

Pennsylvania State University

Annual Progress Report: 2011 Formula Grant

Reporting Period

July 1, 2013 – June 30, 2014

Formula Grant Overview

The Pennsylvania State University received \$6,637,701 in formula funds for the grant award period January 1, 2012 through December 31, 2015. Accomplishments for the reporting period are described below.

Research Project 1: Project Title and Purpose

Development of Bioinformatics Methods for Medical Research – The purpose of this project is to develop and implement a wide array of bioinformatics tools for the analysis of large biological datasets, including those related to genomic, epigenetic, protein, and metabolomics experiments.

Duration of Project

1/1/2012 – 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 2: Project Title and Purpose

How Cells Respond to Stress – Recent studies have highlighted the fact that persistence of cancer cells depends on their successful adaptation to local stresses induced by local hypoxia and acidification. Moreover, our recent data indicating that eukaryotic cells exhibit stress induced mutations may suggest a mechanism for increased mutations rates in cancer cells. We plan to determine the mechanisms underlying stress resistance and stress induced mutation in a tractable model system. These components could provide novel targets for therapeutic intervention and our experimental system provides a test bed for evaluating the feasibility of such approaches.

Anticipated Duration of Project

5/1/2012 – 12/31/2015

Project Overview

We have recently shown that the stress-responsive transcription factor, Msn2, serves as an integrator of input from multiple stresses and does so in an unusual way. The activity of the transcription factor is modulated by the frequency with which it enters and exits the nucleus in a damped oscillatory mode, rather than by modulation of the amplitude of the factor in the nucleus in proportion to the intensity of the incident stress. We are pursuing these observations to sort out experimentally the calculus that the cell uses to integrate the inputs from the stress sensitive pathways, using microfluidic based single cell observations. Moreover, we are exploring the mechanistic basis for this oscillatory behavior and the extent to which other transcription factors engage in similar behavior. We also observed that genetically identical yeast cells exhibit a wide spectrum of responses to the same stress condition. We propose to test the hypothesis that this variation in response increases the fitness of the strain by allowing different cells to mount a different response to the same condition. Finally, we have shown that yeast cells exhibit a stress-induced increase in mutation frequency, a process that is dependent on Msn2/4. We propose to extend these studies by determining the means by which this increased mutation frequency occurs. Recent studies have highlighted the fact that persistence of cancer cells depends on their successful adaptation to local stresses induced by local hypoxia and acidification. Accordingly, we plan to extend these studies by introducing our mutation assay into mammalian cells to assess the effect of stress on mutation frequency in various cell lines.

Aim 1. Integration and coordinate regulation of the stress response. Msn2 and Msn4 integrate input regarding external stresses using a complex calculus. Cells regulate Msn2 activity through concerted cycles of nuclear-cytoplasmic relocalization. Different inputs are translated into outputs predominantly by modulating the frequency of those Msn2 oscillations. We plan to address the mechanistic basis of the integration of stress inputs as well as the means by which the oscillations are regulated by those inputs. Second, we plan to test the hypothesis that the oscillatory behavior of Msn2 in response to stress serves to provide a coherent output of a cohort of stress responsive genes. Third, we plan to correlate the stress-induced modification of Msn2 behavior with stress-induced posttranslational modification of the protein.

Aim 2. Stress induced mutation (SIM). We have recently obtained data suggesting that yeast cells exhibit an increased frequency of mutation when subjected to different stresses. In order to provide a better means of detecting stress induced mutations, we are developing a reporter assay that will provide a more immediate readout of mutational events than are currently employed. We propose to conduct a comprehensive genetic analysis of SIM, focusing initially on examining the components of stress response, DNA replication, and DSB repair pathways by analyzing the effects on SIM of genetically eliminating their functions.

Principal Investigator

James R. Broach, PhD
Professor and Chair
Department of Biochemistry and Molecular Biology
Pennsylvania State University College of Medicine
500 University Drive
Hershey, PA 17033

Other Participating Researchers

Jacob Hornick – employed by Pennsylvania State University College of Medicine
Erika Shor, PhD – employed by Princeton University, Princeton, NJ, Vasudha Bharatula –
employed by Pennsylvania State University College of Medicine
Mark Rutledge – employed by Princeton University

Expected Research Outcomes and Benefits

The experiments proposed in this application should answer several critical questions about how cells sense stress and what they do in response to stress. In particular, we should learn what are the pathways that transmit information on stress – many of which are conserved across all eukaryotes – and what is the role of nuclear oscillations in stress response and how extensive is this strategy. Moreover, we hope to explain why cells exhibit different behaviors in identical conditions. Does this enhance survival in an uncertain world? Finally, we hope to understand stress induced mutations and whether this process might explain the high rate of mutation accumulation in cancer cells. These observations should help us formulate novel approaches to addressing a number of diseases, including cancer.

Summary of Research Completed

Aim 1. Integration and coordinate regulation of the stress response.

In the previous reporting period, we determined the signaling network responsible for regulating the entry of Msn2 into the nucleus in response to stress and provided a model to account for the oscillatory behavior of Msn2. In this current reporting period, we have examined what happens once Msn2 enters the nucleus and how the outcome depends on the nature of the oscillations.

When Msn2 enters the nucleus, it can bind to stress response elements (STREs) within the genome to alter transcription of genes neighboring the sites. Approximately 3000 STREs reside upstream of yeast genes, but only a fraction of these serve as binding sites for Msn2. Many of these sites are likely occluded by positioned nucleosomes that prevent access to Msn2. Moreover, since each cell contains only 100-200 Msn2 molecules, formation of stable Msn2 complexes with a large number of STREs within a single cell's genome would not be possible. One explanation for the rapid dynamics of Msn2 localization may be to facilitate sampling of multiple sites by individual Msn2 molecules. Whether different stresses affect the selection of

different subsets of sites – either by modifying Msn2’s DNA binding recognition region or by altering the accessibility of different sites – had not been extensively explored.

We examined binding of Msn2 to genomic sites by chromatin immunoprecipitation followed by DNA sequencing in response to the nutritional stress of transferring cells from rich media to media lacking glucose. We also measured the global nucleosome architecture by sequencing micrococcal nuclease resistant fragments from chromatin isolated before and after application of the stress, both in the presence and absence of Msn2, in order to address the extent to which Msn2 binding influences and is influenced by nucleosomes. Finally, we assessed the sufficiency and necessity of Msn2 binding on changes in expression of each associated gene, to determine the effect on transcription elicited by Msn2 binding.

ChIP-Seq identified few Msn2 binding sites prior to the carbon source downshift and a large number after the downshift. We computationally identified sites of Msn2 binding and then hand annotated each of the peaks to identify the genomic features associated with each site. This process yielded 273 distinct and robust peaks of bound Msn2, distributed over 268 genes. Most of these sites corresponded to promoter regions and were highly enriched for STRE sequences, particularly clusters of STREs. While most of the Msn2-binding sites identified upon nutrient downshift overlap those identified following oxidative stress, we noted that a number of stress-induced Msn2 binding sites identified after oxidative stress were absent upon nutrient downshift, and vice versa. Those genes in whose promoters Msn2 bound following oxidative stress but not after glucose downshift were enriched for those involved specifically in response to treatment with oxidizing agents. Similarly, those genes in whose promoters Msn2 bound following glucose downshift but not in response to oxidative agents were highly enriched in those involved in carbohydrate metabolism. Thus, in addition to binding to promoters of genes involved in a common stress response, Msn2 binds specifically to a subset of genes associated with response to the particular initiating stress.

How might Msn2 binding be responsive to specific stimuli? One possibility is that posttranslational modification of Msn2 dictated by different stress signaling pathways modifies the binding specificity of the protein (Figure 1A). In this context, different stresses yield different patterns of Msn2 nuclear entry and exit, patterns that may well be perceived by different promoters in different ways. We found in our study that Msn2 responsive genes had quite divergent delay times in response to Msn2 activation when Msn2 was largely cytoplasmic and exhibited only random bursts of nuclear occupancy. This difference was essentially eliminated when Msn2 resided predominately in the nucleus (Figure 2). This divergence in response times may reflect the distinction recently described by Hansen and O’Shea, who noted that some genes respond to both sustained and pulsatile Msn2 nuclear localization (fast promoters) while others responded only to sustained Msn2 localization (slow promoters). Another possible scenario to explain stress specific Msn2 promoter binding is that transcription factors responsive to specific stress signals, e.g. Yap1 for oxidative stress or the Hap1-4 complex for glucose downshift, bind to promoters in a stress-specific manner and stimulate chromatin clearance at those promoters, allowing access of Msn2 to STRE sites that would otherwise be occluded (Figure 1B). This may represent an example of cooperative binding of transcription factors by sequential unwrapping of DNA from nucleosomes, such that binding of a transcription factor to its cognate site near the periphery of a positioned nucleosome provides access to a

binding site for a second transcription factor that would be otherwise buried under the interior of the nucleosome.

To explore these possibilities, we examined the nucleosome positions globally before and after glucose downshift in both an *MSN2 MSN4* and an *msn2 msn4* strain. In this way, we were able to identify promoters in which expression change was associated with nucleosome repositioning and determine whether that repositioning was dependent on Msn2 and/or Msn4. Two examples of genes in which transcriptional activation, Msn2 binding and nucleosome depletion from the promoter were coincident and in which nucleosome depletion was dependent on Msn2 and/or Msn4 are profiled in Figure 3A,C. Similarly, we identified several genes at which transcriptional repression, Msn2 binding and nucleosome acquisition are coincident and in which nucleosome acquisition is dependent on Msn2 (Figure 3B,D). From these studies, we conclude that a significant function of Msn2 is to expose promoter regions during gene activation and to occlude promoter regions during gene repression. Thus, Msn2 is not simply a passive respondent to remodeling promoted by other agents but plays an active role in restructuring the nucleosome depleted region during transcriptional reprogramming. Moreover, these results partially explain the differences in induction kinetics among different genes. Namely, most of the genes that exhibited rapid induction kinetics did not show extensive remodeling of the nucleosomes around the promoter. On the other hand, those genes that showed slow induction exhibited extensive remodeling. Accordingly, we conclude that the rate limiting step in induction involves conversion from a closed to an open promoter configuration and those genes already in an open configuration require less time for induction than those that require conversion.

Aim 2: Stress induced mutation (SIM).

As noted in last year's progress report, we were able to define genetically the interconnections of the stress response, the error-prone DNA repair pathway and the double strand break repair pathway. Specifically, we were able to show that the stress response shifted the relative contribution of translesion synthesis versus double strand break repair following induction of DNA damage and thus increased the mutation frequency following damage. Accordingly, these studies essentially completed the research under this aim in the initial proposal.

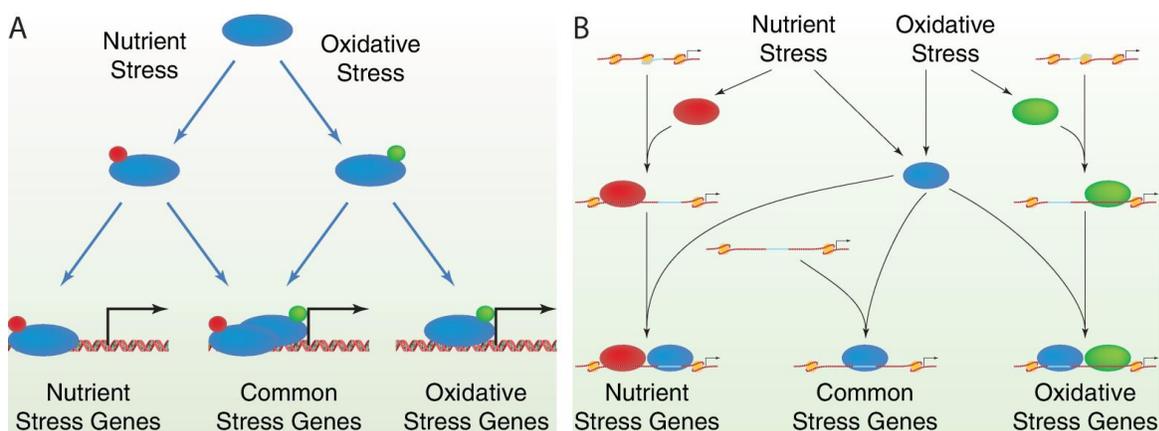


Figure 1. Possible mechanisms for stress specific binding of Msn2 to different sets of genes. *A:* Different stresses could result in distinct modifications (red versus green dots) of Msn2 (blue oval), which could alter the binding specificity or the nuclear occupancy dynamics and which would lead to interaction with different but overlapping sets of stress responsive genes. *B:* Different stresses could activate Msn2 (blue oval) as well as a stress-specific transcription factor (red oval for nutrient stress, green oval for oxidative stress). Those genes with STREs lying in nucleosome free domains would bind Msn2 under either condition. However, binding of the stress specific transcription factor could partially unwrap adjacent nucleosomes (beige ovals) to reveal additional, previously inaccessible STREs to which Msn2 could bind.

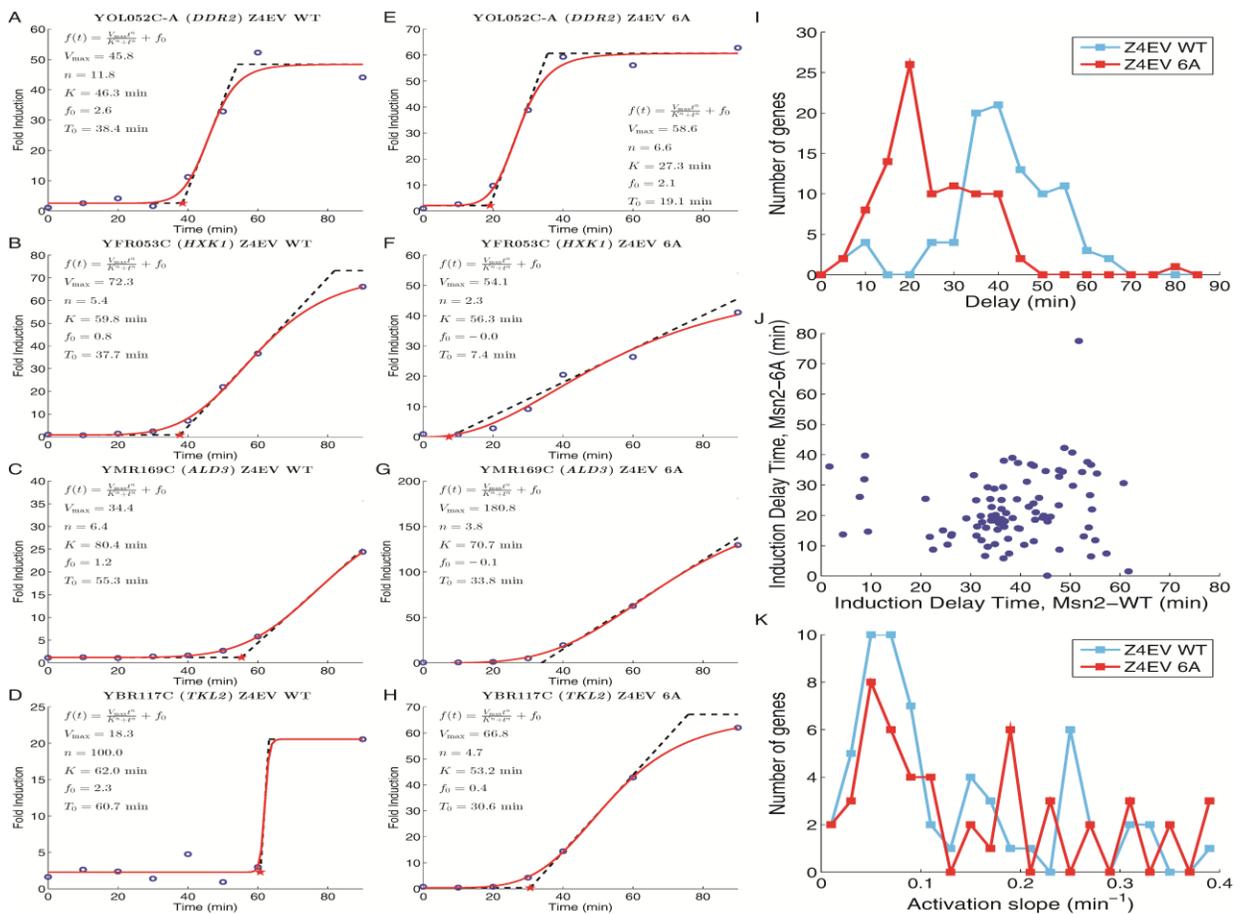


Figure 2. Different genes exhibit different induction kinetics in response to Msn2. *A-D*: Induction kinetics for four Msn2-inducible genes following estradiol addition to strains containing wild type *MSN2* under control of the hybrid Z₄EV transcription factor. Blue circles: fold increase in gene transcript levels relative to that immediately prior to estradiol addition; Solid red line: best fit of the data for each gene to the function $f(t) = f_0 + V_{max} \cdot t^n / (K^n + t^n)$; Dashed line: tangent line to the curve at $f(t) = V_{max}/2$, whose extrapolation to the x-axis provides the measure of the time delay in response to Msn2 induction. *E-H*: Induction kinetics for the genes shown in A-D following estradiol addition to strains containing *MSN2*^{6A} under control of the hybrid Z₄EV transcription factor. *I*: Histogram of time delay values for the 96 genes induced more than two-fold in both the *MSN2* wild type and *MSN2*^{6A} strains and whose induction values are reasonably fit by the Hill curve. Blue line: delay values in the *MSN2* wild type strain; Red line: delay values in the *MSN2*^{6A} mutant strain. *J*: Scatter plot of the delay time for each gene in I in the *MSN2*^{6A} strain relative to that in the *MSN2* wild type strain. *K*: Histogram of rates of induction, i.e., the slope of the tangent line to the fitted curve at $f(t) = V_{max}/2$, for the 96 genes in I. Blue line: delay values in the *MSN2* wild type strain; Red line: delay values in the *MSN2*^{6A} mutant strain.

