# IGS NEWSLETTER

SPRING 2015

UNIVERSITY of MARYLAND
SCHOOL OF MEDICINE
INSTITUTE FOR GENOME SCIENCES

801 W BALTIMORE ST, BALTIMORE MD 21201 | WWW.IGS.UMARYLAND.EDU | 410.706.1481

# Greetings Colleagues,

In my last Director's column in fall 2014, I mentioned highlights of our infectious disease research. Now, I'd like to talk about our work with human genomics and its implications for personalized medicine. This research is capturing more attention since President Obama's January 2015 launch of a new "Precision Medicine" Initiative, an effort to "revolutionize how we improve health and treat disease." His 2016 budget includes a new $215 million investment for the NIH, NCI, FDA and others to support this effort. We believe that genomics will play a major role in precision medicine, and here at SOM, our faculty are involved in a number of research programs that are focused on better understanding the impact of DNA sequence variants on human biology. **Dr. Scott E. Devine** is part of the 1000 Genomes Project, an initiative to identify variants in the human genome that occur at a frequency of at least one in 50 people. Genetic variation explains part of why people vary in their risk for disease (we'll update you about this exciting project later this summer). **Dr. Timothy O'Connor** has been leading research that applies exome analysis to gain new understandings of rare population variations. His work helps scientists look at trends of diseases among various ethnic groups. **Dr. Julie Hotopp** is conducting innovative research involving lateral gene transfer (LGT), the transmission of genetic material between organisms in the absence of sexual reproduction. Her research may indicate novel linkages between LGT and cancers. We will be hearing more about our work with human biology and genomics in upcoming newsletter issues.

Many of these efforts are in collaboration with clinician-scientists at UMB, including the Program in Personalized and Genomic Medicine (PPGM) and the Greenebaum Cancer Center at SOM, as well as the School of Nursing. We are working together to apply genomic analysis in innovative ways to better understand the genomic basis of pain, addiction, and various cancers.

We are hosting **leading genomic scientists** in our **seminar series**, which continues through June. We also offer advanced hands-on bioinformatics and metagenomics **workshops** for those investigators interested in exploring clinical applications and genomics. Please check our educational and **event schedule** for these upcoming programs. There is a wide array of topics and we hope you can join us.

As always, I welcome your feedback,

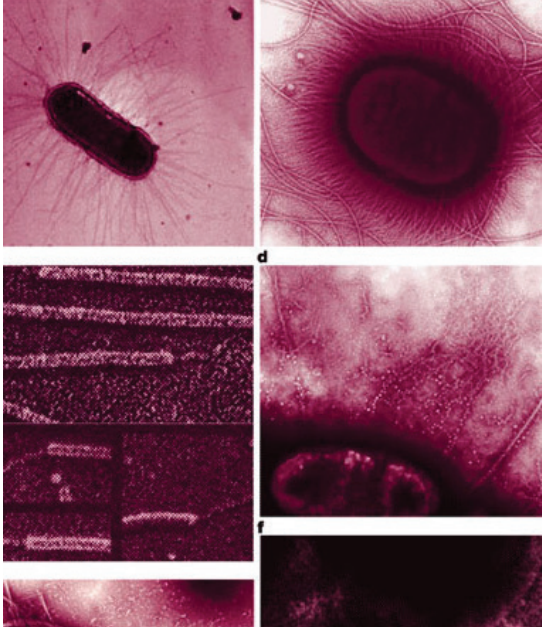
Claire M. Fraser, PhD

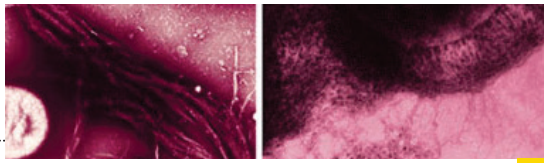
*Professor of Medicine and Microbiology and Immunology*
*Director, Institute for Genome Sciences*
*University of Maryland School of Medicine*

UNIVERSITY of MARYLAND
SCHOOL OF MEDICINE
INSTITUTE FOR GENOME SCIENCES

# "Host, Pathogen and the Microbiome: Determinants of Infectious Disease Outcome"

## Update on IGS' NIAID-Funded Grant





In 2014, IGS was awarded a five year, $15.2 M grant from the National Institute of Allergy and Infectious Diseases (NIAID) called the IGS Genome Center for Infectious Disease (GCID). The IGS GCID applies large-scale genomics and bioinformatics approaches to investigate virulence, drug resistance, and immune evasion, as well as host-microbiome interactions. Claire Fraser, PhD, is leading the grant as Principal Investigator and Administrative Core Director, and Drs. David Rasko and Owen White are the co- Principal Investigators.

Among other aspects of the grant, the funding supports research programs focused on three areas – host/bacterial pathogens and the microbiome; the genomic analysis of fungal pathogenesis; and integrated genomics research in parasitic tropical diseases.

