associated with ALS. In both cases, we identified a mutation that had previously been linked to ALS. In one case, the mutation resided in one copy of the patient's SOD1 gene and created a D to A transition at amino acid position 80. This mutation had

previously been linked to ALS and, in particular, to a particularly indolent form of the disease. In examining the medical record of this patient, we find that the disease had indeed progressed relatively slowly, with the patient still alive with reasonable lung function eight years after diagnosis of the disease. This observation suggests that genomic analysis of fALS patients could, in some cases, provide prognostic information on the progression of the disease, information that could be valuable to the patient and the patient's family. In the second patient, we identified a six base-pair deletion, resulting in a two amino deletion of glycines 223 and 224. This deletion has also been previously linked to fALS. These results confirmed our capabilities to identify causative mutations in ALS patients.

c9ORF72. A recent report suggested that a hexanucleotide repeat expansion in the c9ORF72 gene was associated with a significant number of fALS and sALS patients. Accordingly, we developed an assay to determine the repeat length of this repeat element and then applied that assay to a number of our ALS patients' genomes. We have found 2 examples of repeat expansion in 137 genomes examined. This is significantly lower than the 33% of repeat expansions in fALS patients and 10% of repeat expansions of sALS patients previously reported. We conclude that our patient population is significantly distinct from those analyzed previously or that the reported percentage is an over estimation of the natural occurrence.

sALS trios. An effective mechanism for identifying mutations that might contribute to the onset of disease in patients who have no family history of the disease is to compare the sequence of the genome of the affected patient to that of their parents. This analysis identifies mutations that might have arisen de novo in the patient and which might affect a gene in a way that leads to ALS. This analysis also identifies genes in which the patient inherited different defective alleles from each of his parents but for which each parent was heterozygous and thus carried one normal version of the gene. We would assume that in this case the loss of function of the gene in question would be deleterious but loss of only one of the alleles in an individual would not be deleterious. Such a situation, referred to as compound heterozygosity, would appear to be a case of sporadic ALS, since neither family would exhibit a history of ALS, since all family members would likely carry at least one normal copy of the gene. To accomplish this goal, we have collected samples from seven patients and both of their parents and completed exome sequencing of all twenty-one individuals. We have completed analysis of three of these trios and the results are discussed below.

Trio 1. We identified no *de novo* mutations in this patient but we did identify compound heterozygous variants in two genes, whose known functions suggest a possible involvement in ALS. One gene, DYNC2LI1, encodes the light chain of dynein type 2. Mutations of dynein have been shown to produce motor disorders similar to ALS. Moreover, the product of DYNC2LI1 has been shown to interact with SOD1, the product of gene mutations which are linked to fALS. The second gene, COL6A1, encodes collagen VI. Mutations in this gene have been associated with myopathy disorders. Finally, using array based technology, we identified in the patient a de novo 3 Mb deletion in one homolog of chromosome 13, which spans at least thirteen genes. One of those genes encodes a heat shock protein that associates with mutant SOD1 in the spinal cord of transgenic mice. Thus, analysis of this trio revealed several possible causative genes of sALS.

Trio 2. We identified compound heterozygosity in two genes in this patient, NRP2 and RP1L1. NRP2 encodes neuropilin 2, which may play a role in cardiovascular development, axon guidance and tumorigenesis. RP1L1 is involved in retinitis pigmentosa and encodes a protein that binds to microtubules. However, its expression is limited to the retinal neurons, which are not affected in most cases of ALS.

Trio 3. We identified compound heterozygosity in two genes in this patient, MYO7A and B4GALNT3. MYO7A encodes myosin 7, mutations of which have been linked to myodystrophies. B4GALNT3 encodes an enzyme that transfers N-acetylgalactosamine onto glucosyl residues to form N,N'-diacetyllactosediamine (terminal structure of cell surface N-glycans).

Research Project 10: Project Title and Purpose

Investigation of Molecular Contributors to Early Onset Parkinson's Disease in Identical Twins - Identification of subtle genetic differences in groups of identical twins (one with Parkinson's disease and one healthy) will allow the identification of previously unknown mechanisms that may underlie disease pathology. This knowledge will then allow the investigators to have insight into the cause of the disease. As Parkinson's disease is considered to be caused by a combination of genetic and environmental factors, identical twins provide a very well controlled platform for the investigation of the molecular differences, hence the possible cause of the disease. Identifying subtle genetic alterations in the affected twin(s) will allow to us to obtain information about the changes that resulted in the formation of Parkinson's disease in the affected individual.