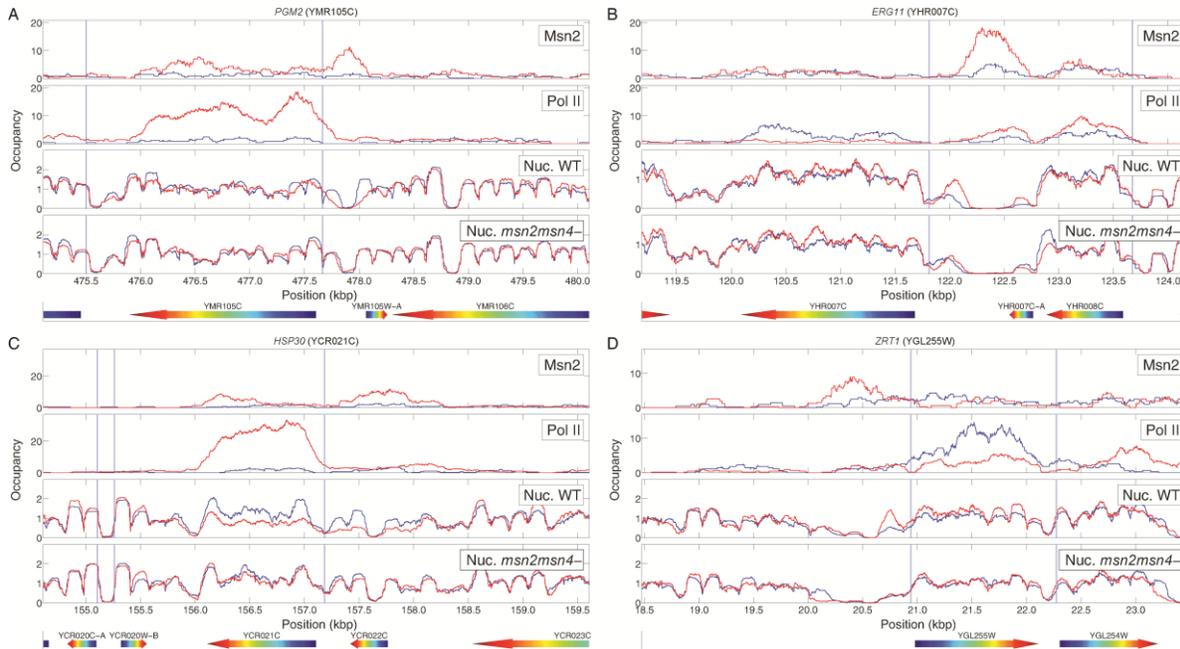


Figure 3. Msn2 promotes nucleosome repositioning over gene promoters. Shown are binding profiles of Msn2 (row 1) and PolII (row 2) as well as the nucleosome profiles (rows 3 and 4) over four different Msn2-regulated genes and the surrounding regions along the genome. Each panel shows a binding profile before the glucose-to-glycerol switch (blue lines) and 20 minutes after the transition (red lines). The nucleosome profiles were obtained for both an *MSN2 MSN4* (row 3) and an *msn2 msn4* strain (row 4). Two of the genes, *PGM2* and *HSP30*, are induced by Msn2 following the nutrient downshift, while the other two, *ERG11* and *ZRT1*, are repressed. In order to be able to compare the occupancy profiles obtained from different experiments with different sequencing coverage, we normalized all profiles such that the average occupancy across each chromosome in each experiment equals 1.

Research Project 3: Project Title and Purpose

Anal Human Papillomavirus Infection in HIV-infected Women – Human papillomavirus (HPV)-related anal cancer is an emerging health problem in people infected with human immunodeficiency virus (HIV). Few longitudinal studies have been conducted to understand the natural history of anal HPV infection in HIV-infected women. We propose this project with the goal to prospectively investigate risk factors associated with incident and persistent anal HPV infection in a cohort of rural HIV-infected women. This project will provide new evidence on the epidemiology and immunopathogenesis of anal HPV infection, and will help clinicians to plan appropriate strategies for anal cancer prevention in HIV-infected women.

Duration of Project

7/1/2012 – 6/30/2014

Project Overview

Human papillomavirus (HPV) is the most common sexually transmitted infection and is an established causal agent of anogenital cancers. HPV-related anal cancer is one of the top three emerging cancers among people infected with human immunodeficiency virus (HIV+) in the era of highly active antiretroviral therapy (HAART). Anal HPV infection is highly prevalent in HIV+ women, but few longitudinal studies have been conducted to understand the acquisition, persistence, and progression of anal HPV infection in HIV+ women. Lack of such knowledge is an important limitation to developing strategies to tackle the unmet medical needs for clinical management of anal HPV infection and anal cancer prevention. As HIV+ women live longer in the HAART era, the lifetime risk of anal cancer is increasing. Therefore, more research is needed to better understand the natural history of anal HPV infection in HIV+ women.

The goal of this project is to define the rates of and risk factors associated with incident and persistent anal HPV infection in a longitudinal cohort of HIV+ women living in a rural environment. The central hypothesis is that high risk behaviors and the lack of serum HPV antibodies are associated with the incident and persistent anal HPV infection in HIV+ women. The objectives of this project are to (1) assess the rate and determinants of newly acquired anal HPV infection in HIV+ women; and (2) examine the rate and risk factors associated with persistent anal HPV infection in HIV+ women. To achieve the aims, we are proposing a longitudinal study to follow 100 rural HIV+ women at 4-month intervals to assess anal HPV infections over time. Findings from this project will improve knowledge in the field of host and viral factors for anal HPV infection and will help researchers to develop guidelines for clinical management of anal infection in HIV+ women.

Principal Investigator

Ping Du, MD, PhD
Assistant Professor
Pennsylvania State University College of Medicine
500 University Drive
Hershey, PA 17033-0850

Other Participating Researchers

Lan Kong, PhD; Craig M. Meyers, PhD; John J. Zurlo, MD; Eugene J. Lengerich, VMD, MS; Neil D. Christensen, PhD – employed by Pennsylvania State University College of Medicine

Expected Research Outcomes and Benefits

Expected research outcomes: To successfully accomplish this project, we have put together an interdisciplinary team of investigators with expertise in epidemiology, HPV immunology and biology, HIV medical care, and biostatistics. At the completion of this project, we expect to have identified risk factors related to anal HPV infection and to provide new evidence on the epidemiology, immunity, and biology of anal HPV infection in HIV+ women. The findings will have great impacts on practice paradigms. The information can help clinicians to adapt practice

to seek either more or less aggressive strategies for management of anal HPV infection, to focus the prevention efforts on high-risk HIV+ women, and to improve physician-patient communication about anal HPV infection. The project will also provide valuable information on host immunity against HPV infection. The results can be used to better inform HPV vaccination to prevent anal HPV infection in populations of HIV+ women, other high-risk groups, or older women in general.

Benefits: HPV-associated cancers have become an emerging health issue for HIV+ people due to immunosuppression and co-infection with HPV. However, anal HPV infection is understudied in HIV+ women. It is also anticipated that rural HIV+ women have limited knowledge about HPV infection. Through this project, HIV+ women will gain more knowledge on HPV-associated cancers. The project will help medical science to further understand the etiology of HPV-associated diseases in order to reduce the health impacts of anal HPV infection in HIV+ women.

Summary of Research Completed

Overview: During this year we have recruited 2 additional HIV-infected women in this study bringing the total enrollment to 104 women. We collected longitudinal data on their sexual practices, general health behaviors, HIV/AIDS-related medical information, HPV infection status (at the oral cavity, the vagina, and the anus), and anti-HPV antibody status. We completed the data collection on June 30, 2014. We have also developed the statistical models to examine factors associated with incident (new) anal HPV infection. Given the complex nature of the longitudinal data with repeated measures, we are in the process of finalizing the statistical model to evaluate risk factors related to persistent anal HPV infection.

Aim 1: Assess the incidence and determinants of anal HPV infection in HIV+ women.

The characteristics of the study participants, the presence of anti-HPV antibodies at the baseline, and the incidence rate of anal HPV were described in the previous progress report. The updated results are presented in Table 1. A total of 78 HIV+ women returned for follow up visits. About 35% of women had at least one incident HPV infection (anal HPV infection: 28%). We detected a total of 63 incident anal HPV infections, 70 incident vaginal HPV infections, and 23 incident oral HPV infections among women with repeated visit(s) (one woman can be infected with multiple HPV types and with multiple times). Since the last progress report, we have conducted and finalized the following analyses:

1. We tested the longitudinal serum samples using cell-infection neutralization assays with native HPV16, 18, and 31. The results indicated that at the baseline, the seropositivity to HPV16, 18, and 31 was 54%, 42%, and 26%, respectively. However, during the follow up visits, nearly 50% of women who were seropositive at the baseline visit became seronegative, suggesting that HPV antibodies acquired from natural infection were waning over time. Interestingly, five women who were seronegative to HPV16, 18 or 31 at the baseline became seropositive, but the corresponding HPV infection was not detected. We are investigating if this result could reflect a possible reactivation of a latent HPV infection.

2. We developed the statistical models to examine risk factors associated with incident anal HPV

infection. Risk factors included in the models were separated into time-invariant variables (e.g., race-ethnicity, education, and medical history) and time-varying variables (e.g., sexual practices, HIV viral load, and CD4 counts). We used a Poisson or negative binomial regression (if overdispersion existed) model for the multivariate analyses. Poisson regression counts were adjusted for varying time period from the baseline to the last follow-up visit by using log interval time as an offset. Bivariate analysis was performed first and variables that were significant at p -value <0.2 were then entered into the multivariate model using a forward selection approach. HPV infection status at the baseline was included in the model regardless of significance. We also conducted an additional analysis focusing on the overall incident HPV infection with high-risk types.

The results showed that a history of anogenital warts- an indication of previous HPV infection- was significantly associated with incident anal HPV infection (Table 2). The baseline HPV infection status and current smoking behavior were significantly related to the overall incident HPV infection, incident HPV infection with high-risk types, and incident vaginal HPV infection. We did not identify significant factors for incident oral HPV infection due to the small sample size of oral HPV infection. These results suggest that HIV-infected women who had an HPV infection before are still at risk for acquiring new HPV infections, possibly because some women continue to engage in high-risk behaviors for HPV infection. HIV primary care physicians should provide routine health education about HPV infection and HPV-related cancers to HIV-infected women and focus on HIV-infected women who have a history of HPV infection for behavioral interventions. HPV vaccination may be beneficial for HIV-infected women as the natural HPV antibodies may not be protective against new HPV infections.

Aim 2: Define the rate of and risk factors associated with persistent anal HPV infection.

We defined a persistent anal HPV infection as continuous detection of the same type of HPV DNA from the anal swab for ≥ 2 visits. During the follow up visits, a persistent anal, vaginal, or oral HPV infection was detected among 20%, 27%, and 2% of HIV-infected women, respectively. Persistent anal HPV infection with high-risk HPV types was found in 11% of women. For the statistical model, we classified the study outcome as: persistent anal HPV infection, transient anal HPV infection, and no anal HPV infection. We are developing a multi-level logistic regression model with both fixed and random effects to evaluate factors associated with the persistent anal HPV infection.

Table 1. Characteristics of HIV+ women in the study (N=104)

Selected characteristics	% or Median (range)
Age (in years)	46 (20-66)
Non-Hispanic white	52%
Married/living with a partner	32%
Living in central PA >10 years	79%
In low socioeconomic status	60%
Self-reported HPV knowledge ("Nothing/only heard of")	61%
Length of HIV diagnosis (>10 years)	67%
Current CD4 count (cells/ml)	527 (26-1,542)
Current low HIV RNA viral load (<=200 copies/ml)	83%
Currently on cART	82%
History of AIDS complications	24%
History of abnormal cervical cytology	54%
History of anogenital warts	26%
>10 lifetime number of sexual partners	36%
Use of injection or non-injection drugs	21%
Current smoker	50%
Use of smokeless tobacco	2%
Regular alcohol consumption	27%
HPV infection at baseline (overall)	60%
Infection with high-risk (HR) types	30%
Anal HPV infection	51%
Vaginal HPV infection	43%
Oral HPV infection	10%
Overall incident (new) HPV infection	35%
Incident infection with HR types	20%
Incident anal HPV infection	28%
Incident vaginal HPV infection	23%
Incident oral HPV infection	10%
Overall persistent HPV infection	31%
Persistent infection with HR types	17%
Persistent anal HPV infection	20%
Persistent vaginal HPV infection	27%
Persistent oral HPV infection	2%

Table 2. Factors associated with incident HPV infections

Incident HPV infection	Significant predictors Incidence rate ratio (95% confidence interval) and p-value
Overall incident infection	Baseline HPV infection status (Yes vs. No): 1.18 (1.01, 1.38), p=0.02 Current smoking (Yes vs. No): 4.81 (2.05, 11.2), p=0.007 Travel history to meet sexual partners (Yes vs. No) 0.33 (0.13, 0.79), p=0.01
Incident infection with high risk HPV types	Baseline HPV infection status (Yes vs. No): 1.82 (1.18, 2.83), p=0.006 Current smoking (Yes vs. No): 14.0 (2.89, 61.6), p=0.0004 Travel history to meet sexual partners (Yes vs. No): 0.16 (0.06, 0.47), p=0.0009 Having >=2 sexual partners in the past 12 months (Yes vs. No): 0.41 (0.18, 0.92), p=0.03
Incident anal HPV infection	History anogenital warts (Yes vs. No): 4.75 (1.34, 16.9), p=0.02
Incident vaginal HPV infection	Baseline vaginal HPV infection status (Yes vs. No): 1.84 (1.15, 2.91), p=0.01 Current smoking (Yes vs. No): 16.4 (5.00, 83.1), p<0.0001 History of oral contraceptive use (Yes vs. No): 0.23 (0.07, 0.74), p=0.01 HIV viral load (Detectable vs. Undetectable): 0.12 (0.03, 0.61), p=0.01
Incident oral HPV infection	No significant factors identified

We are preparing grant proposals and multiple manuscripts using the results obtained from this study. Since the last progress report, we have submitted two NIH grant applications and one manuscript, and have presented research findings at three conferences, as follows:

Grant Applications:

NCI R03 (Du, P.I.)

12/01/2014-11/30/2016

Title: The roles of glutathione in the natural history of HPV infection in HIV+ women

The goal of this project is to examine the association between low glutathione level and HPV infection in HIV+ women.

Total amount requested: \$153,000

Status: pending

NIH/U01 (Du, Meyers, Wu, Multiple P.I.)

07/01/2015-06/30/2020

Title: Racial disparities in HPV-associated cervical cancer

The major goal of this project is to integrate the biological mechanism model into the population model to better address cervical cancer disparities between black and white women.

Total amount requested: \$3,414,228

Status: submitted in November 2013; scored, to be resubmitted in November 2014

Manuscript under review:

Du P, Brendle S, Camacho F, Zurlo J, Christensen N, Meyers C. Comparisons of VLP-based ELISA, neutralization assays with native HPV, and neutralization assays with PsV in detecting HPV antibody responses in HIV-infected women. (submitted to *PLOS ONE* on June 16, 2014)

Manuscripts in preparation:

Archibeque A, Zurlo J, Lengerich E, Meyers C, Christensen N, Du P. Factors associated with anal human papillomavirus infection in HIV-infected rural women.

Du P, Camacho F, Kong L, Zurlo J, Lengerich E, Meyers C, Christensen N. Longitudinal assessment of anal HPV infection in rural HIV+ women: Implications for anal cancer prevention.

Du P, Camacho F, Zurlo J, Lengerich E, Meyers C, Christensen N. Oral HPV infection in rural HIV-infected women: Unclear health impact?

Presentations:

Archibeque A, Du P. Factors associated with the prevalence of anal human papillomavirus infection in HIV-infected rural women. Public Health Day, Department of Public Health Sciences, Penn State University College of Medicine. Hershey, PA: May 9, 2014.

Du P, Camacho F, Zurlo J, Lengerich E, Meyers C, Read A, Christensen N. Longitudinal assessment of anal human papillomavirus (HPV) Infections in HIV+ women. Women's Health Research Day, Penn State Center for Women's Health Research. Hershey, PA: April 8, 2014.

Du P. Human papillomavirus (HPV) Infections in HIV-infected women: Implications for prevention of HPV-related cancers. Penn State Hershey Cancer Institute Interdisciplinary Research Conference. Hershey, PA: February 18, 2014.

Research Project 4: Project Title and Purpose

Structural Studies of Virus and Receptor Interactions – Cryo-electron microscopy and X-ray crystallography will be used to pursue a structural study of the neurotropic enterovirus 71. The virus will be reconstructed in 3-D alone and complexed to receptor to investigate receptor use and the pathogenic consequences. There are no known structures, little understanding of virulence determinants, and no vaccine for EV71. The aim of our study is to support other ongoing efforts by elucidating the capsid structure to understand viral function, especially the mechanism of receptor recognition.

Duration of Project

7/1/2012 – 6/30/2014

Project Overview

Enteroviruses are members of the picornavirus family, which cause a wide spectrum of disease. Several enteroviruses, including enterovirus 71 (EV71), are frequently associated with a mild exanthematous infection called hand, foot and mouth disease. However, only infections involving EV71 can progress to severe neurological disease, including aseptic meningitis, fatal encephalitis and acute flaccid paralysis. EV71 currently is a major public health issue across the Asia-Pacific region, but being an RNA virus, it lacks a proofreading mechanism and is evolving rapidly with new outbreaks occurring regularly. It is recognized as an emerging infectious disease with pandemic potential.

Little is known about the genetic determinants of pathogenicity of enteroviruses like EV71 and to what extent the structural genes contribute, either by mediating receptor interaction or by directing some aspect of viral uncoating or disassembly. There have been detailed mutational studies, genotypic comparisons of clinical isolates, comprehensive binding assays, and the recent discovery of specific receptors, but no structural studies. Structural studies will complement and supplement the ongoing investigation of these pathogenic viruses.

This study will directly examine the structure of EV71 virus capsids by cryo-electron microscopy (cryoEM) and X-ray crystallography. The contribution of receptor binding to pathogenicity will be investigated by cryoEM reconstructions of virus-receptor complexes. Elucidating attachment by examining the structure of the virus interacting with cellular components can lead to a more complete understanding of the viral life cycle. Such a study may also clarify aspects of viral evolution and aid in the advance of structurally based anti-viral therapies and vaccines. Specifically, the study will examine the molecular mechanisms of specific receptors recognized by an enterovirus with pandemic potential.

Aim 1: Identify the contribution of molecular components of receptors to tropism by investigating receptor usage of emerging pathogens with pandemic potential.