**David Rasko**, Assistant Professor, Microbiology & Immunology, spoke to us about the GCID progress.

***Dr. Rasko, you've been involved with previous versions of NIAID-funded research. If you were going to summarize this latest "generation," what are some distinctions?***

These projects extend the knowledge that we generated from the previous iterations of this funding program, including the **Genome Sequencing Center for Infectious Diseases.** In this project, we are using novel sequencing technologies and data analyses techniques and applying them in unique ways to derive biological and translational findings.

Each of the projects includes the use of novel technologies, as well as samples obtained directly from patients. Previous genomic studies had utilized, for the most part, prototype isolates, however in these studies we are examining samples that are closer to the patient. This provides both an exciting new avenue for the application of genomic medicine, but also creates technical hurdles with the examination of uncharacterized organisms in the host.

Additionally, our center aims to examine the biology of the host-pathogen interaction on a deeper level to understand the biology of the system. This depth of knowledge will provide the potential for the identification of diagnostic and therapeutic targets.

***The GCID is still in early stages, but what can be shared about the research at this point?***

The bacterial project is focused on examining enteric pathogens with multi-omic approaches, including genomics, transcriptomics, proteomics, metabolomics, and others. Two pathogens will be examined, *Vibrio cholerae*, the agent of cholera, and Enterotoxigenic *Escherichia coli* (ETEC). Each of these pathogens causes significant human mortality and morbidity. We are leveraging the placebo arms of two ongoing human vaccine challenge trials to track host-pathogen interaction before, during and after pathogen-induced disease.

Studying these detailed host-pathogen interactions from a systems biology perspective was not possible five years ago.

**(cont. on next page)**

**View papers already published from this funding source**

Technological advancements made it possible. This project takes advantage of the collaboration between the **Center for Vaccine Development (CVD)** and IGS on the University of Maryland, Baltimore campus.

The fungal project, led by **Dr. Bruno**, addresses invasive fungal infections, which are devastating and have extremely high mortality. Moreover, there are very few effective antifungal therapies and no vaccines to prevent them. This project aims to understand the complex interaction between the host and three very different fungal pathogens to allow us to develop desperately needed novel drugs.

The parasite project focuses on malaria and filariasis, both of which are mosquito-borne diseases. A primary and very exciting new emphasis of this project led by **Dr. Silva**, is the characterization of genetic variation in natural populations, and its impact on drug resistance and vaccine efficacy. The filarial project, led by **Dr. Hotopp**, focuses on identifying novel drug targets, which are needed to develop treatment for filariasis.

Also, a major component of this NIAID–funded Center is to leverage IGS educational department and establish new workshops and educations initiatives will be establishing workshops and a large component is continuing educational initiatives to expand application of genomics in research projects that impact global health.

# Lateral Gene Transfer

## Update on Julie Hotopp's Research

Dr. Julie Hotopp, Associate Professor, Department of Microbiology and Immunology at the University of Maryland School of Medicine and the Institute for Genome Sciences, is leading innovative research involving lateral gene transfer (LGT) - the transmission of genetic material between organisms in the absence of sexual reproduction - and human somatic cells.

Dr. Hotopp led a team that found evidence that LGT is possible from bacteria to the cells of the human body, known as human somatic cells. She and her team found that the bacterial DNA was more likely to integrate in the genome in tumor samples than in normal, healthy somatic cells. The phenomenon might play a role in cancer and other diseases associated with DNA damage.

It is possible that LGT mutations play a role in carcinogenesis, or they could simply be passenger mutations. The tumor cells could be proliferating so rapidly that they may be more permissive to lateral gene transfer. It is also possible that the bacteria are causing these mutations because they benefit the bacteria.

Dr. Hotopp is also a research scientist with the University of Maryland Marlene and Stewart Greenebaum Cancer Center.

> **"LGT from bacteria to animals was only described recently,"** says Dr. Hotopp, "**Studies applying this approach to additional cancer genome projects could lead us to a better understanding of the mechanisms of cancer.**"

# Scientists Applying **Exome Analysis** Gain New Understanding of **Rare Population Variations**

Timothy O'Connor, PhD, Assistant Professor, Department of Medicine at SOM and IGS, has led a team using large-scale resequencing analysis to quantify rare and common variations of populations.

Population structure has important implications for understanding global variation in disease prevalence. Although broad-scale patterns of population structure among continental groups are well understood, delimiting recently emerged and fine-scale population structure has received comparatively less attention. As rare variants are more geographically restricted compared to common variants, they may provide a powerful resource to delineate fine-scale patterns of population structure. Their analyses demonstrate that rare variation contains considerable information about fine-scale population structure, and will be a powerful tool to understand recent population demographic history.

# TRACY HAZEN'S **K22 & E. coli**

Tracy Hazen, PhD, a Research Associate in the Rasko Lab, recently received a K22 award from the National Institutes of Health, National Institute for Allergy and Infectious Diseases to investigate the global virulence regulons of *Esherichia coli*.

Tracy joined the Institute for Genome Sciences in 2011 as a postdoctoral fellow in Dr. David Rasko's lab and became a Research Associate in 2014. Her focus has been investigating the diversity of virulence mechanisms of enteropathogenic *E. coli* (EPEC), which is a causative agent of moderate to severe diarrhea among children in developing countries.