Anticipated Duration of Project

5/29/2013 – 12/31/2014

Project Overview

The main goal of this project is to identify novel genetic causes for Parkinson's disease. In order to achieve this goal, we will perform Next Generation Sequencing that will allow the sequencing of the whole genome of selected subjects. The subjects will be pairs of identical twins, with one twin in each pair having been diagnosed with early-onset Parkinson's disease and the other being healthy. Our approach will provide us with enough technical capability to discover any acquired variations in the genomes of the affected individuals. The subjects are currently being followed by a movement disorder specialist who is part of this project; hence we have myriad clinical data regarding our subjects. As these subjects are also part of an ongoing project, their blood samples are currently stored. The genomic DNA will then be sequenced by deep sequencing. The automated Apollo 324 library preparation system and Illumina MiSeq platform will be employed for sequencing. Several bioinformatic platforms will be utilized for the analysis of the genomic sequences (Galaxy, BaseSpace).

Principal Investigator

Kent E. Vrana, PhD
Elliot S. Vesell Professor and Chair, Department of Pharmacology
Penn State College of Medicine
500 University Drive
Hershey, PA 17033-0850

Other Participating Researchers

Mechelle M. Lewis, PhD; Izel Tekin, BS; Xuemei Huang, MD, PhD - employed by Penn State College of Medicine

Expected Research Outcomes and Benefits

This project will help identify Parkinson's disease specific genetic alterations in a very well controlled subject group, and using this information we will be able to obtain novel targets for the cause of disease. These targets can be utilized in the future for the development of novel and efficient therapies. Since it has been described 200 years ago, there is not a single therapy that can cure or reverse the course of this disease. The identification of novel molecular causes will aid in the process of battling with this disorder. The data to be obtained from this translational project will provide an important step towards understanding the cause of Parkinson's disease.

Summary of Research Completed

We are currently completing sequencing data analysis and awaiting final next-generation sequencing data from an additional set of twins. Specific progress is detailed in the sections below.

<u>Sequencing of identical twins</u>
Three pairs of identical twins have been consented and their genomic DNA has been sequenced. These identical twins are noteworthy in that only one of the pair has Parkinson's disease (i.e., discordant disease manifestation on a nearly identical genetic background). The whole genome sequencing approach was employed to obtain genomic sequence information. We have utilized a non-amplification method to prepare subject-specific libraries. Whole genome sequences were obtained at 30x coverage for each base pair in 5 of the 6 subjects. The sixth sequencing failed to provide reliable data and will be repeated along with the final set of twins. At the present time, an experienced bioinformatician/statistician has been recruited to assist with the genome analysis (Dajiang J. Liu, PhD; Department of Public Health Sciences, Penn State College of Medicine). In the meantime, a library preparation has been performed on a fourth set of twins, and these samples are waiting sequencing time on the instrument.

Optimization of the bioinformatic analysis of the whole genome sequencing approach has proven to be more time-consuming than was previously anticipated. However, we believe these issues have been resolved by the recruitment of a new bioinformatics faculty member and we expect these studies to be complete within the next four months, including any follow-up experiments

and sequencing analysis.

In parallel to the progress described for the twin sequencing above, we have performed whole exome sequencing on two phenotypically distinct subject groups – early and late onset sporadic Parkinson's disease patients in order to investigate molecular contributors to PD. Comprehensive bioinformatic sequence analysis has just commenced, but we have already performed targeted gene/pathway specific analyses.