Aim 2: Solve the atomic structure of enterovirus 71 by X-ray crystallography.

Principal Investigator

Susan Hafenstein, PhD
Assistant Professor
Pennsylvania State University Hershey Medical Center
500 University Drive
Hershey, PA 17033

Other Participating Researchers

Hyunwook Lee, MS – employed by Pennsylvania State University Hershey Medical Center

Expected Research Outcomes and Benefits

We will solve the 3-D reconstruction of the pathogenic enterovirus 71 (EV71) by cryo-electron

microscopy and X-ray crystallography. EV71 is a member of the enterovirus A species, which contains several important pathogens. There are no known structures for any member of that species; therefore our structures will add important knowledge to the field of structural virology. We will also solve the structure of the virus interacting with a known receptor. Receptor recognition directly contributes to tropism and may affect virulence. All of the structural studies hope to aid and complement other ongoing research of EV71 to understand how the virus invades the central nervous system (CNS) and also to develop a vaccine.

Summary of Research Completed

Aim 1: Identify the contribution of molecular components of receptors to tropism by investigating receptor usage of emerging pathogens with pandemic potential.

Our work with the PSGL-1 receptor-virus interactions is continuing in collaboration with Yorihiro Nishimura and Jeff Bergelson. The advancement of new cryoEM Direct Electron Detector (DED) technology will likely lead to success for these studies. We have access to DED data collection through a productive collaboration with James Conway at the University of Pittsburgh. Application of the DED is enhancing cryoEM map resolution to nearly 3Å. With this approach, cryoEM will likely supplant crystallography in biological applications. The cryoEM map of the virus-receptor will allow us to make out details of peptide binding and identify the footprint of PSGL-1. Data collection time has been reserved for EV71 incubated with a PSGL-1 peptide.

During the reporting period using the structures we have mapped specific capsid residues that are essential for binding to PSGL-1, and have generated PB and non-PB viruses that are genetically identical except for one or two capsid residues involved in binding PSGL-1. We found that PB isolates (but not non-PB isolates) stimulate human peripheral blood mononuclear cells to release inflammatory cytokines that have been implicated in the pathogenesis of severe EV71 disease (Figure 1).

We demonstrated previously that a single capsid residue, VP1-145, regulates virus interaction with PSGL-1. Among approximately 1700 EV71 sequences in GenBank, 99% have either G, Q, or E at VP1-145. Viruses with VP1-145G or Q bind PSGL-1 (PB, 18% of sequenced viruses), whereas viruses with VP1-145E did not (non-PB, 82%). VP1-145 makes contact with an adjacent lysine residue, VP1-244K, which is highly conserved among EV71 isolates (present in 99.8%). Based on analysis of available crystal structures, we suspected that VP1-145 regulates PSGL-1 interaction by altering the orientation of VP1-244K, and the exposure on the virus surface of its positively charged patches. We found that mutation of VP1-244K to a non-charged residue abolished virus binding to PSGL-1, and mutation of an adjacent lysine, VP1-242K, reduced binding significantly (Figure 2).

Aim 2: Solve the atomic structure of enterovirus 71 by X-ray crystallography.

EV71-1095 strain has been successfully crystallized. The crystallization drops were prepared by mixing 2 µL EV71 at 5 mg/mL in PBS with 0.4 µL 0.2 M sodium citrate, 0.1 M Tris pH 8.5, and 30% (vol/vol) PEG 400. The well solution contained 1.8 M sodium acetate and 0.1 M Bis-Tris

propane, pH 7.0. Crystals formed within 1 wk.

Soaking of drug NF449:

Crystals were soaked by addition of 10 μ L of NF449 at a concentration of 20 mg/ml (in the 1.5 times concentrated crystallization solution) to the crystallization drop and incubation for at least 4 h prior to freezing in N₂(l).

Soaking of drug NF110:

Crystals were soaked by addition of 10 μ L of NF110 at a concentration of 20 mg/ml (in 1/2 diluted PBS, 50% glycerol) to the crystallization drop and incubation for 20 hours prior to freezing in N₂(l).

Crystals were transported to Penn State University Park where they were tested at the X-ray crystallography core facility. Diffraction was successful and data collection has been reserved at the Cornell High Energy Synchrotron Source (CHESS).

Manuscripts: The studies have led to two manuscripts:

Shingler KL, Cifuentes 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 Journal of Virology October 2014 (Corresponding Author)

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

Grant applications: The successful structure function study will allow us to pursue additional research and resulted in this grant submission:

NIH 1 R01—Bergelson, Hafenstein, Winkler (Multiple PI)—12/01/15-11/30/20

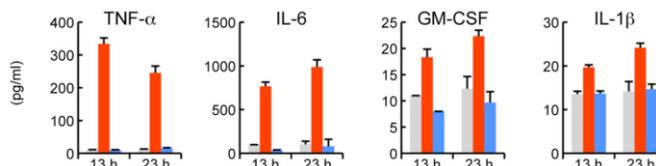
“EV71 receptors at the 5-fold vertex: functions, structures, antiviral targets”—\$534,709

The goal is to engage a highly interdisciplinary group of investigators with a distinctive range of scientific expertise— in virology, structural biology, and organic chemistry—to define how EV71 interacts with receptors and enters cells, and to apply this information to the generation of new small molecules with anti-viral activity.

Role: co-PI

Figure 1. PB EV71 stimulates cytokine production.

Human peripheral blood mononuclear cells were exposed to medium alone (Mock), or to isogenic PB and non-PB isolates differing at a single capsid residue (VP1-145) critical for PSGL-1 interaction (10 TCID₅₀ per cell). Cells were incubated at 37°C, supernatants were collected at 13h and 23h, and cytokine concentrations were determined by Luminex assay.



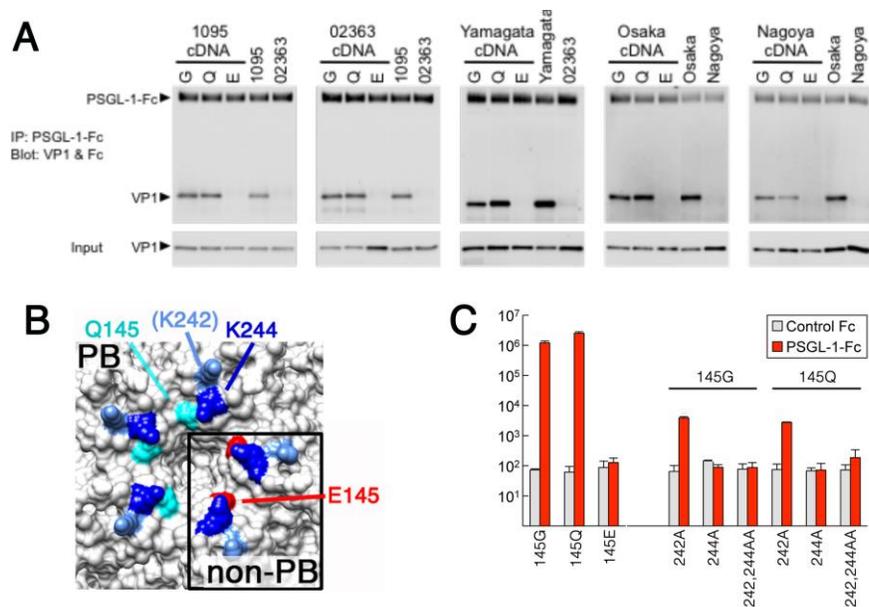


Figure 2. Residues at the 5-fold axis control virus interaction with PSGL-1. (A) Viruses with VP1-145 G or Q (but not E) bind PSGL-1. Mutations at VP1-145 were introduced into cDNA clones encoding PB isolates (1095, Osaka, and Yamagata) and non-PB isolates (0263 and Nagoya). Viruses were incubated with PSGL-1-Fc fusion protein, then precipitated with protein G beads. PSGL-1 fusion protein was detected by immunoblotting with anti-Fc antibody, and virus was detected with antibody to viral VP1. Wild-type isolates are included at the right of each panel as PB and non-PB controls. (B) Residues important for PSGL-1 binding surround the 5-fold symmetry axis: in the non-PB isolate (inset), VP1-244K (dark blue) VP1-244K is tipped toward VP1-145E (red); in the PB isolate (VP1-145Q, cyan), VP1-244K faces outward. VP1-242K is in light blue. (C) VP1-244K and VP1-242K are important for PSGL-1 binding. Binding to PSGL-1-Fc, or to a control Fc protein was detected by RT-PCR for viral genomes. Viruses with VP1-145 G or Q are PB, and bind specifically to PSGL-1-Fc; virus with VP1-145E does not bind. Mutation of VP1-244K to A reduces binding 10,000-fold, and mutation of 242K reduces binding 100-fold.

Research Project 5: Project Title and Purpose

T-Cell Immunity to Polyomavirus Infection – Polyomaviruses silently infect most humans, but can cause life-threatening disease in the setting of depressed immunity. No effective antiviral therapies are available. Because these viruses only infect their natural host reservoirs, we have limited understanding of the immunological mechanisms required to contain them. Studies proposed in this project use mouse polyomavirus to define determinants guiding differentiation of effective antiviral CD8 T cell responses that are needed to keep these smoldering infections in check.

Anticipated Duration of Project

5/1/2012 – 12/31/2015

Project Overview

The long-term goal of our research is to understand the mechanisms responsible for generating and maintaining effective T cell responses to persistent “smoldering” viral infections, with a focus on those caused by polyomaviruses (PyV). We will focus our efforts toward studying the

evolution of virus-specific T cell responses from acute to persistent stages of MPyV infection. We have recently demonstrated that naïve Mouse PyV (MPyV)-specific CD8 T cells are recruited during persistent infection, yielding memory cells that differ phenotypically and functionally (e.g., costimulation and CD4 T cell requirements) from those primed earlier in infection. Our new data show that temporal differences in recruitment qualitatively modulate memory anti-MPyV CD8 T cells. A central unresolved question is whether antiviral T cell memory in persistently infected hosts is predominantly maintained by antigen (Ag)-driven proliferation of memory T cells that differentiate from cells primed early in infection or by ongoing recruitment of naïve virus-specific T cells. This issue has implications for developing strategies to preserve/promote T cell-mediated control of persistent infections in humans as thymic output diminishes, as with aging. Our *central hypothesis* is that the functional integrity of the memory virus-specific CD8 T cell compartment is regulated by changes in viral Ag, inflammation, and CD4 T cell help over the course of MPyV infection. To test this hypothesis, we propose the following Specific Aims:

Aim 1: To define determinants regulating the contribution of acute vs. persistent infection-recruited antiviral CD8 T cells to the memory population.

Aim 2: To determine molecular signaling mechanisms underlying the functional differences between acute and persistent infection-recruited memory MPyV-specific CD8 T cells.

Principal Investigator

Aron E. Lukacher, MD, PhD
Chair of Microbiology and Immunology
The Pennsylvania State University
500 University Drive
Hershey, PA 17033-0850

Other Participating Researchers

Qingsong Qin, PhD; Shwetank Verma, PhD; Ge Jin, BS— employed by Pennsylvania State University College of Medicine

Expected Research Outcomes and Benefits

Studies proposed in this project will provide a new conceptual framework for understanding virus-specific T cell memory in persistent infections that will help resolve seemingly contradictory observations in the field – that (a) continuous Ag is required to maintain memory antiviral T cells [aka, “antigen addiction”], that (b) maintenance of antiviral T cell memory relies on naïve Ag-specific T cell recruitment during persistent infection [our work], and that (c) long-lived memory virus-specific T cells are also recruited during acute infection [our new data and that of Ann Hill’s and Douglas Fearon’s groups]. *We currently envision a “conveyor belt” scenario where ongoing priming of naïve virus-specific T cells is required to resupply the pool of deteriorating antiviral CD8 T cells that were generated at early stages of infection, but, as viral load/inflammation diminishes over time, bona fide memory T cells progressively emerge and assume a larger role in maintaining T cell memory.* The studies proposed here to define the

determinants and mechanisms that operate to shape the contributions of acute and persistent infection-recruited CD8 T cells to chronic memory will directly test this working hypothesis. Importantly, these studies will lay the groundwork for exploring innovative interventions to accelerate viral clearance and foster differentiation of effective antiviral memory T cells in the setting of persistent viral infections.

Summary of Research Completed

Specific Aim 1 seeks to define the roles of viral antigen and virus-associated inflammation in controlling the generation of memory CD8 T cell responses to MPyV infection. Given early unpublished data from our lab that the natural killer (NK) cell cytotoxic response peaked at day 4 after infection, and published literature implicating NK cells as an early source of T cell help, we asked if NK cells were responsible for augmenting progression of naïve anti-MPyV CD8 T cells to memory cells. To this end, we first confirmed by flow cytometric analysis that NK cell activity peaked in magnitude, expression of activation markers, and elaboration of interferon (IFN)- γ at day 3-4 after MPyV infection. We then transiently depleted mice of NK cells using an NK1.1 monoclonal antibody, but observed no effect on recruitment of naïve MPyV-specific CD8 T cells. However, compared to control mice given rat IgG, NK cell-depleted mice had an MPyV-specific CD8 T cells with elevated expression of CD127 [(interleukin) IL-7 receptor- α chain] and lower expression of KLRG1, a phenotype indicative of differentiation biased toward memory precursor T cells. We also began studies to analyze the cytokine transcriptome in the spleen during MPyV infection, using SYBR green quantitative PCR, and found a 4-fold increase in type I IFN mRNA by day 4 after infection that was (unexpectedly) sustained for at least one month. Intrinsic type I IFN signaling has been shown by other groups to promote virus-specific CD8 T cell responses, making IFN-I a prime candidate for modulating recruitment of MPyV-specific CD8 T cells during acute infection and preventing dysfunction and loss during persistent infection.

We previously reported that a subdominant CD8 T cell response directed to the D^b-restricted Large T Antigen epitope LT638 exhibited the following unusual characteristics: (1) peak expansion 2 days earlier than the dominant CD8 T cell response to the D^b-restricted LT359 epitope; (2) lack of CD127 surface expression; (3) no detectable cells by tetramer staining and intracellular cytokine staining by 12 days after infection; (4) reappearance of LT638-specific CD8 T cells at >50 days postinfection in ~1/3rd of B6 mice, often reaching levels equivalent to those of memory LT359 CD8 T cells. Using the RMA/S cell peptide stabilization assay, we determined that the LT638 peptide exceeded the binding affinity of the LT359 peptide by 10-fold. We hypothesized that high/sustained TCR stimulation of LT638-specific CD8 T cells biased their differentiation toward terminal effectors. My departmental faculty colleague and collaborator, Todd Schell, developed and maintains a transgenic mouse line expressing a TCR directed to an epitope from Large T Antigen of SV40, termed "Site V", which differs from LT638 by only 3 of 9 amino acids. Interestingly, the Site V CD8 T cells are only detectable in mice immunized by tumor cells expressing a mutant SV40 LT lacking the three dominant CD8 T cell epitopes; for this reason, the Site V response has been given the moniker "immunorecessive". Todd Schell's group previously showed that analog peptides with specific amino acid substitutions at the amino-terminal (P1) position bound better to D^b and had

increased stimulatory capacity for Site V CD8 T cells. This system provided us with the opportunity to directly test the hypothesis that TCR signal strength per se regulates CD8 T cell differentiation to persistent MPyV infection. We obtained a panel of synthetic oligopeptides such that each amino acid in the LT638 peptide was replaced by the corresponding residue in Site V, as well as getting peptides with the P1 substitutions having accentuated Site V T cell stimulatory ability. RMA/S peptide stabilization assays revealed that this peptide panel covered a wide range of peptide binding affinities for D^b. Using intracellular IFN- γ production as a readout for T cell stimulation, we found that these analog peptides fell into a profile of stimulatory capacity for Site V CD8 T effector cells that paralleled their binding affinity to D^b. By site-directed mutagenesis, we altered the LT638 epitope sequence in polyomavirus genomic DNA to express the Site V epitope, and did the same for each of these analog epitopes; we successfully isolated each of these mutant viruses. To date, we have determined that infection by each of these viruses did not change the dominant MPyV LT359-specific CD8 T cell response. Thus, we now have the reagents to use B6 mice made chimeric with adoptively transferred naive Site V CD8 T cells to directly test (as virus-associated inflammation and innate/adaptive antiviral immunity for other MPyV T cell epitopes is held constant for each mutant virus) the hypothesis that strong TCR signal strength drives terminal differentiation of MPyV-specific CD8 T cells.

IFN- γ has been shown to operate as an antiviral effector cytokine for primary human kidney epithelial cells infected in vitro by the human BK polyomavirus. We reported that IFN- γ was a dominant antiviral effector mechanism employed by infected mice to contain MPyV. Our in vivo studies using IFN- γ receptor (R) KO mice, however, could not exclude potential global effects of IFN on host immune responses (e.g., IFN-mediated upregulation of MHC molecules and direct effects on T cell expansion and differentiation). We sought an approach to selectively render infected host cells refractory to IFN receptor signaling. To this end, we initiated a collaboration with Benjamin tenOever's laboratory (Mt. Sinai School of Medicine) to insert into the MPyV genome a micro(mi)RNA coding sequence to target host cell STAT1, an essential transduction molecule for type I and type II IFN receptor signaling. Because this "artificial" miRNA is perfectly complementary to the STAT1 mRNA, the STAT1 transcript will be targeted for degradation and thereby silence production of new STAT1 protein. Fortuitously, a 2013 publication demonstrated that a small intron in the MPyV early region tolerated insertion of a 34-bp loxP sequence without affecting viral viability. Using the InFusion cloning strategy, we inserted an 80-bp miRNA for STAT1 or a control scrambled (SCR) sequence (plasmids for each were provided by the tenOever lab) into this intron, and have isolated mutant MPyVs. At this time, we have confirmed by sequencing that these viruses retain these intron insertions, and have determined that the mRNA primary transcript is processed to produce the ectopic miRNAs.