For Tracy's K22 application, she hypothesized that the unique combination of the genetic diversity of the *E. coli* adherence factor (EAF) plasmid along with the chromosomal content of the EPEC isolate influences the EPEC virulence regulons, which may impact the severity of infection caused by some EPEC isolates. Her research may provide insights into how other pathogenic *E. coli* use similar virulence mechanisms to cause disease, such as enterohemorrhagic *E. coli* (EHEC), a public health burden in the U.S. Tracy received an impact score of 10 on this very competitive K22 award. The award is intended to support post-doctorates as they begin independent research careers, and will be an extension of work that she has done so far in the Rasko lab.

# ILRI

## INTERNATIONAL COLLABORATIONS

### Consortium to Develop East Coast Fever Vaccine

The Kenya-based International Livestock Research Institute (ILRI) is leading a multi-national consortium to help develop an affordable vaccine for African farmers to protect their livestock against the deadly tick-transmitted parasite *Theileria parva*. The disease it causes, East Coast fever (ECF), kills over one million cattle each year and is a significant contributor to hunger and poverty in sub-Saharan Africa. Joana Silva, PhD, Assistant Professor, Microbiology & Immunology, the Institute for Genome Sciences at the University of Maryland School of Medicine, is working with the organization.

ILRI received $11M from the Bill and Melinda Gates Foundation to fund the first four years of this project, for which the ultimate goal is the development of a new ECF vaccine and have it available within the next 10 years. The existing ECF vaccine costs as much as $12 per animal, and its production if very complex and labor-intensive. "We're developing a new-generation vaccine that will be more affordable to farmers and take less time to produce," said Vish Nene, PhD, director of ILRI's vaccine biosciences and leader of the ECF vaccine project.

The Consortium includes another U.S.-based institution, Washington State University, and the United States Department of Agriculture, as well as the Institute of Tropical Medicine Antwerp, in Belgium, three UK-based institutions, and a Malawi-based center for research on tick-born diseases.

Joana leads the genomics and transcriptomics components of the project. These tasks include the comprehensive description of all protein-coding sequences in the genome, the timing of their expression in the course of the parasite's life cycle, and the characterization of protein diversity in the parasite population. All these factors are essential to the development of an effective and fully protective vaccine, and feed directly to the work of the other groups in the consortium.

# IGS Informatics Project

IGS is also working with ILRI on informatics support for their joint research on livestock pathogens *Theileria parva* and *Mycoplasma mycoides*. The scientific project team is headed by Drs. Joana Silva and **Hervé Tettelin**, and Anup Mahurkar directs the informatics requirements. The group is providing genome assembly, analysis and data dissemination.

The team is creating and maintaining a **portal for the *T. parva* research community** where genomic resources, annotation and protocols can be shared. The newly completed genome re-annotation will be available through a web-based genome visualization tool.

*Mycoplasma mycoides* is a bacterial species of the genus *Mycoplasma*, and is a parasite that lives in ruminants (cattle and goats), causing lung disease. For the *Mycoplasma* project, IGS is collaborating with ILRI scientists on annotation and comparative genomics, then the data will be available to the *Mycoplasma* research community. Instances of the IGS web-based Sybil comparative genomics package will be loaded for mining of genes unique or shared among isolates. IGS will also be providing reverse vaccinology analysis and interpretation to the ILRI scientists in order to identify novel potential vaccine candidates against these mycoplasmas.

"**Increasingly, research communities focusing on organisms of interest find it useful to pool complementary expertise and disseminate genomics and bioinformatics data,**" explained Dr. Tettelin. "With our robust tools and our team's multi-decade bioinformatics experience, we are optimizing exploitation of the data towards curing disease and organizing data dissemination to be most accessible and impactful for these communities."

## THIRD ANNUAL FRONTIERS IN GENOMICS LECTURE

Thursday March 26, 2015 11 am-noon

Jay Shendure, MD, PhD

IGS began the "Frontiers in Genomics" lecture series for our fifth anniversary in 2012 to recognize and discuss innovative scientific discovery in the field of genomics, and we have continued this tradition by hosting scientific leaders in the field. In 2015, we are pleased to be hosting Dr. Shendure, who will speak about "Next Generation Human Genetics." Dr. Shendure is an Associate Professor of Genome Sciences at the University of Washington School of Medicine. He trained under Dr. George Church, Professor of Genetics at Harvard Medical School, who has pioneered the specialized fields of **personal genomics** and **synthetic biology**. Dr. Shendure's laboratory applies new DNA sequencing technologies for human genetics and genomics, focusing primarily on the following: next-generation human genetics; genome contiguity and completeness; massively parallel functional analysis; molecular tagging; synthetic biology and translational genomics. Dr. Shendure states, "Next generation DNA sequencing which is effectively emerging as a broadly enabling 'microscope' for the measurement of biological phenomena in the 21st century." His work is at the forefront of genomics research and as such he has been selected as our **3rd Frontier in Genomics speaker**.