Recent literature has pointed to the noradrenergic system as a player in the development and progression of PD. One of the key components in PD pathology, alpha-synuclein protein, has been shown to form aggregates in the locus coeruleus of the midbrain before the substantia nigra (the area where the dopaminergic neurons reside). In addition to being an early target for neuronal pathology, these noradrenergic neurons provide input into the striatum and regulate the basal ganglia function. In PD patients, NE signaling and uptake at nerve terminals have been shown to be drastically decreased, indicating a clear lack of the release of NE to the synapse. The lack of NE at the early stages of PD is thought to underlie the pre-symptomatic depression that presents itself in the pre-motor stages of the disease. Lately, noradrenergic drugs have shown improvements of PD symptoms, however, these data warrant confirmation and future work. In animal models, where dopaminergic cells are targeted for destruction with MPTP or 6-OHDA, an alpha-2 adrenergic receptor agonist has improved locomotor function. These data, along with decreased locomotion in DBH double knock out animals, emphasize the importance of NE in basal ganglia function. The noradrenergic system is therefore a good target for the understanding of PD pathology, and variations within its components may underlie different PD characteristics. In view of these facts, we have analyzed the whole exome sequencing with a specific focus on biogenic amine-specific genes. This focus has permitted us to obtain SNP relationships without requiring sophisticated sequence analysis algorithms. These data are provided in the table below. We have confirmed the sequence genotypes for the tyrosine hydroxylase gene and had previously determined that the V81M polymorphism is associated with increased risk of "freezing of gait" – a severe symptom of the disease. Neither of these two SNPs is uniquely associated with early versus late disease onset in this patient population. Statistical analyses of these data are ongoing at this time.

Finally, two CURE grant supported papers have been published:

Tekin I, Roskoski R, Jr., Carkaci-Salli N, Vrana KE. Complex molecular regulation of tyrosine hydroxylase. J Neural Transm. In press, 2014. PMID: 24866693.

Tekin I, Vrana KE. Caveat emptor: SNP reporting in pharmacogenomics. Pharmacology. 2013;92(5-6):319-23. Epub 2013 Dec 12. PMID: 24356117

Table: List of Monoaminergic Gene Variants of the Early and Late Onset PD Subjects

Subject ID	Age of Onset	TH		DDC		COMT	DRD2		DRD3	DBH				
		V81M NP_000351	R202A NP_000351	M17V NP_001229817	R384Q NP_001229817	V158M NP_001128634	I183V NP_057658	H413Y NP_057658	G9S NP_000787	V26M NP_000778	N201S NP_000778	A211T NP_000778	A318S NP_000778	R549C NP_000778
007	47.77	HOM	WT	HOM	WT	HET	WT	WT	WT	WT	WT	WT	WT	WT
035	48.73	WT	WT	HOM	WT	HOM	WT	WT	WT	WT	HET	WT	WT	WT
55	39.00	WT	WT	HOM	HET	WT	WT	WT	WT	WT	HET	WT	WT	WT
067	52.84	WT	WT	HOM	WT	HOM	WT	WT	WT	WT	WT	WT	WT	WT
092	51.91	HET	WT	HOM	WT	HET	WT	WT	WT	WT	WT	WT	WT	WT
108	51.52	HET	WT	HOM	WT	HET	WT	WT	WT	WT	WT	HET	WT	WT
109	52.79	HET	WT	HOM	WT	HET	WT	WT	WT	WT	WT	WT	WT	WT
144	49.18	HET	WT	HOM	WT	HET	WT	WT	HOM	WT	WT	WT	WT	WT
151	47.70	HET	WT	HOM	WT	HET	WT	WT	WT	WT	WT	WT	WT	WT
153	49.43	HET	HET	HOM	WT	HOM	WT	WT	WT	WT	WT	WT	WT	HET
181	46.35	HET	WT	HOM	WT	HET	WT	WT	HOM	WT	WT	WT	HET	WT
014	65.32	HET	WT	HOM	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT
030	65.96	HOM	WT	HOM	WT	HET	WT	WT	WT	WT	WT	WT	WT	HET
062	67.79	WT	WT	HOM	WT	HET	WT	WT	WT	WT	WT	WT	WT	WT
070	74.21	HET	WT	HOM	WT	HET	WT	WT	WT	WT	WT	WT	WT	HET
130	66.88	WT	WT	HOM	WT	HOM	WT	WT	WT	WT	WT	WT	WT	WT
141	65.94	WT	WT	HOM	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT
148	67.35	WT	WT	HOM	WT	HOM	WT	WT	WT	WT	WT	WT	WT	HET
164	68.74	WT	HET	HOM	WT	HET	WT	WT	WT	HET	WT	WT	WT	WT
169	67.84	HET	WT	HOM	WT	HET	HET	HET	WT	WT	WT	WT	WT	HET
174	74.89	HET	WT	HOM	WT	HET	WT	WT	WT	WT	WT	WT	WT	WT

TH = tyrosine hydroxylase; DDC = dopa decarboxylase; COMT = catechol-O-methyltransferase; DRD = dopamine receptor; DBH dopamine-beta-hydroxylase

Table – cont.: List of Monoaminergic Gene Variants of the Early and Late Onset PD Subjects