A goal of Specific Aim 1 is to explore the impact of MPyV infection on the quantity and quality of the memory anti-MPyV CD8 T cell response. To achieve this, we had constructed a recombinant MPyV carrying two loxP sites in cis, with one loxP inserted into the noncoding control region immediately upstream of the early region transcriptional start site and the other loxP inserted into a unique BlnI site in the overlapping middle T-large T antigen coding sequence. An early report showed that a mutant MPyV with a 66-bp deletion at this BlnI site did not suffer replication defects; however, we found that mutant viruses constructed with a loxP site at this position were attenuated and that we could not detect MPyV DNA by Taqman

quantitative PCR in mice one month after infection. Based on the recent report noted above describing the ability of an early region intron to tolerate insertion of a loxP site, we inserted the second loxP site at this position rather than at the BlnI site. At this position, we also inserted a loxP site having a single nucleotide change that ablated its ability to be recognized by cre. We designated the dual loxP virus “A2.dflx”, and the control loxP virus “A2.cflx”. Using a replication-defective adenovirus vector carrying cre to transduce cre into polyomavirus permissive cells, we confirmed that cre ablated the genomes of A2.dflx, but had no effect on replication by parental A2 strain virus or the control A2.cflx virus. We further created a dflox MPyV where the dominant LT359 epitope sequence was replaced by site-directed mutagenesis into the corresponding D^b-restricted "Site I" SV40 large T antigen epitope. We previously showed that MPyV mutants expressing this Site I epitope are recognized in vivo by CD8 T cells adoptively transferred from the TCR-I transgenic mouse. We currently maintain a colony of B6 background mice transgenic for a taxomifen-inducible cre recombinase. These unique reagents will allow us to test the hypothesis that de novo recruitment of naive MPyV-specific CD8 T cells during persistent infection maintains the pool of memory CD8 T cells.

Finally, to accomplish a major goal of *Specific Aim 2*, we have successfully applied multiparametric flow cytometry with phospho-site specific antibodies (Phospho-Flow) to study activation of TCR signaling molecules at the single-cell level. Using purified memory TCR-I CD8 T cells at day 30 post-transfer to acutely or persistently infected mice, we found that a markedly higher frequency of persistent infection-recruited cells phosphorylated the ERK MAP kinase than did those TCR-I memory cells recruited during acute infection.

Research Project 6: Project Title and Purpose

Computational Tool Development for Supporting Biomedical Research – We will develop an integrated software system for biomedical applications of human genetic-variation data, ranging from single-base differences in DNA sequence, to variation in the number of copies of a gene, to differences in abundance of certain gene transcripts, to perturbations in a network of protein interactions. The goal is to provide, in an easily accessible form, the tools needed to reliably infer biological function and consequences of variation from human genome data. The choice of tools to implement will be driven by the needs of collaborating physicians at Penn State’s College of Medicine.

Anticipated Duration of Project

1/1/2012 – 12/31/2015

Project Overview

The project is organized into the following six Specific Aims.

1. Analyze single-nucleotide polymorphisms
 - (1.1) *Uniformly determine SNPs from sequence data*
 - (1.2) *Identify statistically significant SNPs in case-control data*
 - (1.3) *Identify tag SNPs*

2. Predict phenotypes of protein variants
 - (2.1) *Link variants to recorded phenotypes*
 - (2.2) *Predict the likelihood that an amino-acid variant has a phenotype*
 - (2.3) *Identify systemic properties of a set of protein variants, such as effects on interacting networks of proteins*
3. Analyze gene-centric quantitative features
 - (3.1) *Reduce primary experimental data to gene-related quantities, such as copy number, expression level, or methylation status*
 - (3.2) *Compare gene-related quantities to find significant similarities or differences*
4. Predict phenotypes of non-coding variants
 - (4.1) *Analyze at high resolution sequence-census data for epigenetic features*
 - (4.2) *Integrate epigenetic data with other functional genomic data*
 - (4.3) *Integrate variation and epigenetic data*
 - (4.4) *Analyze allele-specific binding of transcription factors*
5. Classify disease status or risk
 - (5.1) *Classify disease status and/or predict risk from multiple factors (environmental and genetic)*
 - (5.2) *Develop machine-learning tools for more complicated phenotypic patterns*
6. Develop methods for GWAS of understudied genomic regions
 - (6.1) *Facilitate analysis of existing GWAS for chromosome X, Y, and mitochondrial variants*
 - (6.2) *Develop and apply bioinformatics and statistical analysis tools for these regions*

Principal Investigator

Kateryna Makova, PhD
Professor
Penn State University
305 Wartik Lab
University Park, PA 16802

Other Participating Researchers

Qunhua Li, PhD; Ross Hardison, PhD; Debashis Ghosh, PhD; Marylyn Ritchie, PhD; Anton Nekrutenko, PhD; Webb Miller, PhD – employed by Penn State University

Expected Research Outcomes and Benefits

A huge amount of spending by NIH is predicated on the belief that genetic differences among humans play a large role in the differences in disease susceptibility, and that genetic differences among tumors are critical for prognosis of cancer survivability. These investments now make it possible to determine the genetic make-up of each individual and individual cancers. However, research is needed to learn how to translate that information into diagnoses customized to the individual patient. Since the amount of data gathered for a patient is potentially huge, it is absolutely necessary to employ computers for its analysis. Ultimately, doctors will have this information and these methods at their fingertips, but only a dedicated research effort by the biomedical community will bring this to fruition.

This research project will make a substantial contribution to that goal. In particular, we will work with clinicians to identify their computational needs for utilizing these large and rich data sets, locate or produce the best computational tools, provide those tools in an easy-to-use computer environment, and work with the clinicians to ensure these resources are optimally used to improve human health.

Summary of Research Completed

Aims 1, 3, 4: The Ritchie lab has continued to develop the knowledge source and analysis strategy for low frequency DNA sequence variants, along with common SNPs, to look for association of genetic variation and disease phenotypes. In our work within the Library of Knowledge Integration (LOKI) and BioBin, we have done three primary projects. First, we have spent a good amount of time evaluating different knowledge sources to expand LOKI. As we have learned through the ENCODE project, the majority of the human genome is composed of regulatory, functional elements. The current implementation of LOKI is biased toward biological knowledge related to coding regions of the genome. Thus, it is an important goal to expand LOKI to include regulatory information. The data generated by ENCODE has been processed and analyzed by many different research groups and the important functional annotations are being cataloged in several databases such as RegulomeDB and HaploReg. We have been extracting the functional elements from these two sources, along with the UCSC genome browser catalog of ENCODE elements, to pull the non-redundant information into LOKI. This will make the next generation of LOKI include coding and non-coding regions of the genome for annotation and binning by BioBin.

Second, we are trying to determine how to most effectively combine the coding and non-coding regions into 'omic-units or 'omic-modules. The basic premise is that we expect that all of the non-coding, regulatory regions will be linked to one or more coding regions (those coding regions that they regulate). Therefore, we would like to link together the components of gene regulation with the genes they regulate to build larger functional elements – which we are calling 'omic-units or 'omic-modules. To achieve this, we are also expanding the gene region knowledge in LOKI using GENCODE, which is the reference human genome annotation. Using GENCODE, along with the other described ENCODE functional element databases, we will be able to construct the relationships between coding and non-coding regions to use for downstream analysis of genotype-phenotype relationships.

Finally, we have continued to evaluate BioBin with the earlier version of LOKI on different datasets. We have started testing BioBin using targeted gene sequence data, such as the PGRNSeq platform which we are using in the eMERGE network (electronic Medical Records and Genomics). We have started testing BioBin to look for somatic mutations in The Cancer Genome Atlas (TCGA) data. The TCGA is a tremendous resource for methods development and discovery of important genetic and genomic factors associated with one or more cancers. We have access to the complete TCGA through dbGaP and will be using this for the next year of the project.

Aims 2, 5: No work on these aims was done during this funding period.

Aim 6: The Makova laboratory has been continuing developing tools for accurate identification of genetic polymorphisms in mitochondrial DNA. Mitochondrial DNA (mtDNA) variation is characterized by heteroplasmy, the presence of several alleles in an individual. The manifestation of mtDNA diseases depends on heteroplasmy frequency, yet we lack the information about the prevalence of heteroplasmy in human populations. Furthermore, the heteroplasmy transmission between generations cannot be readily predicted due to the poorly understood mitochondrial bottleneck during oogenesis. In this most comprehensive study of heteroplasmy transmission to date, we analyzed previously generated sequencing data on blood and buccal mtDNA for 39 healthy mother-child pairs (156 samples) of European ancestry identifying 98 point heteroplasmies with minor allele frequency (MAF) $\geq 1\%$. A strong correlation in heteroplasmy MAF was observed between two tissues of the same individual, but a weaker correlation between mothers and children, supporting a mtDNA germline bottleneck with a size of only ~29-35 mtDNA molecules. We also estimated the germline mtDNA mutation rate at 1.3×10^{-8} mutations/site/year – lower than in previous pedigree studies but, remarkably, in agreement with phylogenetic studies, thus solving a long-standing controversy. Eight pathogenic heteroplasmic variants were found, leading to a frequency of carriers of one in eight unrelated individuals, the highest reported for mtDNA. Our results, obtained from a set of healthy individuals, provide a robust baseline for studying mitochondrial genome dynamics and have clear implications for predicting transmission of disease-causing variants heteroplasmies. All computational tools developed for this project are part of the Galaxy toolshed.

Publications:

1. (in Press) Rebeca Campos-Sanchez; Aurelie Kapusta; Cedric Feschotte; Francesca Chiaromonte; Kateryna D. Makova. Genomic Landscape of Human, Bat and Ex Vivo DNA Transposon Integrations Molecular Biology and Evolution 2014; doi: 10.1093/molbev/msu138
2. Benjamin Dickins, Boris Rebolledo-Jaramillo, Marcia Shu-Wei Su, Ian M. Paul, Daniel Blankenberg, Nicholas Stoler, Kateryna Makova, and Anton Nekrutenko. 2014 Controlling for contamination in re-sequencing studies with a reproducible web-based phylogenetic approach BioTechniques BioTechniques, Vol. 56, No. 3, March 2014, pp. 134–141
3. Kuruppumullage Don P, Ananda G, Chiaromonte F, Makova KD. 2013. Segmenting the human genome based on states of neutral genetic divergence. Proceedings of the National Academy of Sciences USA 110 (36): 14699-14704; doi:10.1073/pnas.1221792110
4. Ananda, G., E. Walsh, K. D. Jacob, M. Krasilnikova, K. A. Eckert, F. Chiaromonte, K. D. Makova. 2013. Distinct mutational behaviors distinguish short tandem repeats from microsatellites in the human genome. Genome Biology and Evolution 5: 606-620.
5. Moore CB, Wallace JR, Wolfe DJ, Frase AT, Pendergrass SA, Weiss KM, Ritchie MD. Low frequency variants, collapsed based on biological knowledge, uncover complexity of population stratification in 1000 genomes project data. PLoS Genet. 2013 Dec;9(12):e1003959. doi: 10.1371/journal.pgen.1003959. Epub 2013 Dec 26. [PM 24385916](#), [PMC 3873241](#)

Presentations made during this funding period: 9 oral and 19 poster presentations

Research Project 7: Project Title and Purpose

A Pilot Study of Personalized Medicine for Pediatric Asthma – We propose to conduct a pilot study to evaluate the effectiveness of personalized asthma care for children, a new treatment approach for this complex disease. The control arm of the trial will involve the current standard primary care asthma management guided by the National Institutes of Health’s National Asthma Education and Prevention Program (NAEPP), and the intervention arm will involve personalized asthma management (NAEPP management enhanced by incorporating genetic and environmental information).

Duration of Project

1/1/2012 – 6/30/2014

Project Overview

Asthma is one of the most common, chronic diseases of childhood affecting approximately ten percent of US children and adolescents. Asthma is a complex disease without a uniform presentation, and asthma triggers, treatment responsiveness, and severity vary by individual. This varied disease manifestation can be a challenge for clinicians managing asthma. The National Institutes of Health’s National Asthma Education and Prevention Program (NAEPP) issued the first guidelines for the management of asthma to address this challenge two decades ago, and the most recent revision was released in August 2007. The guidelines standardize clinical management of asthma with minor modifications for differences in disease expression. This standardized approach has helped improve asthma management, but standardization does not address a key issue – the differential response to treatment.

Investigators have identified genetic polymorphisms associated with reduced medication response. The presence of genetic differences which can affect response to treatment suggests that an individualized therapeutic plan may be better than the relatively uniform, standardized approach found in the current NAEPP guidelines. Social and environmental factors have also been identified which play a role in the exacerbation of asthma. The difference in triggers for exacerbation manifests in differences in disease expression. While the NAEPP guidelines note some of these factors as issues to address in management, the main focus of the NAEPP guidelines is on medical management.

The objectives of this project are: 1) to generate data needed for a formal power calculation, 2) assess the feasibility of the study design, and 3) to provide pilot data for a grant application for a randomized controlled trial (RCT) of personalized asthma care.

Principal Investigator

Adam J. Spanier, MD, PhD, MPH
Assistant Professor of Pediatrics and Public Health Sciences
Penn State University College of Medicine
500 University Drive

Hershey, PA 17033-0850

Other Participating Researchers

Julie Vallati, LPN; Jennifer Stokes, RN; Jessica Beiler, MPH; Heidi Watts, RN; Amyee McMonagle, RN; Patricia Carper, RN; Amy Shelly, LPN; Allen Kunselman, MA – employed by Penn State University College of Medicine

Expected Research Outcomes and Benefits

We expect that this study will provide us with the data needed for a formal power calculation, will permit us to assess the feasibility of the intervention and study design, and will provide pilot data for a grant application for a randomized controlled trial of personalized asthma care. The participants in both arms could experience improved asthma care through the guidance provided to their Primary Care Provider. In addition participants will learn of their allergen sensitizations if they have not received skin testing before. It is also possible that participants may experience no direct benefit.

Summary of Research Completed

Recruitment: Our initial enrollment did not occur as quickly as we anticipated. Prior to June 30, 2013 and the approval of a one-year no cost extension for this project, we had recruited 11 of 40 subjects. During the current reporting period (July 1, 2013-June 30, 2014), we conducted another direct mailing to clinic patients and put flyers in some of the subspecialty clinics (pediatric allergy and pediatric pulmonary) whereas our initial recruitment efforts had focused on the general pediatric clinic alone. As a result, we nearly doubled our recruitment, obtaining a total of 21 participants in the study. One of the main objectives of this pilot project was to assess the feasibility of the study design and we have learned some valuable information about recruitment strategies for pediatric asthma studies at the medical center. While we did not achieve our recruitment goal, we believe we have enough data for a formal power calculation. Additionally, we have valuable experience in recruitment and have learned that the recruitment base for a larger study will need to include the area outside of Hershey.

Study Population Descriptive and Outcome Data: We have just completed data cleaning and reran our genotyping. Over the last couple weeks, we compiled some of the baseline data and are starting to evaluate outcome data. Our recruitment enrolled a largely white sample of children with a mean age of 10.5 years (Table 1). The majority of the children that were enrolled have allergic asthma. Surprisingly, the forced expiratory flow in one second (FEV1), an objective measure of airway function, was fairly normal in the children we enrolled. The mean of the percent predicted FEV1 for these children was 102.3 with a range of 71-118%. This suggests that either the asthma was under better control for the children we enrolled or this measure is not a sensitive outcome for us to use in a larger trial. We have other pulmonary outcome measures which we have not evaluated yet. Our main outcome variable is the number of symptom free days over the previous two weeks. We collected this data at baseline and planned to look at the change over the three months. Unfortunately, we were not able to collect these data on a few of the participants and many did not answer this question clearly. However,

we were able to obtain data on a large number of the participants at baseline. The mean number of symptom free days was 7.1 and the range was 1 to 14. This is an excellent spread and bodes well for interventions. Based on this preliminary information it appears that we will need to clarify the question so that all parents answer it. We will also need to make sure to move it earlier in the questionnaire so it is not skipped or missed. This was another important lesson that we learned from this pilot study.

Skin Testing and Genotyping: All of our participants have demonstrated allergy to at least one of the antigens we tested demonstrating the important role of allergy in pediatric asthma (Table 2).

The genotyping of the participants was initially challenging due to a lack of a sample of control genes to which we could compare the clustering of the alleles. Our initial run results changed a little as we enrolled more participants. Genomic DNA was purified from saliva collected by Oragene DNA kit (DNA Genotek, Inc., Ottawa, Ontario, Canada). Following normal isolation protocol, concentrations were obtained via NanoDrop ND-1000 (ThermoFisher Scientific, Wilmington, DE USA) and samples were normalized to 5 ng/uL. Using predesigned and validated TaqMan® SNP Genotyping Assay (AppliedBiosystems, Foster City, CA) for 7 targets of interest, DNA was amplified in 384-well plate reactions with total volume of 5 uL using an ABI QuantStudio 12KFlex and Genotyping setup, with the following conditions: 50°C for 2 min; 95°C for 10 min; and 40 cycles of 95°C for 15 seconds and 60°C for 1 min followed by post-read stage. Reactions included Taqman Genotyping Master Mix (1× final concentration), 900nM for each primer, 200nM for each probe, and 10 ng of DNA. Negative controls (no DNA template) were run on every plate. Genotypes were assigned by the automatic calling feature of the allelic discrimination option in QuantStudio 12K Flex Software v1.2.2 (Applied Biosystems, Foster City, CA).

As we obtained more samples our ability to determine the allele clustering improved, and our results stabilized. We have been able to genotype the 7 planned gene variations (Table 3). This information was shared with the “intervention” group so that the differences in medication response could be taken into account in medical management decisions. We are just starting to unblind the data to look at any differences in outcomes between the intervention and control groups.

Early Conclusions: In summary, the no cost extension has permitted us to achieve several important goals. First, our enrollment nearly doubled. Second, we have learned many valuable lessons which will help with the construction of the larger project and serve as a formal basis for power calculations. In addition, we have learned important factors about recruitment: letters work better than ads in the clinic, school time slows enrollment, and allergy season/ allergy testing method affects enrollment (participants do not want to withhold the allergy medication as required for formal skin testing). We have also learned that skin testing might pose challenges for enrollment and perhaps serum sensitization determination would be a better choice in a larger project. Lastly, we have collected important control samples for genotyping that will increase the turn-over for results for participants and will help in allelic discrimination determinations.