**Reservations** are encouraged – **igs-event@som.umaryland.edu** to ensure a place.

FEATURE

# Supporting Genomics Analysis

## at the University of Maryland, Baltimore

Sean Daugherty, Senior Bioinformatics Analyst
Heather Huot Creasy, Senior Bioinformatics Analyst

*As large-scale sequencing becomes faster, the ability to handle complex informatics analysis becomes ever more critical. Working closely with IGS faculty, UMB academic researchers, private industry researchers and the GRC core are the people working in the Informatics Research Center (IRC) core, headed by Anup Mahurkar, and Dr. Owen White.*

*In this issue, we're featuring two people in the IGS informatics team – Heather and Sean. Both were part of the informatics team at The Institute for Genome Research (TIGR), they share an office and (sometimes) work on joint grants and projects, but often support different faculty and projects.*

### Sean Daugherty

Senior Bioinformatics Analyst, IRC

Sean was hired at IGS soon after the Institute launched in 2007. His background is in biochemistry and, as informatics became more complex over the years, he's learned new tools, techniques and software, continually updating his informatics skills through professional literature, classes and conferences.

Informatics support for faculty research has always been a key part of Sean's work. As IGS has increased its educational outreach curriculum with a variety of genomic-themed workshops, Sean has helped with the logistics, set-up and teaching the workshops. Workshop topics range from programming (PERL, R and database software) to a week-long overview of genomics to complex transcriptomic classes. Over the years, workshop participants have included graduate students and faculty from various UMB professional schools, from other academic research centers, from pharma and industry and from different federal agencies, such as the FDA and USDA.

Sean is part of the informatics support for the **GCID** grant, both in the administrative core for the educational workshops, and in the data dissemination for the research. Faculty leading the GCID projects are producing important data and new insights that can impact the treatment of infectious diseases, so sharing the data in public repositories is a key and challenging part of the project and Sean's expertise adds to its success.

In addition to providing informatics support for the workshops and UMB faculty research, Sean also works on analytical support for external customers, who schedule IRC support through fee-for-service arrangements. Both the GRC and the IRC are UMB core facilities, with a menu of services that can be purchased by any researcher. Customers for the analytical services include academic and industrial researchers, as well as federal agencies.

Sean's greatest challenge? "**Staying current in informatics is always a challenge. The equipment and tools are continually evolving, with new features and approaches, so we need to stay in the midst of the action,**" said Sean. For fun and relaxation, Sean is an avid golfer and likes outdoor activities with his wife and dog.

## Heather Huot Creasy

### Senior Bioinformatics Analyst

Heather started at IGS in September 2007, a member of the Institute's initial bioinformatics group. In the early days, challenges involved evaluating tools and developing pipelines to be used by the new bioinformatics team in supporting IGS faculty. As faculty projects grew in complexity, she became more involved with project management support – for the **Human Microbiome Project (HMP)** and other internal faculty genomic, metagenomic and data coordination projects.

Her background is bacterial phylogenetics and comparative genomics. While working in the lab after her master's degree, Heather realized that her primary interest was delving into the informatics of analyzing data. Informatics was just growing as a field and she was fortunate enough to join the team of biocuration analysts at TIGR. This was an exciting time when curation was moving from largely manual to increasingly automated processes, which required evaluation and customization of new software tools and pipelines.

Working closely with Jacques Ravel's group on the genomics of chlamydia, there are new challenges specific to analyzing this "temperamental" bacteria that cannot be grown outside of human cells, therefore resulting in higher levels of human contaminated sequence, and less bacterial target sequence.

Heather loves her informatics work – "**Every genome or metagenome I analyze is a new puzzle to solve,**" she explains.

Outside of IGS, Heather enjoys running, tackling her first half marathon this month. She is married to Todd Creasy, another TIGR/IGS alumni, and they have two children, Julia and Ben.

## ANTIBIOTICS, MICROBIOTA AND INFECTION

Bing Ma, PhD, a Research Associate in the Ravel Lab, recently received a Developmental Research Project Award to study the effects of antibiotics, microbiota and infections.

The vaginal microbiota is the first line of defense against infections. The effects of antibiotic treatment on this natural barrier are unknown but often trigger a dysbiosis (microbial imbalance), which can put a woman at risk of further infection and adverse outcomes. One such infection is chlamydia, caused by *Chlamydia trachomatis*. Interestingly, the rate of recurrence of chlamydial genital infection after treatment is common and reaches 20 to 30%. Dr. Ma's project is testing whether antibiotic treatment for genital chlamydia leads to vaginal microbiota that increases the risks to subsequent infections. The findings gained will lead to a microbiome-based diagnostics test to identify women at higher risk for sexually transmitted infections (STIs), as well as the development of improved therapeutics and preventive strategies, such as combination therapies with probiotics that would facilitate the re-establishment of a beneficial and protective vaginal microbiota.

# SPOTLIGHT ON (GRC)

The Genomics Resource Center (GRC) continues to expand its capabilities and project portfolio. As part of our contract with the U.S. Food and Drug Administration (FDA) to sequence, assemble, and annotate pathogens in support of the development and expansion of a comprehensive, curated public reference database, we are developing a new pipeline for Ebola virus sequencing and analysis. We have also initiated several new projects to sequence large animal and plant genomes using the Pacific Biosciences platform. These larger projects were made possible by our recent upgrade to the new P6-C4 chemistry. This new chemistry, combined with improved software, has increased read lengths by more than 30% and doubled overall throughput. In June, we will host the Pacific Biosciences East Coast User Group Meeting for the third consecutive year. Please join us to hear about this exciting technology and its expanding applications.