Subject ID	Age of Onset	ADRA1A					ADRA2A	ADRB2		SERT (SLC6A	HTR1A	HTR1B	HTR2A		
		I200S - NP_150646	I322V - XP_005273472	C347R - NP_150646	R427L - NP_150646	E433Q - NP_150646	R240H - NP_000672	G16R - NP_000015	E27Q - NP_000015	G56A - NP_001036	I28V - NP_000515	F124C - NP_000854	T25N - NP_000612	D49N - NP_001159419	H452Y - NP_000612
007	47.77	WT	HET	HET	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT
035	48.73	WT	WT	HET	WT	WT	WT	HET	HET	WT	WT	WT	WT	HOM	WT
55	39.00	WT	WT	HET	WT	WT	WT	HET	HET	WT	WT	WT	WT	HOM	WT
067	52.84	WT	WT	WT	WT	WT	WT	HET	HET	WT	HET	WT	WT	WT	HET
092	51.91	WT	HET	HOM	WT	WT	WT	HET	HET	WT	WT	WT	WT	HOM	WT
108	51.52	WT	WT	HOM	WT	WT	WT	HET	HOM	WT	WT	WT	WT	HOM	WT
109	52.79	WT	WT	WT	WT	HET	HOM	HET	HET	WT	WT	WT	WT	HOM	WT
144	49.18	WT	WT	HOM	WT	WT	WT	HOM	HOM	WT	WT	WT	WT	HOM	WT
151	47.70	HET	WT	WT	WT	WT	WT	HET	HET	WT	WT	WT	WT	HOM	WT
153	49.43	WT	WT	WT	WT	WT	WT	HOM	HOM	WT	WT	WT	WT	WT	WT
181	46.35	WT	WT	HOM	HOM	WT	WT	HET	HOM	WT	WT	WT	WT	HOM	WT
014	65.32	WT	HET	HET	WT	WT	WT	HET	HET	WT	WT	WT	WT	HET	WT
030	65.96	WT	HET	WT	WT	WT	WT	HET	HOM	WT	HET	WT	HET	WT	WT
062	67.79	WT	HET	HET	WT	WT	WT	WT	WT	WT	WT	WT	WT	HOM	WT
070	74.21	WT	WT	HOM	WT	WT	WT	HET	HOM	WT	WT	WT	WT	HOM	WT
130	66.88	WT	WT	HET	WT	WT	WT	HET	HET	WT	WT	WT	WT	WT	WT
141	65.94	WT	WT	HET	WT	WT	WT	WT	WT	WT	WT	WT	HET	WT	WT
148	67.35	WT	WT	HET	WT	WT	WT	HET	HET	HET	WT	HET	WT	WT	WT
164	68.74	WT	WT	HOM	WT	WT	WT	HET	HET	WT	WT	WT	WT	HOM	WT
169	67.84	WT	WT	HET	HET	WT	WT	HOM	HOM	WT	WT	WT	WT	WT	WT
174	74.89	WT	WT	HET	WT	WT	WT	WT	HOM	WT	WT	WT	WT	WT	WT

Subject ID	Age of Onset	HTR3B	HTR5A	HTR6	HTR7
		Y129S - NP_006019	P15S - NP_076917	P375T - NP_000862	P279L - NP_062873
007	47.77	WT	WT	WT	WT
035	48.73	WT	WT	WT	WT
55	39.00	WT	WT	WT	WT
067	52.84	WT	WT	HET	WT
092	51.91	WT	WT	WT	WT
108	51.52	HET	WT	WT	WT
109	52.79	WT	WT	WT	WT
144	49.18	WT	WT	WT	WT
151	47.70	WT	HET	WT	WT
153	49.43	HET	WT	WT	WT
181	46.35	HET	WT	WT	WT
014	65.32	HET	WT	WT	WT
030	65.96	WT	WT	WT	WT
062	67.79	WT	WT	WT	WT
070	74.21	HOM	WT	WT	WT
130	66.88	WT	WT	HET	WT
141	65.94	WT	WT	WT	WT
148	67.35	HET	WT	WT	WT
164	68.74	WT	WT	WT	WT
169	67.84	HET	WT	WT	HET
174	74.89	HET	WT	WT	WT

ADRA =alpha adrenergic receptor; ADRB = beta adrenergic receptor; SERT = serotonin transporter; HTR = serotonin receptor