Table 1: Baseline Demographics (N=21)

	N (%)
Type of housing child lives in	
Single family home	15 (71.4%)
Multifamily home (duplex, townhome, etc.)	4 (19%)
Apartment	2 (9.5%)
Child's health insurance	
Private health insurance	13 (61.9%)
Public health insurance	7 (33.3%)
None	1 (4.8%)
Total household income	
Under \$20,000	5 (23.8%)
Between \$20,000 and \$39,999	2 (9.5%)
Between \$40,000 and \$59,999	1 (4.8%)
Between \$60,000 and \$79,999	1 (4.8%)
Between \$80,000 and \$99,999	3 (14.3%)
Between \$100,000 and \$119,999	2 (9.5%)
Between \$120,000 and \$139,999	1 (4.8%)
Between \$140,000 and \$159,999	1 (4.8%)
Between \$180,000 and \$199,999	2 (9.5%)
Over \$200,000	3 (14.3%)
Child race	
Missing	1 (%)
White	18 (90%)
Black, African-American	2 (10%)
Child of Hispanic or Latino descent	
No	20 (95.2%)
Yes	1 (4.8%)
Marital status of parents	
Married, living together	13 (61.9%)
Married, living apart	1 (4.8%)
Not married, living alone	6 (28.6%)
Other (Widow)	1 (4.8%)

Table 2: Distribution of Allergen Sensitization (N=21)

	N (%)
Dust mite	
No	4 (19%)
Yes	17 (81%)
Cat	
No	4 (19%)
Yes	17 (81%)
Dog	
No	3 (14.3%)
Yes	18 (85.7%)
Mouse	
No	7 (33.3%)
Yes	14 (66.7%)
Alternaria	
No	6 (28.6%)
Yes	15 (71.4%)
Aspergillus	
No	7 (33.3%)
Yes	14 (66.7%)
Cladosporium	
No	4 (19%)
Yes	17 (81%)
Cockroach	
No	6 (28.6%)
Yes	15 (71.4%)
Any allergy	
Yes	21 (100%)

Table 3: Genotype Distributions (N=21)

	N (%)
ADRB2	
AA	9 (42.9%)
AG	7 (33.3%)
GG	5 (23.8%)
ALOX5AP	
AA	8 (38.1%)
AG	9 (42.9%)
GG	4 (19%)
ARG1	
AA	10 (47.6%)
AG	8 (38.1%)
GG	3 (14.3%)
LTA4H	
AA	15 (71.4%)
AG	5 (23.8%)
GG	1 (4.8%)
FCER2	
AA	6 (28.6%)
AG	12 (57.1%)
GG	3 (14.3%)
TBX21	
CC	16 (76.2%)
CG	5 (23.8%)
LTC4S	
AA	14 (66.7%)
AC	5 (23.8%)
CC	2 (9.5%)

Research Project 8: Project Title and Purpose

Immune and Radiologic Correlates of Lung Injury Following Stereotactic Body Radiosurgery – Stereotactic body radiotherapy is a relatively new technique that is being used for the definitive treatment of Stage I Non-small-cell lung carcinoma (NSCLC). The purpose of this project is to assess pulmonary toxicity due to this modality. We examine the temporal relationship between T-cell/inflammatory cytokine profiles and lung function damage by pulmonary function tests as well as radiographic changes (CT Scans as well as CT SPECT Scans). If the T-cell/inflammatory cytokine profiles exhibit manifestation prior to functional and radiographic changes, it can be used as a predictive indicator and hence allow for intervention to lessen the effect of radiation therapy.

Summary of Research Completed

This project was deleted. No health research grant funds were spent on this project.

Research Project 9: Project Title and Purpose

Development of Pharmacophore-based QSAR of Sphingosine Kinase 1 Inhibitors – The long-term goal of this project is to develop potent sphingosine kinase 1 (SphK1) inhibitors (SKI) as effective therapeutic agents for pancreatic cancer (PC). In this regard, we have already identified several novel lead “drug-like” SphK1 inhibitors. However, before any target-specific lead compound is used as a therapeutic agent in human studies, a lead compound must be optimized to maximize the therapeutic index and minimize side effects. To optimize these SKI lead compounds, we will develop 3D pharmacophore models for the SKI lead compounds and perform quantitative structure activity relationship (QSAR) analyses. We will then test the *in vitro* efficacy of the QSAR analysis generated “hit” SKI compounds.

Duration of Project

5/1/2012 – 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 10: Project Title and Purpose

Pharmacoeconomic Analysis of Current and Alternative Reimbursement Models for Actinic Keratoses – This project describes a 3-phase project to investigate *and* implement an innovative method of reimbursement for a common and important skin condition, actinic keratoses (AK). AK are premalignant skin lesions; the management by various means, is the most common outpatient dermatologic procedure and in 2004 accounted for over \$1 billion in direct healthcare costs. The overall goals of this project are to describe the current patterns in AK management, to develop an alternative model for reimbursement of AK and to test this model in a real-world clinical setting to ensure quality patient care and financial feasibility.

Anticipated Duration of Project

7/15/2013 – 6/30/2015

Project Overview

Aim 1. Define the current regional management patterns and costs of actinic keratosis (AK). The goal of Aim 1 is to establish the pattern of resource utilization for individuals with AK and with Highmark insurance. This insurance cohort was chosen since patients treated at Hershey Medical Center are included and billing data will be available for comparison when the alternative model is applied and assessed in Aim 3. The patterns of resource utilization will be compared to a larger proprietary database, MarketScan[®], which will serve as a measure of the generalizability.

Aim 2. Develop a bundled care reimbursement model for actinic keratosis.

The bundled care model is a method of reimbursement in which a “global fee” is paid for all the care a patient requires for a certain disease in a defined period of time. In this portion of the project, data from Aim 1 will be used to develop a theoretical bundled care model. The model will be assessed by applying it to the Highmark and MarketScan[®] patient datasets.

Aim 3. Evaluate the real-world application of a bundled payment reimbursement model for the management of actinic keratosis.

The bundled care model has been suggested to encourage innovative, cooperative, and efficient care. There have been no studies of alternative reimbursement models for AK or any other dermatologic condition. Utilizing the data from Aims 1 and 2, the goal of Aim 3 is to apply the model in a real-world setting and assess the financial feasibility of this model while ensuring quality patient care. It is critical to assess the feasibility in order to ensure quality patient care and fair provider compensation.

Principal Investigator

Joslyn S. Kirby, MD
Assistant Professor
Milton S. Hershey Medical Center
Dept of Dermatology
500 University Dr, PO Box 850
Mail Code HU14
Hershey, PA 17033

Other Participating Researchers

Douglas Leslie, PhD; Jeffrey Miller, MD, MBA; James Marks, MD; David Shupp, MD; Sara Ferguson, MD; Amy Behe, RN; Sidone Keener, RN; Amy Longenecker, RN; Diane Thiboutot, MD— employed by Milton S. Hershey Medical Center

Expected Research Outcomes and Benefits

With the rise in health care costs associated with an aging population, there is a significant need to preemptively study potential alternative models prior to broad legislative changes or changes in reimbursement. It has been several years since the economic impact of AK has been described. In the early phase of this project, the description of AK-related healthcare utilization and costs will result in a more current estimate for this important and common premalignant condition.

Several pharmacoeconomic studies have evaluated AK treatments; however, none have investigated management with the influence of alternative reimbursement models. This project will utilize both regional and national datasets to develop and apply a novel reimbursement model.

There have been investigation and modeling of alternative reimbursement models however there

is limited literature describing the application and assessment of these models. It is expected that this project, through implementation of the theoretical model, will allow investigators to refine not only the AK-specific model, but will also inform and refine the model-development process for other indications in dermatology as well as other disciplines. Additionally, prior to the implementation phase of this project, a systematic review of AK treatment guidelines will be performed, which will be valuable information for physicians treating AK.

Summary of Research Completed

Aim 1. Define the current regional management patterns and costs of actinic keratosis (AK).

Methods: Claim data were obtained from Highmark Incorporated, a large private insurer for the mid-Atlantic region, and included data for enrollees in the western and central Pennsylvania region. Paid claims for the period of January 2010 to December 2012 were used. Individuals with one or more claims for AK (International Classification of Diseases [ICD]-9 702.X) and that were continuously enrolled were included and their demographic, utilization and cost data were extracted using ICD-9 and common procedural terminology [CPT] codes. Prescription treatments were limited to common dermatologic medications and were identified by drug name and national drug code number. Costs were measured from the perspective of the healthcare system (paid by the patient and the insurer). All costs were adjusted for inflation based on the medical care component of the consumer price index reported by the Bureau of Labor Statistics and are reported in 2010 US dollars. The study was considered exempt by the Penn State Hershey Medical Center Institutional Review Board (IRB).

Analysis: Descriptive statistics were performed for several variables including age, gender, utilization and cost. Comparisons of the continuous outcome variables were made using a *t* test or ANOVA test with a Tukey correction for groups of three or more. Comparisons of proportions were made using the Chi-square test. Cost data are often highly skewed, however, mean costs rather than median costs are more easily interpreted and in large datasets parametric tests of the mean can be robust. Winsorization is a form of data transformation that alters the values above and below a designated percentile to the percentile value to maintain the direction but diminish the effect of outliers. Multivariate linear regression with a generalized linear model with a log link and gamma distribution was used. Marginal effects of the independent variables were calculated. Bootstrapping was performed with samples of 500 observations and 1,000 repetitions to estimate of the 95% confidence interval of the mean. Stata 13 (StataCorp LP, College Station, TX) was used for multivariate analyses and bootstrapping. SAS 9.3 (SAS Institute, Inc., Cary, NC) was used for all other analyses. All statistical tests were two-sided and a *p*-value less than .05 was considered statistically significant.

Results: The study population included a total of 95,294 individuals who were continuously enrolled for the three-year period (January 2010 to December 2012) and had at least one claim for AK. There were over 400,000 claims for the study population and 89.1% (383,030/430,050) of these were associated with a dermatologist. The mean (standard deviation [SD]) age for the population was 66.5 (12) years and 90.6% (86,304) were 50 years or older. Men comprised 53.7% (51,173) of the population. The mean age (SD) for men was slightly higher than for women (68 ± 11 vs. 65 ± 13; *p*<.0001). Non-melanoma skin cancer (NMSC) was diagnosed,

during or prior to the study period, in 13,849 (14.5%) patients. The total cost of AK-related care for the three-year period, comprised of prescription and outpatient claims (visits and procedures), was \$40,719,495. Outpatient visits accounted for \$37,230,232 (91.4%) of the total three-year cost (Table 1). Annual visits were the most common frequency and accounted for 67.1% of all visit frequencies. The next most common visit frequency was twice annual visits, which accounted for 23.2% of all visit frequencies. Destructive procedures [CPT 17000, 17003, and 17004] accounted for 44.9% or \$18,277,451 of the three-year total outpatient cost. The three-year mean (SD; CI) total outpatient cost for those treated with extensive destruction (17004) once or more in the three-year study period was \$948 (1 025; 912-984), compared to \$431 (577; 428-435) for those that never had a claim for 17004 ($p < 0.0001$). Prescription claims for AK-related medications accounted for \$3,489,263 (8.6%) of the total three-year cost for the cohort (Table 1). Only 6.0% or 5,756 patients utilized a topical AK therapy in the three-year period. Prescription field therapy was 2.3 times (OR [95%CI] = 2.3[2.2-2.5], $p < 0.0001$) more frequent for patients with a history of NMSC than those without.

There were statistically significant differences in annual costs by gender, age, and history of NMSC (Table 2). A multivariate analysis was performed and showed that the treatment-related variables with the largest marginal effects on the mean three-year total cost per patient were the use of extensive cryotherapy (17004) or prescription field therapy. Both older age and male gender were associated with higher mean three-year total cost per patient. All other variables being equal, men had a mean cost that was about \$56 higher, and for every year of age the mean three-year cost increased by \$0.44. The mean (SD) and median annual costs per person were \$281.78 (342.80) and \$187.62, respectively and three-year costs per person were \$447.74 (604.28) and \$258.49, respectively. After Winsorization at the first and ninety-ninth percentiles, the mean (SD; CI) three-year total cost per patient was shifted from \$448 (604; 445-453) to \$413 (457; 411-416) ($p < 0.0001$). After Winsorization at the fifth and ninety-fifth percentiles, the mean (SD) three-year total cost per patient was \$384 (348; 382-386) ($p < 0.0001$). If the annual cost per patient was capped at the 99th percentile for each year, the cost savings for 2010, 2011, and 2012 would have been \$406,826, \$421,852, and \$391,819, respectively. Similarly, if the annual cost per patient was capped at the 95th percentile, the cost savings for 2010, 2011, and 2012 would have been \$1,256,364, \$1,324,909, and \$1,289,057, respectively. This work has been submitted as a manuscript.

Aim 2. Develop a bundled care reimbursement model for actinic keratosis.

Methods: Claim data for January 2010 to December 2012 from Highmark Incorporated were used again in this study. A test dataset and a validation dataset were randomly derived from the original dataset of 95,254 patients. The test dataset was used to develop the alternative payment models and the validation dataset was sequestered until the models were finalized. This dataset was then used to cross-validate the performance of the models. Additionally, a 1:1 random sample of patients with a claim for AK will be selected from the MarketScan Clinical Claims[®] database, which is a national sample that is representative of privately insured individuals. Inpatient claims and pathology services were not included. A patient-level episode rather than an episode-level bundled payment was selected as this removes the incentive to repeatedly care for the same issue, potentially on the same person. The most common visit frequency was an annual visit, so a one-year time horizon was used. Median-based payments and unadjusted mean-based

models were developed using descriptive statistics of the test population. Mean-based models with adjustments were based on the intercept of the multivariate model incorporating the variables for which payment adjustment would be made. Costs were measured from the perspective of the healthcare system (by patient and insurer). As above, all costs were adjusted for inflation and are reported in 2010 US dollars. The annual total cost was calculated for each year that the member was enrolled.

One-way sensitivity analyses were performed on the mean-based models and the following variables were allowed to vary from the upper to the lower 95% confidence limit of the mean: unadjusted mean, adjusted mean, gender-based adjustment, and NMSC-based adjustment. The 95% confidence limits were obtained with a bootstrapping analysis that utilized sampling of 500 patients for 1,000 repetitions. A one-way sensitivity analysis was also performed to investigate changes in the proportion of people with a history of NMSC or male gender. Stata 13 (StataCorp LP, College Station, TX) was used for multivariate analyses. SAS 9.3 was used for all other analyses. All statistical tests were two-sided and a *p*-value less than .05 was considered statistically significant. The study was considered exempt by the Penn State Hershey Medical Center IRB.

Results: Several mean- and percentile-based models were developed on the test population and applied to the validation population. The unadjusted mean payment population was \$262 and the adjusted mean payment was \$209.38. The median for annual payment was \$173.17 and the 75th percentile was \$305.84. For some models adjustments were developed for gender and a history of NMSC. The risk-adjusted payments for male gender and NMSC were \$50.87 and \$233.80, respectively. The median-based payments had the largest decrease in total annual spending, while the 75th percentile-based payments had the largest increase in spending for the population (Table 3). Application of these bundled payments to the Highmark validation population showed similar spending patterns, with the 75th percentile-based models had similar higher spending and the median-based payments had the most decrease in spending. The mean-based models had a small increase (0.4-0.7%) in spending compared to the test population. Sensitivity analyses demonstrated that if the payment for the “unadjusted mean” model was varied to the upper and lower 95% confidence limit of the mean then the total cost of the model would increase by 1.1% or decrease by 0.7%. This was repeated for the risk-adjusted mean-based model; changes in the mean base resulted in an increase by 1.1% or decrease by 1.4%. For the mean-based model if the values of the gender and NMSC-history adjustments were varied then the cost increased or decreased by 0.9% and 0.5%, respectively. In median and 75th percentile models if the gender-based adjustment was varied then the cost would vary by 2.2% or 1.4%, respectively. If the NMSC-based adjustment was varied then the cost would vary by 0.3% or 0.2%, respectively. The proportion of men was 53.9% and accounted for 57.0% of the patient-years. If the proportion of men in the study population was increased to 60%, and the ratio of patient-years was preserved, then the change in cost for an increase to 45,740 patient-years and gender-adjusted models would have an increase of \$236,588 or 0.9-1.5% depending on the model. Similarly, if the proportion of people with any history of NMSC was increased to 15% and the ratio of patient-years was maintained then the risk-adjusted model for 9,178 patient-years would increase the total cost of NMSC-adjusted models by \$407,747 or 1.6-2.5%. The analysis will be complete after the models are applied to the Market Scan sample.

Aim 3. Evaluate the real-world application of a bundled payment reimbursement model for the management of actinic keratosis.

Methods: A cohort of dermatologists at the Hershey Medical Center will be recruited to contribute data from their eligible and consenting patient population. The eligible patient population will include adults (over age 18) with AK who are existing patients and consent to participate. Patients and providers will participate in a prospective, observational, pilot study of the implementation of a bundled payment model for AK management. Given the complexity of the billing mechanisms and contracts with insurers these processes will not be changed for this pilot project. The bundled payment model will be implemented as a simulation while preserving the current billing services. This will allow physicians and staff to work within the model construct, preserve department revenue, and collect data on the costs fee-for-service model to which comparison will be made. Subjects will be enrolled for 15 months from the time they enter the study. The primary outcome variable will be the proportion of patients under the care of these dermatologists with total direct costs less than the simulated bundled payment. With 95% confidence and 80% power, approximately 270 patients will be needed to detect a change from 35% to 20% of patients with costs above the bundled payment limit. Patient satisfaction will be assessed using the Consumer Assessment of Healthcare Provider and System survey, which is a validated tool that is used nationally to assess patient care quality. The survey will be administered at baseline and at the end of the study.

Results: Study is ongoing. Modifications that will facilitate patient enrollment have been submitted to the IRB. Physician participants have been recruited.

Table 1. Mean annual utilization and cost per patient for AK management during the three-year study period.

	Patients with 1 or more claims, n (%)	Annual use per patient mean \pm SD	Total annual cost per patient, (\$) mean \pm SD
Outpatient Office Visit			
Any	56,786 (59.6)		
New	15,810 (27.8)	1.0 \pm 0.2	322 \pm 359
Consultation	1,500 (2.6)	1.0 \pm 0.1	437 \pm 458
Established	45,649 (80.4)	1.9 \pm 1.4	322 \pm 321
Prescription Therapy			
Any	5,756 (6.0)		
Fluorouracil	3,141 (54.6)	1.2 \pm 0.5	311 \pm 387
Imiquimod	2,050 (35.6)	1.3 \pm 0.8	823 \pm 574
Diclofenac	613 (10.6)	1.2 \pm 0.7	414 \pm 457
Ingenol	271 (4.7)	1.1 \pm 0.3	847 \pm 1,001
Cryotherapy			
Any	80,851 (84.8)		
One lesion (17000)	78,717 (97.4)	1.3 \pm 0.7	129 \pm 84
Two or more (17003)	50,233 (62.1)	1.3 \pm 0.7	56 \pm 69
15 or more (17004)	5,558 (6.9)	1.4 \pm 0.7	299 \pm 181

SD = standard deviation

Table 2. Mean total three-year and annual total cost per patient for AK management.