Our Illumina platform continues to improve as well. In April, we will take delivery of our first HiSeq4000. This sequencer, the newest announced by Illumina, will increase throughput by 50% while reducing run time by an additional 50%. Each HiSeq4000 will be capable of sequencing 24 human genomes per week. We have also expanded our MiSeq repertoire with the installation of a MiSeq Dx in our CLIA facility for clinical sequencing applications.

The GRC will be hosting a **booth at the annual American Society for Microbiology** (ASM) general meeting in New Orleans from May 30–June 2, 2015. If you're there, please stop by to visit and learn more about our services and capabilities!

Check out our blog (**grcblog.igs.umaryland.edu**) for more up-to-date GRC news.

# Q&A WITH THE CO-DIRECTORS OF THE GRC

### How do I initiate a project with GRC?

It's easy! Contact us via our website (**www.igs.umaryland.edu/grc**) or email (**grc-info@som.umaryland.edu**) and we will set up an initial consultation with you. During this consultation, we will discuss your project goals and expectations and advise on experimental design. From there, we develop a project plan that includes sample requirements, timelines, cost estimates, and deliverables. For large, long-term projects, we schedule additional discussions to finalize the project plan and monitor progress.

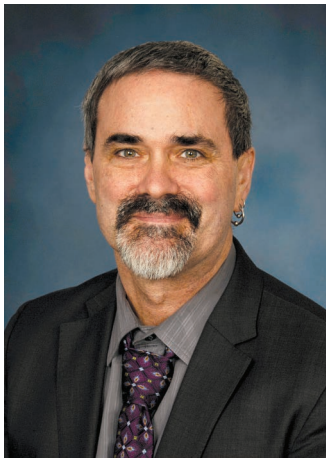
### How long does it take? How much will it cost?

These are the most common questions we hear, but often difficult to answer. Depending on the scope and scale of the project, the timeline can vary from a few weeks to months. Similarly, costs can fall in a wide range. We treat each project separately and develop the best estimates of cost and timelines as part of our consultation with each investigator.

### What about analysis after the sequencing?

IGS does it all. The GRC works closely with our Informatics Resource Center (IRC) through Dr. Owen White and Anup Mahurkar. We work seamlessly with investigators, from their initial project design through the sequencing and analysis. There are experts in IRC who specialize in genome assembly, variant analysis, metagenomics, transcriptomics, and epigenomic analysis. If you are interested in analysis, we include that as part of the project consultation and project plan.

# White Recognized with **Franklin Open Access Award**



IGS is very excited to share that Owen White, PhD, Professor, Epidemiology and Public Health, Associate Director Informatics at the Institute for Genome Sciences (IGS) at the University of Maryland School of Medicine, Co-Director, **Center for Health-related Informatics and Bioimaging (CHIB)**, has been selected to receive the Benjamin Franklin Award for Open Access in the Life Sciences from **Bioinformatics.org**, also known as The Bioinformatics Organization. The Benjamin Franklin Award for Open Access in the Life Sciences is a humanitarian/bioethics award presented annually to an individual who has, in his or her practice, promoted free and open access to the materials and methods used in the life sciences.

Dr. White has been dedicated to open source/open access during his entire scientific career. As a Principal Investigator with the **Human Microbiome Project Data Analysis and Coordination Center (known as the HMP DACC)**, he led the multi-year, multi-institutional aggregation and organization of data for the HMP and other large-scale data informatics projects. In addition to his work with the HMP DACC, Dr. White has relentlessly focused on developing effective displays of uniform and freely available information to the scientific community, which is one of the main challenges presented by large-scale genome sequencing efforts. He has developed several web resources such as the Comprehensive Microbial Resource and Pathema containing annotation from microbial genomes; **Gemina**, a web-based system designed to identity infectious pathogens; and **Sybil**, a web-based open source software package for comparative genomics. Dr. White has developed automated annotation systems for bacterial, eukaryotic and microbiome organisms, as well as several bioinformatics training programs. White will receive the award and deliver a laureate presentation at the **Bio-IT World Conference and Expo** on April 22 in Boston. White is the 14th recipient of the award, and he was selected from a pool of internationally recognized nominees.

**IGS Newsletter** is produced by the **Institute for Genome Sciences** at the **University of Maryland School of Medicine**.

| | |
|---|---|
| Jacques Ravel, PhD | Scientific Editor |
| Sarah Pick | Managing Editor |
| Riham Keryakos | Research Editor |
| Clara Daly | Graphic Designer |

# GRANTS

## Anup Mahurkar,

**Executive Director, Software Engineering & IT, Institute for Genome Sciences**, has been awarded a $250,687 grant from the National Science Foundation, for his collaborative research entitled "ABI Development: VIROME, Bioinformatics Cyberinfrastructure for the Next Wave of Scientific Advancements in Microbiome Research." This three-year award started on July 1, 2014, and is in collaboration with the University of Delaware.

## Lynn M. Schriml,

PhD, Assistant Professor, Department of Epidemiology and Public Health, University of Maryland School of Medicine and the Institute for Genome Sciences, has received a $259,844 award from the National Institutes for Health, National Institute for General Medical Sciences for collaborative Disease Ontology research entitled " Gene Wiki: Expanding the Ecosystem of Community-Intelligence Resources". The project director is Dr. Andrew Su from The Scripps Research Institute. This grant aims to expand standardized disease content and disease-gene-drug target content within the Wikipedia and Wikidata resources. This four-year award started on July 15, 2014.

## Hervé Tettelin,

PhD, Associate Professor, Department of Microbiology & Immunology, University of Maryland School of Medicine and Institute for Genome Sciences, received a $361,723 award from the National Institute of Allergy and Infectious Diseases. This multi-PI RO1 grant is led by the University of Texas Health Sciences Center, San Antonio and aims to study the formation of cardiac microlesion during invasive pneumococcal disease. This three-year award started on November 1, 2014.

## Jacques Ravel,

PhD, Professor, Microbiology & Immunology, School of Medicine, Associate Director Genomics, Institute for Genome Sciences has received a $331,067 five-year award from the National Institutes of Health (NIH) for a multi-institutional project led by the University of Louisville that aims at developing and evaluating griffithsin-based rectal microbicides for HIV. The project entitled PREvention of Vitral ENTry (PREVENT) started July 14, 2014.

Dr. Ravel has also received a five-year $79,229 award from the NIH to contribute vaginal microbiome expertise for a project entitled "Study of Women's Health Across the Nation (SWAN)", which is led by the University of California Davis and is entering its 16th year. This fourth phase of the project in part aims at understanding the changes in the vaginal microbiota in women throughout the menopausal transition and in older age, and started September 15, 2014.

Dr. Ravel received another award from the NIH for a five-year $602,811 study led by New York University to study the role of the oral microbiome in pancreatic cancer risk using genomics approaches. The project began September 19, 2014.

# Introducing IGS Post-Docs

**ELVIRA MITRAKA** is a Postdoctoral Fellow working on the Human Disease Ontology under the mentorship of Lynn Schriml, Associate Professor, Department of Epidemiology and Public Health, SOM and IGS. She got her doctorate at the University of Crete researching Vector-borne diseases. Her main research interests are Ontologies, Infectious Diseases and Data Interoperability.

**VINCENT LAM** has a research interest in cancer tumorigenesis. He received his medical degree from University of California, San Francisco and is currently a Hematology/Oncology Clinical Fellow here at University of Maryland working with Scott Devine, Associate Professor, Medicine.

**WEI SONG** is a Ph.D. researching Human Evolutionary Genetics and Disease Association under the mentorship of Timothy O'Connor, Assistant Professor, Medicine. Wei got his doctoral degree from University of North Carolina at Charlotte.

**TAYLOR RICHTER** is a Ph.D. researching the impacts of pathogens on human gut microbiome under the mentorship of David Rasko, Assistant Professor, Microbiology & Immunology. She got her doctorate from Scripps Institution of Oceanography.

# PUBLICATION LIST




Bachmann NL, Fraser TA, Bertelli C, Jelocnik M, Gillett A, Funnell O, Flanagan C, Myers GS, Timms P, Polkinghorne A: **Comparative genomics of koala, cattle and sheep strains of Chlamydia pecorum.** *BMC genomics* 2014, **15**:667. PMID 4137089.

Bachmann NL, Sullivan MJ, Jelocnik M, Myers GS, Timms P, Polkinghorne A: **Culture-independent genome sequencing of clinical samples reveals an unexpected heterogeneity of infections by Chlamydia pecorum**. *Journal of clinical microbiology* 2015. PMID.

Brotman RM, He X, Gajer P, Fadrosh D, Sharma E, Mongodin EF, Ravel J, Glover ED, Rath JM: **Association between cigarette smoking and the vaginal microbiota: a pilot study.** *BMC infectious diseases* 2014, **14**:471. PMID 4161850.

Brumbaugh AR, Smith SN, Subashchandrabose S, Himpsl SD, Hazen TH, Rasko DA, Mobley HL: **Blocking Yersiniabactin Import Attenuates Extraintestinal Pathogenic Escherichia coli in Cystitis and Pyelonephritis and Represents a Novel Target to Prevent Urinary Tract Infection.** *Infection and immunity* 2015. PMID.

Chancey ST, Agrawal S, Schroeder MR, Farley MM, Tettelin H, Stephens DS: **Composite mobile genetic elements disseminating macrolide resistance in Streptococcus pneumoniae.** *Frontiers in microbiology* 2015, **6**:26. PMID 4321634.

Chancey ST, Bai X, Kumar N, Drabek EF, Daugherty SC, Colon T, Ott S, Sengamalay N, Sadzewicz L, Tallon LJ, Fraser CM, Tettelin H, Stephens DS: **Transcriptional Attenuation Controls Macrolide Inducible Efflux and Resistance in Streptococcus pneumoniae and in Other Gram-Positive Bacteria Containing mef/mel(msr(D)) Elements.** *PloS one* 2015, **10**(2):e0116254. PMID 4335068.