Characteristic	Mean total cost per patient, (\$)				Mean annual cost per patient (\$)			
	Mean	SD	p-value	Median	Mean	SD	p-value	Median
Gender								
Female	347	458	<0.0001	209	248	292	<0.0001	173
Male	497	653		295	307	374		202
Age Group								
Less than 30	185	203	<0.0001	121	177	181	<0.0001	113
30-39	270	380		170	224	282		154
40-49	324	433		200	248	298		173
50-59	396	541		234	276	341		186
60-69	443	639		255	292	383		191
70-79	470	566		285	281	310		189
80 and older	457	572		271	289	330		191
Diagnosis of NMSC								
No	352	428	<0.0001	218	244	266	<0.0001	176
Yes	875	986		586	445	534		291

NMSC = non-melanoma skin cancer, SD = standard deviation

Table 3. Comparison of Several Theoretical Bundled Payment Models based on development from the test population

Base Model	Model Adjustments	Total 3-year spending	Difference (% of actual)*	Patient-years within model payment**, n(%)	Providers within model payment^, n(%)
Total cost of care [actual]		18,788,963	[ref]	n/a	
75th percentile payment	None	21,860,526	+3,063,611 (+16.3)	53,662 (75.1)	377 (75.9)
	Gender- and NMSC-	25,688,894	+6,891,623 (+36.7)	58,053 (81.2)	412 (82.9)
50th percentile payment	None	12,377,672	-6,415,793 (-34.1)	35,429 (49.6)	192 (38.6)
	Indirect payment (\$64)	16,952,200	-1,836,9763 (-9.8)	46,990 (65.2)	308 (62.0)
	Gender- and NMSC-	16,206,040	-2,587,781 (-13.8)	44,008 (61.6)	280 (56.3)
Mean payment	None	18,726,974	-68,801 (-0.4)	49,239 (66.9)	342(68.8)
	Gender- and NMSC-	18,794,223	-540 (-0.003)	49,239 (68.9)	329 (66.2)
	Gender- and NMSC- and 2% discount	18,418,338	-376,309 (-2.0)	48,659 (68.1)	326 (65.6)

NMSC = nonmelanoma skin cancer, **indicates that if patients' actual annual cost did not change then it would be less than the bundled model payment cost (denominator = 71,477).

^indicates that if these providers had the same/similar patients and costs of care then the spending for the providers patient group would be less than the bundled model (denominator 497 providers). A negative difference indicates that less was spent with the alternative model; a positive difference indicates the alternative model had a higher cost

Research Project 11: Project Title and Purpose

Development and Testing of a Novel Simulation Technology for Fracture Treatment – The purpose of the project is to develop and test a novel simulation technology for complex bone fracture treatment. The simulation technology will estimate the biomechanics involved in the fixation and repair of complex fractures. Simulations will be validated by biomechanical experiments. Potential uses include determining the optimal surgical fixation strategy for repair of difficult fractures in civilian and military patients. The project also aims to test the ability of the simulation to convey fundamental treatment concepts in a group of orthopaedists.

Anticipated Duration of Project

7/15/2013 – 6/30/2015

Project Overview

Surgical treatment of severe long bone and periarticular fractures are routinely encountered in trauma centers following high energy incidents. Additionally, they are increasingly common in the modern battlefield and are now the largest source of long term disability in military personnel. Surgical fixation in order to promote fracture union and optimal recovery can be extremely challenging. Despite robust education of orthopaedic surgeons, a deep understanding of the interaction of biomechanical variables inherent in fracture fixation is often lacking, leading to suboptimal repair and potentially failure of fracture healing or early implant failure. Our overall hypothesis is that a novel computer simulation technology, developed in this project, can improve treatment of these fractures through improved cognitive understanding by orthopedic surgeons. The hypothesis will be tested through the following two aims:

Aim 1: To create and validate a novel computer simulation technology which enables interactive 3D visualization of how repair variations affect clinically important results. Computer models of three fracture fixation scenarios will be constructed using commercially available finite element software. Clinically important parameters such as strain at the healing fracture gap will be estimated during physiologic activity. Experimental validation will be carried out using biomechanical tests in human cadaver specimens or synthetic surrogates. The custom simulation software will then be developed.

Aim 2: To develop a novel training program based on the novel simulation technology, and to test its effectiveness in a small group of orthopaedists. An expert- and self-led program will be developed. Twenty orthopaedists will go through the training. A control group of ten orthopaedists will undergo an alternative traditional training intervention. An analysis of covariance (ANCOVA) model will be used to assess differences between experimental and control groups.

Use of the simulation technology developed in this project includes potentially being able to determine the optimal personalized treatment strategy for patients suffering complex bone fractures.

Principal Investigator

Gregory S Lewis, PhD
Assistant Professor
Penn State College of Medicine
500 University Dr
Mail Code H089
Hershey, PA 17033

Other Participating Researchers

John S. Reid, MD; Kevin P. Black, MD; Ling Rothrock, PhD; Elizabeth H. Sinz, MD; Hwa Bok Wee, PhD; Evan Roush, BS – employed by Pennsylvania State University
Sandra Jarvis-Selinger, PhD – employed by Univ. of British Columbia, Canada

Expected Research Outcomes and Benefits

Each year in the United States alone, hundreds of thousands of patients experience poor healing following a fracture. Extremity fractures have become the leading source of long term disability in our military personnel. Proper fixation of severe fractures is challenging because of the complex biomechanics involved. We believe the simulation technology developed in this project potentially can change the way in which these cognitive skills are taught in medical schools and in other learning environments. In addition we expect the developed technology could be used to determine optimal treatment for a patient based on their individual 3D imaged bone shape and fracture characteristics.

Summary of Research Completed

Aim 1: To create and validate a novel computer simulation technology which enables interactive 3D visualization of how repair variations affect clinically important results.

We successfully applied for hardware and electronic CAD files from industry (Synthes) for two fracture plates and associated screws:

We are first constructing the software to simulate fracture repair for a simple transverse diaphyseal fracture. A corresponding experimental validation is being set up also. We successfully applied for obtaining CAD files and actual hardware for two different plates (quantity 5 of each plate), and locking and non-locking type screws (quantity 40+ for each). The CAD files are in an STL format which gives the precise 3D surface geometry of each part. Obtaining the files and hardware was a multiple-month process that involved formal applications to Synthes, follow up communications, and intellectual property agreements.

Creation of finite element models

The CAD files were imported into commercial software (CAD: Solidworks; finite element analysis: Abaqus, Dassault Systems Waltham MA) in order to construct finite element computer models of the biomechanics of fracture repair. In addition to models of the hardware, a model of the shaft of a long bone (such as the tibia) with transverse fracture was created. The models were meshed with the software using several different meshing densities and both hexahedral and tetrahedral 3D elements.

Careful planning of strategy was completed for the modeling approach in order to make the models more conducive to incorporation in the simulation. For example, we have experimented with making finite element meshes and partitioning so that fracture repair configurations could be adjusted through change of the materials of certain elements (e.g. at a possible screw location) without changing the meshes themselves.

Complex material behavior was modeled for the bone, given its anisotropic nature and the role this might play in fracture fixation biomechanics. Specifically, a transverse isotropic model was used and proper values for material properties were identified from the literature. This modeling process required review of the elastic behavior of bone and learning how an appropriate model

can be incorporated within the software. Additionally in order to model the stage of tissue healing, the fracture gap was modelled as either fibrous, cartilaginous, or mineralized bone with the different material properties obtained from the literature.

Another aspect of the modeling that needed to be figured out was how to accurately model the interactions between the plate and bone, screws and bone, and especially screws and plate. The two different types of screws, locking and non-locking, behave very differently because the locking screws are rigidly interlocked with the plate. The conventional non-locking screws were modeled to enable frictional sliding between their heads and the plate, which required some adjustments to the contact conditions until the model worked correctly. Boundary conditions describing constraints and loading were applied based on reviewing the literature and considering the physiological conditions of a typical patient post-trauma surgery.

Simulations of the fracture repair construct under mechanical loading were successfully completed for dozens of different combinations of plate type, screw type, plate and screw material, screw number and locations, size of the fracture, stage of healing, and loading conditions. Interesting results were obtained, for example showing the large effects on plate and screw stresses of switching from locking to nonlocking screws (Fig. 1). We determined how to save the resulting stresses, strains, and displacements for all exterior surface elements into a database that could be retrieved by the new simulation software, described below.

Development of the simulation software

Custom-written code was written in Matlab (Mathworks, Natick MA). This code drove a new graphical user interface, which enables the user of the software to interact with the finite element model results described above (Fig. 2). The new software shows the 3D bone, fracture, and hardware and the view can be rotated in 3D and zoomed in and out. The software program shows typical surgical options which are available to a surgeon for fracture fixation surgery, including length of plate, screw type, plate and screw material, and screw number and locations. Additional options are available to change the load applied to the construct, the stage of tissue healing, as well as how to view the results.

When the user changes one or more of these options and clicks the update button, the software retrieves numerical results from the finite element modeling database described earlier. Results are converted to color representations for stress and strain, and colors are displayed at the correct location on the model. In addition, code was developed so that displacement results are applied to the nodes of the model and the model deforms realistically.

Development of this software involved several versions, with new versions developed to add capability and speed. The current version of the software is able to update the model results in less than 0.5 seconds on a moderate-performance laptop computer (Intel Core 2 Duo 2.66 GHz) upon clicking the update button.

Experimental validation

Design of physical biomechanical experiments for validating the above finite element modeling

has been completed. Synthetic bones were purchased. Fixtures for axial and torsional loading of the constructs were designed and fabricated from aluminum and steel, and preparations are underway for actual testing.

Aim 2: To develop a novel training program based on the novel simulation technology, and to test its effectiveness in a small group of orthopaedists.

The training program has not been explicitly developed, but we have had meetings with experts in simulation and curriculum design to start this process. In addition the software was planned so that fundamental concepts involved in fracture fixation are conveyed to the user.

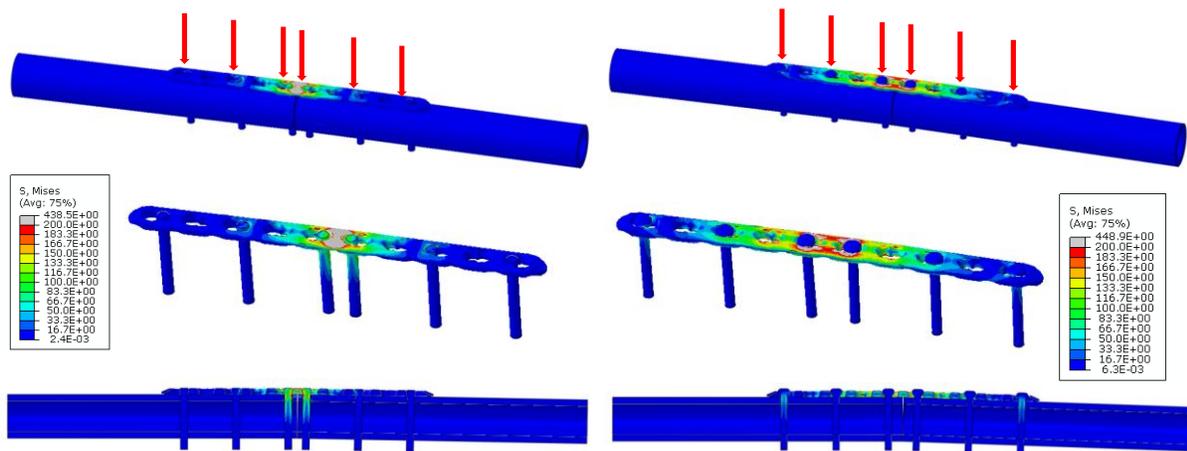


Figure 1. Finite element models and example results for stress under axial compressive loading. A construct using locking screws (left side) is compared with a construct using traditional non-locking screws (right side).

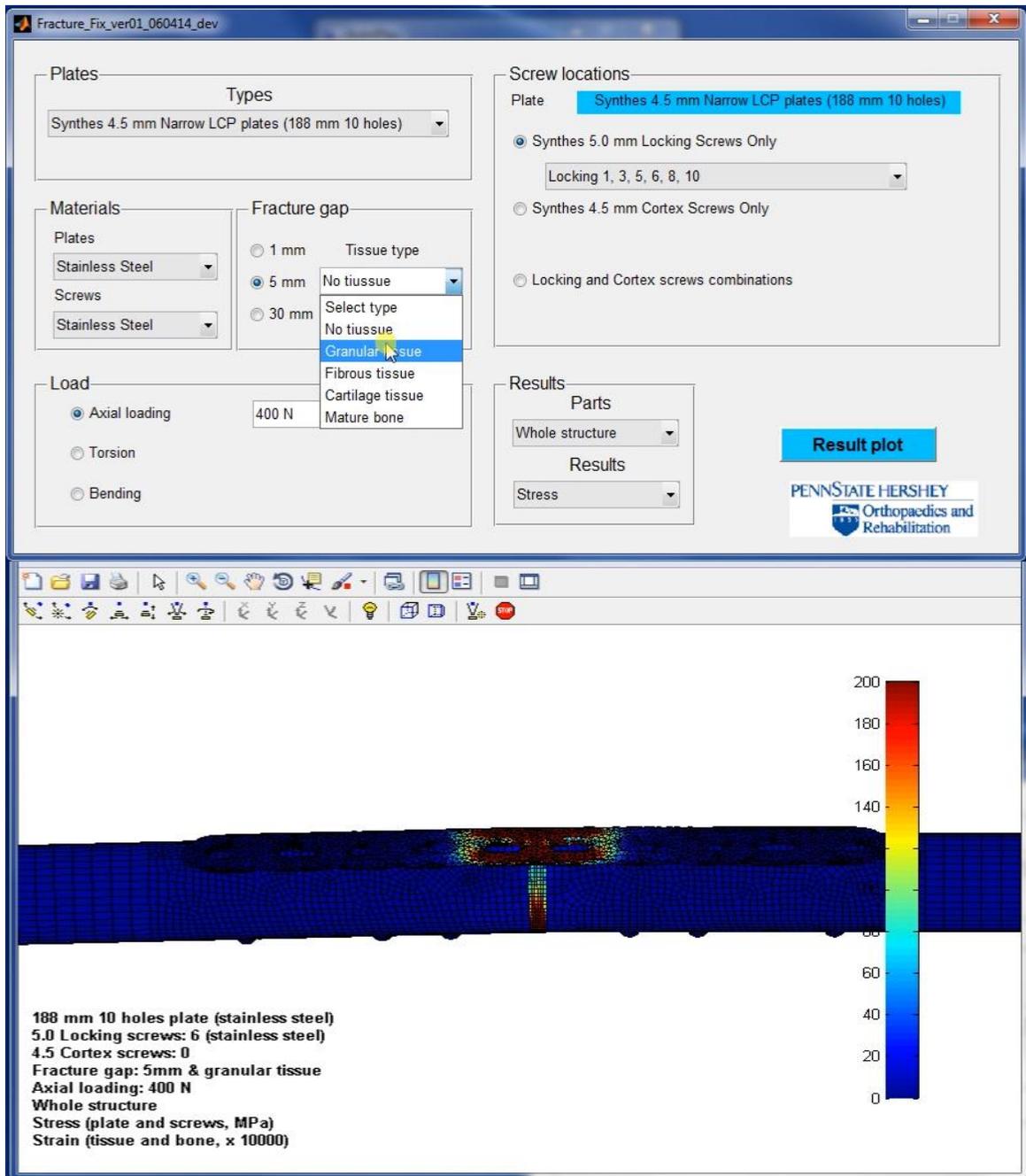


Figure 2. Newly developed graphical user interface software which allows user to make adjustments to fracture fixation and immediately update results for stress, strain, and displacement and visually interact with the model in 3D.

Research Project 12: Project Title and Purpose

Does Enhanced Cannabinoid Receptor 1 (CB₁) Signaling Increase the Risk of Drug Abuse? The primary focus of my laboratory is to determine how activation of the endocannabinoid system modulates alcohol and opiate addiction. We have produced mutant mice expressing a hypersensitive for cannabinoid receptor 1 (CB₁). These S426A/S430A mutant mice also show an exaggerated and prolonged response to the treatment with Δ^9 -THC, the principal psychoactive component of marijuana. The exaggerated and prolonged response to Δ^9 -THC and endocannabinoids establishes these mice as a novel model for testing the effects of enhanced CB₁-mediated endocannabinoid signaling on drug addiction *in vivo*. This project will determine whether “overactive” cannabinoid signaling potentiates reward and dependence for alcohol and morphine. Completion of this study will shed light on whether abuse of these drugs might be potentiated by consumption of marijuana or medical conditions such as obesity that lead to increased systemic levels of endocannabinoids.

Anticipated Duration of Project

7/15/2013 – 6/30/2015

Project Overview

Alcohol and opioid addiction afflict eight million and two million Americans, respectively. Currently available treatments, while helpful to some addicts, are still plagued by high rates of relapse. Much work has shown that the endocannabinoid system is involved in modulating dependence on, tolerance to, and the motivation to consume alcohol and opioids. Antagonism or genetic deletion of cannabinoid receptor 1 (CB₁) signaling attenuates self-administration of and dependence on these drugs of abuse. However, the effects of chronically upregulated endocannabinoid signaling on reward and dependence for these drugs are less well understood.

Specific Aims:

Specific Aim 1: Does enhanced cannabinoid signaling modulate the rewarding effects and motivation to consume alcohol and morphine?

Specific Aim 2: Is dependence for morphine and alcohol increased in S426A/S430A mice?