Chibucos MC, Mungall CJ, Balakrishnan R, Christie KR, Huntley RP, White O, Blake JA, Lewis SE, Giglio M: **Standardized description of scientific evidence using the Evidence Ontology (ECO)**. *Database : the journal of biological databases and curation* 2014, **2014**. PMID 4105709.

Crabtree J, Agrawal S, Mahurkar A, Myers GS, Rasko DA, White O: **Circleator: flexible circular visualization of genome-associated data with BioPerl and SVG.** *Bioinformatics* 2014, 30(21):3125-3127. PMID 4201160.

Di L, Pagan PE, Packer D, Martin CL, Akther S, Ramrattan G, Mongodin EF, Fraser CM, Schutzer SE, Luft BJ, Casjens SR, Qiu WG: **BorreliaBase: a phylogeny-centered browser of Borrelia genomes.** *BMC bioinformatics* 2014, **15**:233. PMID 4094996.

Doi Y, Hazen TH, Boitano M, Tsai YC, Clark TA, Korlach J, Rasko DA: **Whole-genome assembly of Klebsiella pneumoniae coproducing NDM-1 and OXA-232 carbapenemases using single-molecule, real-time sequencing.** *Antimicrobial agents and chemotherapy* 2014, **58**(10):5947-5953. PMID 4187971.

Eppinger M, Pearson T, Koenig SS, Pearson O, Hicks N, Agrawal S, Sanjar F, Galens K, Daugherty S, Crabtree J, Hendriksen RS, Price LB, Upadhyay BP, Shakya G, Fraser CM, Ravel J, Keim PS: **Genomic epidemiology of the Haitian cholera outbreak: a single introduction followed by rapid, extensive, and continued spread characterized the onset of the epidemic**. *mBio* 2014, **5**(6):e01721. PMID 4222100.

Field D, Sterk P, Kottmann R, De Smet JW, Amaral-Zettler L, Cochrane G, Cole JR, Davies N, Dawyndt P, Garrity GM, Gilbert JA, Glockner FO, Hirschman L, Klenk HP, Knight R, Kyrpides N, Meyer F, Karsch-Mizrachi I, Morrison N, Robbins R, San Gil I, Sansone S, Schriml L, Tatusova T, Ussery D, Yilmaz P, White O, Wooley J, Caporaso G: **Genomic standards consortium projects**. *Standards in genomic sciences* 2014, **9**(3):599-601. PMID 4148985.

Fuhrman BJ, Feigelson HS, Flores R, Gail MH, Xu X, Ravel J, Goedert JJ: **Associations of the fecal microbiome with urinary estrogens and estrogen metabolites in postmenopausal women.** *The Journal of clinical endocrinology and metabolism* 2014, **99**(12):4632-4640. PMID 4255131.

Hoffmann DE, Fraser CM, Palumbo F, Ravel J, Rowthorn V, Schwartz J: **Probiotics: achieving a better regulatory fit**. *Food and drug law journal* 2014, **69**(2):237-272, ii. PMID.

Hooven TA, Randis TM, Daugherty SC, Narechania A, Planet PJ, Tettelin H, Ratner AJ: **Complete Genome Sequence of Streptococcus agalactiae CNCTC 10/84, a Hypervirulent Sequence Type 26 Strain.** *Genome announcements* 2014, **2**(6). PMID 4276828.

Huttenhower C, Knight R, Brown CT, Caporaso JG, Clemente JC, Gevers D, Franzosa EA, Kelley ST, Knights D, Ley RE, Mahurkar A, Ravel J, Scientists for Advancement of Microbiome R, White O: **Advancing the microbiome research community.** *Cell* 2014, **159**(2):227-230. PMID 4221798.

Ioannidis P, Lu Y, Kumar N, Creasy T, Daugherty S, Chibucos MC, Orvis J, Shetty A, Ott S, Flowers M, Sengamalay N, Tallon LJ, Pick L, Dunning Hotopp JC: **Rapid transcriptome sequencing of an invasive pest, the brown marmorated stink bug Halyomorpha halys.** *BMC genomics* 2014, **15**:738. PMID 4174608.

Johnson BA, Seneviratne C: **Alcohol-medical drug interactions.** *Handbook of clinical neurology* 2014, **125**:543-559. PMID.

Khmaladze E, Birdsell DN, Naumann AA, Hochhalter CB, Seymour ML, Nottingham R, Beckstrom-Sternberg SM, Beckstrom-Sternberg J, Nikolich MP, Chanturia G, Zhgenti E, Zakalashvili M, Malania L, Babuadze G, Tsertsvadze N, Abazashvili N, Kekelidze M, Tsanava S, Imnadze P, Ganz HH, Getz WM, Pearson O, Gajer P, Eppinger M, Ravel J, Wagner DM, Okinaka RT, Schupp JM, Keim P, Pearson T: **Phylogeography of Bacillus anthracis in the country of Georgia shows evidence of population structuring and is dissimilar to other regional genotypes.** *PloS one* 2014, **9**(7):e102651. PMID 4105404.