This project will give us new insight into the role that “overactive” endocannabinoid signaling plays in modulating alcohol and opioid addiction. *We will test the hypothesis that increased activation of the endocannabinoid system potentiates reward and dependence for ethanol and morphine.* This goal will be accomplished using a novel line of CB₁ receptor “knock-in” (S426A/S430A) mice that exhibit exaggerated and prolonged responses to the acute effects of both anandamide and delta-9-tetrahydrocannabinol (Δ^9 -THC). Alcohol and morphine dependence will be determined by measuring the severity of somatic symptoms associated with withdrawal from these drugs in S426A/S430A mutants. The rewarding properties of alcohol and ethanol will be measured by performing conditioned place preference for these drugs. Microdialysis will be performed to measure morphine and alcohol-stimulated dopamine release in the nucleus accumbens shell of S426A/S430A mice. Due to the widespread recreational and therapeutic use of cannabis, our study investigating whether “overactive” cannabinoid signaling

might increase the risk of addiction for other drugs has importance relevance for human health.

Principal Investigator

Daniel Morgan, PhD
Assistant Professor
Penn State College of Medicine
Mailcode H187, Room C2850
500 University Drive
Hershey, PA 17033

Other Participating Researchers

David J Marcus, BS ,Caitlin Nealon, Jessica Biegler, BS – employed by Pennsylvania State University

Expected Research Outcomes and Benefits

We expect to see an overall enhancement in the rewarding effects of morphine and alcohol in S426A/S430A mutants in the conditioned place preference (CPP) assay (~10-15% increase in preference for the drug-paired side for the S426A/S430A mutants). We expect that S426A/S430A mutants will exhibit increased ethanol-induced dopamine release in the nucleus accumbens shell. We also expect that withdrawal from ethanol and morphine will be more severe in S426A/S430A mutants.

This work will explore the possibility that excess cannabinoid signaling (ECS) increases addiction to other drugs of abuse such as alcohol and morphine. The levels of endocannabinoids in the plasma and/or tissue are increased in a variety of disease states including obesity, neurodegeneration, and cancer, making the results of this project relevant to assessing the drug abuse potential of patients suffering from these disorders. This project also has important implications for the “gateway hypothesis” that marijuana usage leads to dependence on other drugs of abuse since excessive cannabinoid system activation occurs in chronic users of marijuana. Our mutant mice mimic the increased ECS occurring in patients with elevated basal endocannabinoids levels and also users of marijuana. If we determine that these mice have an increased risk for alcohol and morphine addiction then this suggests that drug addiction risk may be increased for these human populations. As a consequence, the goals of our project have significant translational implications for human drug addicts.

Summary of Research Completed

Specific Aim 1: Does enhanced cannabinoid signaling modulate the rewarding effects and motivation to consume alcohol and morphine?

Experiment 1: Test the hypothesis that conditioned place preference (CPP) for alcohol is greater in S426A/S430A mutants.

Experimental Design and Methods:

The rewarding effects of ethanol, cocaine, and morphine have been determined by measuring conditioned place preference using a three-chamber conditioned place preference apparatus (Med-Associates, St. Albans, VT). A single day of pre-conditioning habituation was conducted during which mice were allowed to freely explore all three chambers of the apparatus for 30 minutes per day. Following habituation (Day 3), a 30 minute pre-test where mice had access to all three chambers was performed. Mice showing a pre-conditioning place preference of >65% of time spent in either large chamber or >40% time spent in the neutral chamber were excluded from further analysis. During the conditioning phase mice were administered either saline or 2 g/kg of ethanol via intraperitoneal (i.p.) injection and placed in one of the two context-specific sides of the apparatus for 5 minutes (post-injection). On alternating days, mice were injected with saline and placed in the other context-specific side chamber. Once daily conditioning sessions were done for 10 consecutive days, post-testing for ethanol place preference was performed the day after the last conditioning trial by allowing mice access to all three chambers to determine amount of time spent in the ethanol and saline-paired chambers.

Results:

We have determined that CPP (reward) for 2 g/kg ethanol is abolished in S426A/S430A mutant mice (N=14). In contrast, we do observe modest but significant CPP in wild-type littermate control mice (N=27) (Figure 1). The finding that CPP for ethanol was absent was quite unexpected since our original hypothesis was that ethanol CPP would be enhanced in these mutant mice. One possible explanation for this unexpected finding is that the S426A/S430A mutation might cause down-regulation of the CB₁ receptor or endocannabinoid levels in the nucleus accumbens. Dopaminergic signaling from the ventral tegmental area (VTA) to the nucleus accumbens is required for acute reinforcing effects of ethanol and other abused drugs in the CPP paradigm. In our initial published report we did not observe widespread down-regulation of CB₁ or endocannabinoids (2-arachidonoyl glycerol and N-arachidonylethanolamine) in cerebellum, striatum, hippocampus, and cortex of S426A/S430A mutant mice relative to wild-type controls. However, we have not specifically examined CB₁ and endocannabinoid levels in the nucleus accumbens.

Experiment 2: Test the hypothesis that conditioned place preference (CPP) for morphine is greater in S426A/S430A mutants.

Experimental Design and Methods:

The rewarding effects of morphine and cocaine were also assessed using the CPP paradigm as described above with the following noted exceptions. During conditioning trials mice were given saline and either 10 mg/kg morphine or 10 mg/kg cocaine. The duration of each daily conditioning trial for cocaine and morphine was 30 minutes and the number of days used for conditioning was eight (cocaine) and ten (morphine).

Results:

We detect normal and robust CPP for 10 mg/kg morphine that is not different between S426A/S430A mutants and wild-type littermate controls (Figure 2). We expected to observe increased CPP for morphine in S426A/S430A mutant mice so this result was not fully anticipated. As a control for our ethanol and morphine CPP results we chose to examine CPP for

cocaine because it is well established that reward for 10 mg/kg cocaine in the CPP paradigm is not mediated by the endocannabinoid system. Therefore, we expected to see robust CPP (positive control) for cocaine that was not different between S426A/S430A mutant and wild-type mice (negative control). This anticipated result is precisely what we observed. In this experiment we find strong CPP for 10 mg/kg that is not different between the S426A/S430A mutant and wild-type mice (Figure 3). The presence of strong CPP for both 10 mg/kg morphine and 10 mg/kg cocaine indicates that the absence of ethanol CPP in S426A/S430A mutants is specific for that particular drug. Moreover, the fact that we are able to detect CPP for ethanol in wild-type mice and robust CPP for morphine and cocaine in both genotypes supports the idea that the lack of CPP for ethanol in S426A/S430A mutant mice is a real effect.

Experiment 3A: Test the hypothesis that ethanol-stimulated dopamine (DA) release in nucleus accumbens is greater in S426A/S430A mutants. Experiment 3B: Perform morphine-stimulated dopamine microdialysis in the nucleus accumbens of S426A/S430A mutants.

Methods and Experimental Design:

In collaboration with Dr. Andras Hajnal's laboratory dialysates from the nucleus accumbens were collected by microdialysis. Guide cannulae and microdialysis probes were implanted in the nucleus accumbens via stereotaxic surgery (coordinates: bregma: +1.1 mm, lateral: +0.6, ventral: -4.0 mm). Postmortem verification of probe location was used to confirm that the microdialysis probe was correctly inserted into the nucleus accumbens. Mice were allowed to recover for 7 days before microdialysis was conducted. Dialysates were collected every 20 minutes from the nucleus accumbens of S426A/S430A mutant and wild-type mice for 60 minutes prior to and 180 minutes after administration of 2 g/kg of ethanol (i.p.) and 10 mg/kg morphine (s.c.).

Results:

Ethanol and morphine-stimulated microdialysates (N=10 per genotype) have been collected. These samples are currently being analyzed for dopamine and norepinephrine content.

Specific Aim 2: Test overall hypothesis that dependence on morphine and alcohol is greater in S426A/S430A mice.

Experiment 4: Test the hypothesis that alcohol dependence will be greater in S426A/S430A mutants.

No progress was made for this experiment in this award period.

Experiment 5: Test the hypothesis that morphine dependence is greater in S426A/S430A mutants.

Experimental Design and Methods:

S426A/S430A mice and wild-type littermates were made dependent on morphine via implantation of 75 mg morphine pellets in the nape of the neck. After 72 hours of morphine administration via the pellet, withdrawal was precipitated by an i.p. injection of 1 mg/kg of naloxone, a mu opioid receptor antagonist. Morphine dependence was determined by measuring

the symptoms of morphine withdrawal including jumping, wet-dog shakes, paw tremors, and diarrhea. Withdrawal symptoms were recorded for 60 minutes and quantified for alternating five minute time bins (5-10, 15-20, 25-30, 35-40, 45-50, and 55-60 minutes).

Results:

In a first cohort of animals we detected a non-significant trend towards more severe precipitated withdrawal from morphine in S426A/S430A mutants compared to wild-type littermate controls. A trend towards increased numbers of naloxone-precipitated jumps, wet-dog shakes, and paw tremors but not diarrhea was observed (Figure 4).

Figure 1. CPP for 2 g/kg of ethanol was examined in S426A/S430A mutants and wild-type (WT) littermates. Pre-conditioning (black bars) and post-conditioning (red bars) place preference was measured before and after ten days of place preference conditioning for 2 g/kg of ethanol. Error bars represent the standard error of the mean (SEM) and data analyses were performed using two-way ANOVA with Bonferroni post-hoc tests (* $p < 0.05$). Sample sizes for each group are in parentheses.

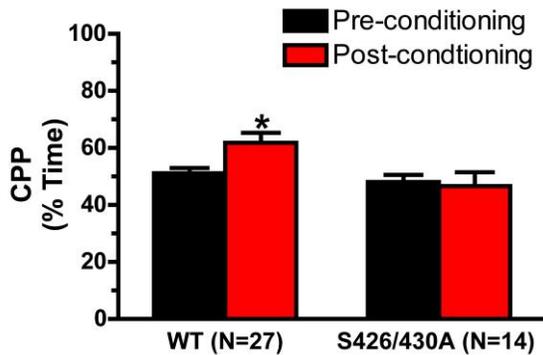


Figure 2. CPP for 10 mg/kg of morphine was examined in S426A/S430A mutants and wild-type (WT) littermates. Pre-conditioning (black bars) and post-conditioning (red bars) place preference was measured before and after ten days of place preference conditioning for 10 mg/kg of morphine. Error bars represent the SEM and data analyses were performed using two-way ANOVA with Bonferroni post-hoc tests (* $p < 0.05$). Sample sizes for each group are in parentheses.

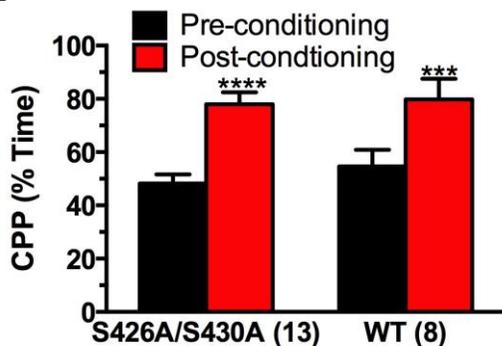


Figure 3. CPP for 10 mg/kg of cocaine was examined in S426A/S430A mutants and wild-type (WT) littermates. Pre-conditioning (black bars) and post-conditioning (red bars) place preference was measured before and after eight days of place preference conditioning for 10 mg/kg of cocaine. Error bars represent the SEM and data analyses were performed using two-way ANOVA with Bonferroni post-hoc tests (* $p < 0.05$). Sample sizes for each group are in parentheses.

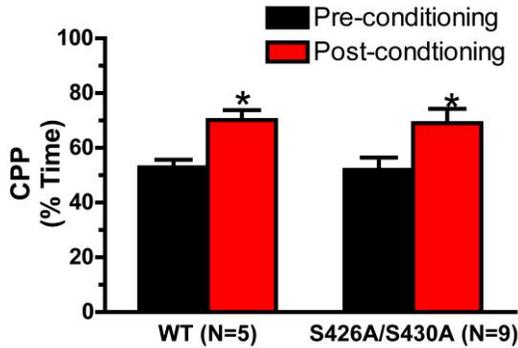
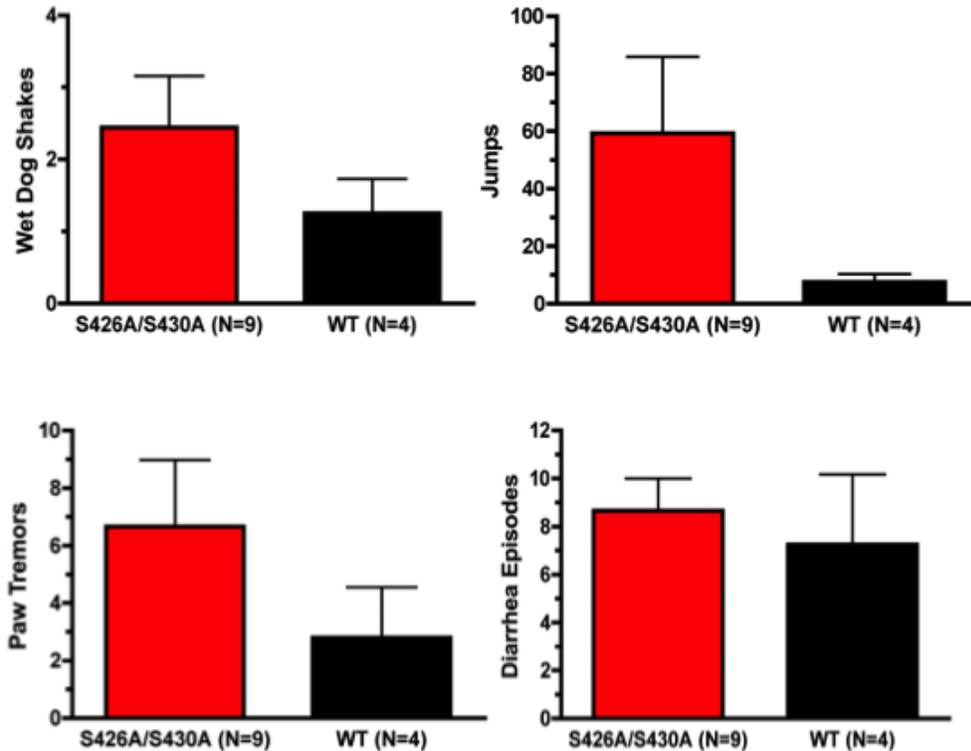


Figure 4. S426A/S430A mutants (red bars) exhibited a non-significant trend toward an increased number naloxone-precipitated wet dog shakes (A), jumps (B), and paw tremors (C) relative to wild-type mice (WT black bars). The number of diarrhea episodes (D) was not different between genotypes. Error bars represent the standard error of the mean (SEM) and data analysis was performed using unpaired Student's t-tests.



Research Project 13: Project Title and Purpose

Mechanisms Underlying Retrovirus Genomic RNA Packaging – The purpose of this research project is to elucidate the molecular mechanisms underlying retroviral RNA encapsidation using the oncoretrovirus Rous sarcoma virus (RSV) as a model system. RSV is unique among retroviruses because viral RNA packaging appears to be initiated in the nucleus of infected cells. Understanding the detailed mechanism of retroviral genome incorporation is important for the development of novel antiretroviral therapies that disrupt genomic RNA encapsidation and to design optimal gene therapy delivery using retroviral vectors.

Anticipated Duration of Project

7/15/2013 – 12/31/2014

Project Overview

Retroviruses commandeer cellular machinery to assemble new virus particles that propagate infection within the host and spread the virus to uninfected individuals. This project focuses on the molecular aspects of retrovirus assembly, which is orchestrated by the retroviral Gag protein. We are investigating the earliest steps in assembly, in which the Gag protein recognizes and selectively packages its RNA genome using the avian Rous sarcoma virus (RSV) as a model retrovirus. Traditionally, retroviral genome packaging was thought to be a late step in the assembly process, occurring at the plasma membrane just prior to escape of the virion from the plasma membrane. Challenging this dogma, we discovered that the RSV Gag protein traffics transiently through the nucleus, a step required for efficient genome packaging. RNA binding serves as a trigger for nuclear export of the viral ribonucleoprotein complex and subsequent transport to the plasma membrane. Our newest data suggest that genome recognition occurs co-transcriptionally, a paradigm-shifting model that will be tested in this project. The specific research aims are:

- Specific Aim 1. Examine whether the subnuclear localization of Gag is RNA-dependent.
- Specific Aim 2. Determine the purity of the subnuclear fractions (nucleoplasm, loose chromatin, dense chromatin, and insoluble/nuclear matrix) using an expanded collection of cellular markers.
- Specific Aim 3. Examine whether the Gag p10 domain associates with chromatin during mitosis.
- Specific Aim 4. Perform preliminary proteomic analysis of Gag-interacting factors in subnuclear fractions.
- Specific Aim 5. Determine whether viral RNAs bound by Gag and Tap are distinguishable using microscopy.

Principal Investigator

Leslie J Parent, MD
Professor, Division of Infectious Diseases, H036
Penn State Hershey Medical Center
500 University Drive
Hershey, PA 17033

Other Participating Researchers

None

Expected Research Outcomes and Benefits

Retroviruses cause cancer and immunodeficiency syndromes in humans and animals. Human retroviral diseases include AIDS, caused by human immunodeficiency virus, and adult T cell leukemia, caused by human T cell lymphotropic virus type I. Animal retroviruses have long served as important model systems for understanding the mechanisms these viruses use to replicate. First among these model retroviruses was Rous sarcoma virus (RSV), which was initially demonstrated in 1910 to induce solid tumors. Despite its study for over a century, our laboratory was the first to discover that nuclear trafficking of the RSV Gag protein was required for genome incorporation, which is essential for virus replication. More recently, we found that RSV Gag localizes to the host chromatin fraction within the nucleus, suggesting a link between chromatin interactions and genome encapsidation. Our work may define a novel paradigm for how retroviruses select and package their RNA genomes.

The purpose of this project is to identify chromatin-bound host factors that contribute to RSV genome incorporation and to dissect the mechanism by which RSV Gag recognizes and selects its genomic RNA at the chromatin interface. By elucidating the host factors that contribute to retroviral genome packaging, our work may provide insights into design of new inhibitors that treat human retrovirus infections. In addition, gene therapy vectors design may be improved by understanding in greater detail the roles host and viral factors play in genomic RNA recognition.

Summary of Research Completed

Specific Aim 1. Examine whether the subnuclear localization of RSV Gag is RNA-dependent. The RSV Gag protein shuttles through the nucleus in a CRM1-dependent manner. We previously identified nuclear localization signals (NLSs) in the MA (matrix) and nucleocapsid (NC) domains and a nuclear export signal (NES) in the p10 domain. Treatment of cells expressing the RSV Gag protein with leptomycin B, a CRM1 inhibitor, or mutation of critical hydrophobic residues in the p10 NES (mutant L219A.Gag) leads to the accumulation of Gag in discrete nuclear foci. To determine whether the Gag foci colocalized with other cellular subnuclear foci, we performed a series of immunofluorescence experiments and found that there was no overlap with splicing speckles, paraspeckles, Cajal bodies, or PML bodies. However, we did find through subcellular fractionation experiments that Gag was present in the chromatin-associated protein fraction. This finding, as well as our observation that RSV RNA and the L219A.Gag protein co-localize to a high degree within the same nuclear foci, led us to hypothesize that RSV Gag may bind to RSV RNA at the site of transcription from the integrated proviral DNA, explaining why Gag is associated with chromatin.

The localization of Gag in nuclear foci, even in the absence of the full-length viral RNA genome, suggested that Gag may tether to a cellular factor in the nucleus. We proposed this tether is most likely a host protein or RNA. Aim 4 focused on potential protein partners of Gag. In Aim 1, we examined whether Gag foci formation was dependent on ongoing RNA synthesis. We treated

cells that had been transfected with a L219A.Gag expression vector with Actinomycin D, an inhibitor of RNA polymerases. We found that the number of foci was markedly reduced within 30 minutes of treatment with Actinomycin D (Fig. 1). With longer treatment (60 minutes and 3 hours), Gag foci were more significantly reduced. Therefore, Gag foci are sensitive to RNA synthesis inhibitors, suggesting that Gag foci depend on ongoing RNA synthesis.

Specific Aim 2. Determine the purity of the subnuclear fractions (nucleoplasm, loose chromatin, dense chromatin, and insoluble/nuclear matrix) using an expanded collection of cellular markers.

We performed a series of experiments using different subcellular fractionation methods to determine which would yield the cleanest fractions with cellular proteins located in the expected fractions. We determined that the subcellular fractionation procedure by Chase et al., PLoS Pathog 7(9): e1002187, reproducibly gave clean fractions for cellular markers (Fig. 2). This procedure employs differential salt concentrations to separate transcriptionally active chromatin (extracted with 150mM NaCl) from heterochromatin (extracted with 500mM NaCl). To determine the purity of the fractions, we performed western blotting with antibodies against GAPDH (cytoplasm), CRM1 (nucleoplasm and cytoplasm), Med4 (nucleoplasm and chromatin), RCC1 (chromatin) and Histone2B (chromatin). We also performed a chromatin fractionation by Henikoff et al., Genome Res 19: 460–469. Using this method, we were able to effectively separate transcriptionally-active and repressed chromatin fractions and we found that RSV Gag was found in both fractions. Cellular markers are being analyzed currently to assess purity of the Henikoff procedure.

Specific Aim 3. Examine whether the Gag p10 domain associates with chromatin during mitosis. The Gag protein of prototype foamy virus, a nonpathogenic retrovirus that infects humans and primates, tethers to chromatin to assist in targeting the incoming reverse transcription complex to the site of proviral integration. Because we found evidence that RSV Gag localizes to the chromatin fraction, we tested the hypothesis that RSV Gag remains associated with chromatin during mitosis. We synchronized cells expressing Gag with serum starvation and examined cells in different phases of mitosis. We found that Gag was adjacent to chromatin, but was not fully colocalized with the cellular DNA, which was stained with DAPI. Therefore, we concluded that RSV Gag is associated to chromatin-associated proteins but Gag does not remain bound directly to chromatin throughout mitosis. Thus the mechanisms of chromatin association for RSV Gag and foamy virus Gag are different.

Specific Aim 4. Perform preliminary proteomic analysis of Gag-interacting factors in subnuclear fractions.

To gain insight into the mechanism underlying Gag-viral RNA interactions in the nucleus, we used a proteomic approach to identify Gag binding partners. In collaboration with Dr. Mamuka Kvaratskhelia (The Ohio State University), we performed an affinity pulldown using His₆-tagged Gag protein purified from E. coli mixed with nuclear lysates from DF1 chicken fibroblast. As a control, we excluded proteins that bound to His₆-Gag.ΔNC, since the NC domain is required for Gag to accumulate in nuclear foci and co-localize with RNA. We identified 23 nuclear proteins that bound specifically to His₆-Gag (Fig. 3). Transcriptional activators were highly enriched (p=0.009, 4.5-fold enriched), as were RNA binding proteins, transcriptional regulators, and chromatin regulators. Interestingly, we found that 4 members of the Mediator complex, which is

a multi-protein complex essential for pol II transcription, were identified in the preliminary pulldown experiment. So far, we have preliminarily confirmed the interaction of Gag with Med20 using co-immunoprecipitation. After examining the transcriptional regulators, we will focus on MAGOH and SNRPE, factors involved in splicing, to determine whether they are involved in the competition between splicing and packaging of full-length viral RNAs in the nucleus.

Specific Aim 5. Determine whether viral RNAs bound by Gag and Tap/NXF1 are distinguishable using microscopy.

This aim has not been completed. We have not yet found an antibody against Tap/NXF1 that recognizes the Tap protein in quail cells. There may be enough difference in the amino acid sequences of the human Tap and the quail Tap proteins that there is not sufficient cross-reactivity. There is a quail Tap expression construct, and we plan to try to express it in QT6 cells to determine whether we see the protein co-localize with RSV RNA in the nucleus. We tried expressing a CMV-driven human Tap expression vector in quail cells, but the protein was undetectable by confocal microscopy. This set of experiments will require further optimization.

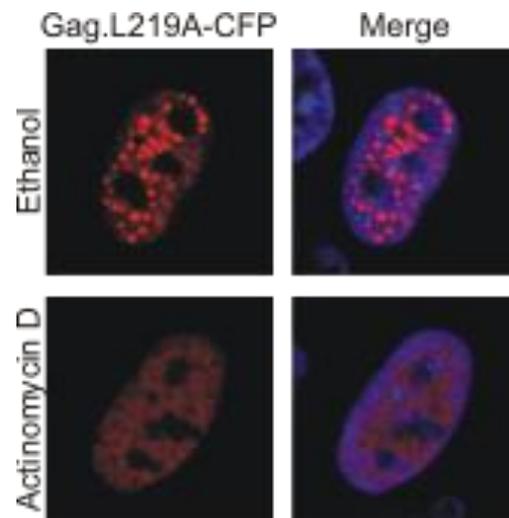


Fig. 1. Confocal microscopy images of cells expressing RSV Gag-CFP (false-colored red) treated with ethanol or Actinomycin B (dissolved in ethanol). The nuclei of the cells were stained blue with DAPI. There are numerous discrete foci of Gag evident in the top (ethanol) panel of cells whereas the Gag protein appears more diffuse with a marked reduction in the number of nuclear foci.

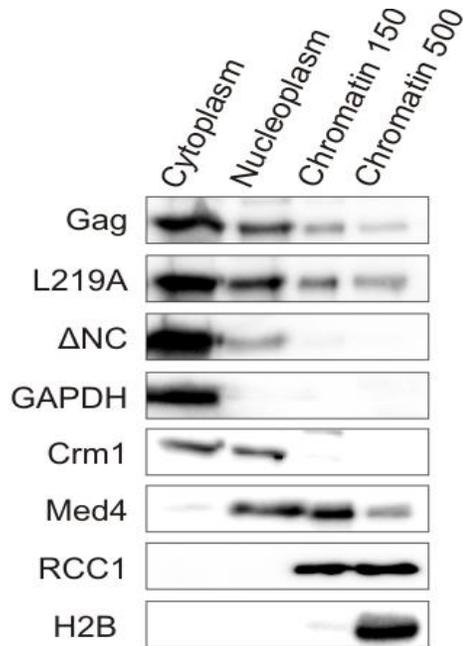


Fig. 2. Immunoblot analysis of chromatin salt extraction fractionation showing Gag, L219A, Gag ΔNC and cellular control proteins to demonstrate the purity of the fractions. The fractionation protocol used in this experiment was from Chase et al., PLoS Pathog 7(9): e1002187 .

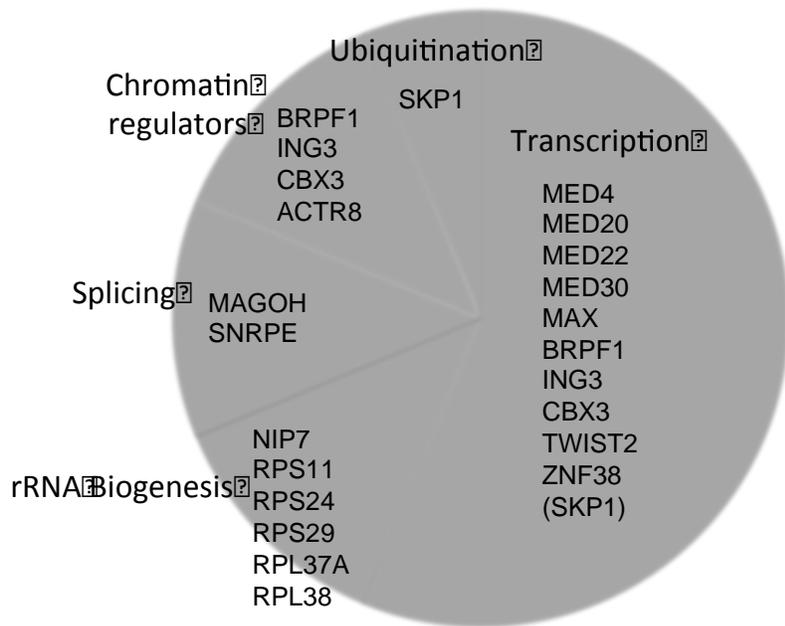


Fig. 3. (Top) Functional classification of His₆Gag-interacting nuclear factors analyzed by DAVID Database for Annotation, Visualization, and Integrated Discovery (www.david.abcc.ncifcrf.gov). SKP1 is involved in ubiquitination and transcription; therefore, it is listed in both categories.

Research Project 14: Project Title and Purpose

Mechanism of Red Cell Invasion by Malaria: Role of Antibodies and Complement – Plasmodium falciparum is the most virulent of all the human malaria parasites and there is no effective vaccine against this parasite. The invasion of red blood cells (RBCs) is an appealing target for vaccine development. However, this goal has been elusive. Although volunteers immunized with parasite antigens produce antibodies that can inhibit the parasite *in vitro*, there is no protection *in vivo*. We believe the explanation for this discrepancy lies in the parasite's ability to use complement activation for its own advantage, but this has never been demonstrated. Therefore, in this study we will test whether complement activation via antibody-dependent or independent mechanisms enhances the invasion of RBCs by the malaria.

Duration of Project

12/17/2013 – 6/30/2014

Project Overview

Despite the fact that experimental vaccines elicit antibodies that inhibit parasite growth and RBC invasion *in vitro*, there is no significant efficacy when vaccinated volunteers are challenged with malaria parasites. We believe the explanation for this discrepancy lies in the parasite's ability to use complement activation for its own advantage. Most vaccines elicit complement fixing antibodies. When complement is activated (C3b and C4b) and opsonizes a pathogen, it allows the pathogen to bind to the complement receptor 1 (CR1) on the RBC surface. For most pathogens this is detrimental, but for malaria it is beneficial because it brings the pathogen to its ultimate target, the RBC. Since most malaria parasites also express a native CR1 ligand called PfRh4, there is the possibility that it acts synergistically with complement. This study is intended to gather additional preliminary data to support the resubmission of an NIH application. The reviewers of our original NIH proposal indicated that the preliminary data was not convincing. This project is intended to address three additional key areas of criticism or reviewer's suggestions.

Specific aims:

- 1) Expand the use of human antibodies: We will test antibodies from a subset of volunteers immunized with anti-merozoite vaccines for their ability to enhance or inhibit RBC invasion in the absence or presence of complement.
- 2) Characterize the interaction with parasite ligands: We will select and screen clones from a *P. falciparum* line that was transfected with a plasmid containing a fragment of the 5' region of the PfRh4 gene for disruption of PfRh4 expression. Once we identify a clone that lacks expression of PfRh4 we propose to use this clone in the resubmission application to determine the contribution of PfRh4 to complement-mediated invasion.
- 3) Gather preliminary data on the suitability of a human CR1 transgenic mouse model: We will determine the effect of passive immunization with anti-malaria antibodies and of complement depletion on the level of parasitemia of malaria-infected human CR1 transgenic mice. If we are successful, we will propose to use this model to investigate more broadly the contribution of complement and antibody-mediated invasion to the efficacy of different avenues of

immunization.

Principal Investigator

José A. Stoute, MD
Associate Professor
Penn State Hershey Medical Center
500 University Drive
Hershey, PA 17033

Other Participating Researchers

Mary E. Landmesser, BS, Sergei Biryukov, MS – employed by Penn State College of Medicine

Expected Research Outcomes and Benefits

If we are correct, our findings will lead to a paradigm shift in the way potential malaria vaccines are screened and designed. If we demonstrate that complement can enhance RBC invasion, growth inhibitory assays will have to account for this to predict in vivo protection. In addition, an important corollary of our hypothesis is that contrary to current dogma, IgG subclasses that are less effective at fixing complement such as IgG2 may be more effective than those that do fix complement such as IgG1 and IgG3. Therefore, malaria vaccines will have to be designed to bias the immune response towards the production of non-complement fixing antibodies. If we are correct, this will be the first time that malaria parasites are shown to use the immune response of the host for their own advantage and constitutes a form of immune evasion.

Summary of Research Completed

This project is currently on hold. No funds were spent on the project during the reporting period.

Research Project 15: Project Title and Purpose

Oncogenic Drivers/Mutations/Biomarkers in Screening Participants for Early Detection of Lung Cancer – The purpose of this project is to identify and validate oncogenic drivers (mutations and biomarkers) for the early detection of lung cancer in individuals at high-risk for lung cancer and characterize the nodules and preneoplastic lesions in screening participants with presence/absence of oncogenic drivers (mutations and biomarkers). Testing for mutations/gene rearrangements as biomarkers for early detection with integration of molecular diagnostics in tissue samples/bronchial lavage will be key to develop and validate personalized management approaches. The current project is designed to pave the way for individualizing therapy for these subjects at high risk for development of lung cancer.

Anticipated Duration of Project

3/17/2014 – 12/31/2015

Project Overview

We hypothesize that approximately 25% of the lung cancer screening participants will have the presence of lung nodules requiring intervention. The tissue/blood from these high-risk individuals will be analyzed for presence/absence of oncogenic drivers (mutations and biomarkers). In addition, the study will characterize the nodules and preneoplastic lesions to develop personalized approaches to management and follow-up and identify/validate new drugable targets. This will subsequently lead to individualized approaches to their management.

Specific Aim 1: To identify and validate oncogenic drivers (mutations and biomarkers) for the early detection of lung cancer in individuals at high-risk for lung cancer. These mutations/biomarkers of interest include those seen in EGFR, HER2, BRAF, MET, MEK1, PIK3CA, PTEN, DDR2.

Specific Aim 2: To characterize the nodules and preneoplastic lesions in screening participants' incidental scans with presence/absence of oncogenic drivers (mutations and biomarkers) in an attempt to develop personalized approaches to management and follow-up.

Principal Investigator

Chandra P. Belani, MD
Miriam Beckner Distinguished Professor of Medicine
The Pennsylvania State University College of Medicine
Penn State Hershey Cancer Institute
500 University Drive, CH72
Hershey, PA 17033-0850

Other Participating Researchers

Anna Salzberg, PhD; Shawn J. Rice, MD; Arthur S. Berg, PhD; Jennifer W. Toth, MD; Rickeshwar Maharaj, MD; Michael F. Reed, MD; Xin Liu, MD, PhD, Christopher Gilbert, MD; Jay Zhu, PhD – employed by Pennsylvania State University College of Medicine

Expected Research Outcomes and Benefits

The aggressive and heterogeneous nature of lung cancer requires efforts to reduce mortality from this cancer through the use of screening. The overall research outcomes and benefits of this project are to identify driver mutations/gene rearrangements as biomarkers for early detection of lung cancers, with integration of molecular diagnostics in tissue samples from the resected/biopsied nodules on screening or incidental CT scans. These studies and analyses will be essential to develop and validate personalized management approaches. We are expecting to precisely outline the key driver mutations/gene rearrangements and define drugable targets that will allow appropriate and personalized management leading to more effective treatment outcomes.

Summary of Research Completed

The PI of this study, Dr. Belani, has initiated a clinical trial entitled “Oncogenic Driver Mutations/Biomarkers in Screening Participants for Early Detection of Lung Cancer” (PSHCI-12-109, IRB #41855). The target accrual is 75 patients and we have successfully enrolled 27 patients to date. These subjects are above 18 years old with smoking history ≥ 30 pack years. Participants with > 20 pack years of smoking are eligible if >60 years of age. Average smoking pack year is 43.1 among these patients. All the patients have solid nodules, 8 patients have solitary nodules and 19 patients have more than one nodule. All of the 27 patients had pathology reports of the nodules, from either bronchoscopy or surgical removal.

Research blood, FFPE-derived lung nodule tissues as well as clinical data were collected from these patients. 10 nodules were benign while 19 nodules were malignant. Benign nodules (n=10) includes fibrosis and chronic inflammation (n=5), neuroendocrine tumor (n=3), squamous metaplasia (n=1) and inflammatory myofibroblastic tumor (n=1); Malignant nodules (n=19) includes adenocarcinoma (n=13), squamous cell carcinoma (n=5) and large cell carcinoma (n=1). Two patients had two lung nodules, one was benign and one was malignant.

Next generation sequencing technology utilizing Illumina HiSeq platform has been performed to identify the genetic mutations in FFPE sample of a lung nodule DNA from one patient (03-008) collaborating with Dr. James Broach’s laboratory (data pending). In addition, the matching DNA derived from patients’ peripheral blood mononuclear cells (PBMC) was used for identification of somatic mutations in this FFPE sample. The project was funded in March 2014 and thus it is too early to provide any results here. The project is underway and is ahead of schedule.