Kibbe WA, Arze C, Felix V, Mitraka E, Bolton E, Fu G, Mungall CJ, Binder JX, Malone J, Vasant D, Parkinson H, Schriml LM: **Disease Ontology 2015 update: an expanded and updated database of human diseases for linking biomedical knowledge through disease data**. *Nucleic acids research* 2015, **43**(Database issue):D1071-1078. PMID.

Kyrpides NC, Hugenholtz P, Eisen JA, Woyke T, Goker M, Parker CT, Amann R, Beck BJ, Chain PS, Chun J, Colwell RR, Danchin A, Dawyndt P, Dedeurwaerdere T, DeLong EF, Detter JC, De Vos P, Donohue TJ, Dong XZ, Ehrlich DS, Fraser C, Gibbs R, Gilbert J, Gilna P, Glockner FO, Jansson JK, Keasling JD, Knight R, Labeda D, Lapidus A, Lee JS, Li WJ, Ma J, Markowitz V, Moore ER, Morrison M, Meyer F, Nelson KE, Ohkuma M, Ouzounis CA, Pace N, Parkhill J, Qin N, Rossello-Mora R, Sikorski J, Smith D, Sogin M, Stevens R, Stingl U, Suzuki K, Taylor D, Tiedje JM, Tindall B, Wagner M, Weinstock G, Weissenbach J, White O, Wang J, Zhang L, Zhou YG, Field D, Whitman WB, Garrity GM, Klenk HP: **Genomic encyclopedia of bacteria and archaea: sequencing a myriad of type strains.** *PLoS biology* 2014, **12**(8):e1001920. PMID 4122341.

Li MD, Wang J, Niu T, Ma JZ, Seneviratne C, Ait-Daoud N, Saadvandi J, Morris R, Weiss D, Campbell J, Haning W, Mawhinney DJ, Weis D, McCann M, Stock C, Kahn R, Iturriaga E, Yu E, Elkashef A, Johnson BA: **Transcriptome profiling and pathway analysis of genes expressed differentially in participants with or without a positive response to topiramate treatment for methamphetamine addiction.** *BMC medical genomics* 2014, **7**(1):65. PMID 4279796.

Liggett SB, Bochkov YA, Pappas T, Lemanske RF, Jr., Gern JE, Sengamalay N, Zhao X, Su Q, Fraser CM, Palmenberg AC: **Genome sequences of rhinovirus C isolates from wisconsin pediatric respiratory studies.** *Genome announcements* 2014, **2**(2). PMID 3968336.

Luo Q, Qadri F, Kansal R, Rasko DA, Sheikh A, Fleckenstein JM: **Conservation and immunogenicity of novel antigens in diverse isolates of enterotoxigenic Escherichia coli.** *PLoS neglected tropical diseases* 2015, **9**(1):e0003446. PMID 4309559.

Mehta SD, Donovan B, Weber KM, Cohen M, Ravel J, Gajer P, Gilbert D, Burgad D, Spear GT: **The Vaginal Microbiota over an 8- to 10-Year Period in a Cohort of HIV-Infected and HIV-Uninfected Women.** *PloS one* 2015, **10**(2):e0116894. PMID 4326357.

Morgun A, Dzutsev A, Dong X, Greer RL, Sexton DJ, Ravel J, Schuster M, Hsiao W, Matzinger P, Shulzhenko N: **Uncovering effects of antibiotics on the host and microbiota using transkingdom gene networks**. *Gut* 2015. PMID.

Neuwald AF: **Protein domain hierarchy Gibbs sampling strategies**. *Statistical applications in genetics and molecular biology* 2014, **13**(4):497-517. PMID.

Obara I, Ulrike S, Musoke T, Spooner PR, Jabbar A, Odongo D, Kemp S, Silva JC, Bishop RP: **Molecular evolution of a central region containing B cell epitopes in the gene encoding the p67 sporozoite antigen within a field population of Theileria parva**. *Parasitology research* 2015. PMID.

O'Connor TD: AdmixKJump: **identifying population structure in recently diverged groups**. *Source code for biology and medicine* 2015, **10**(1):2. PMID 4325960.

O'Connor TD, Fu W, Project NGES, Genetics ESPP, Statistical Analysis Working Group ET, Mychaleckyj JC, Logsdon B, Auer P, Carlson CS, Leal SM, Smith JD, Rieder MJ, Bamshad MJ, Nickerson DA, Akey JM: **Rare variation facilitates inferences of fine-scale population structure in humans**. *Molecular biology and evolution* 2015, **32**(3):653-660. PMID 4327153.

Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Bieda J, Chaemsaithong P, Miranda J, Chaiworapongsa T, Ravel J: **The vaginal microbiota of pregnant women who subsequently have spontaneous preterm labor and delivery and those with a normal delivery at term**. *Microbiome* 2014, **2**:18. PMID 4066267.

Sahl JW, Morris CR, Emberger J, Fraser CM, Ochieng JB, Juma J, Fields B, Breiman RF, Gilmour M, Nataro JP, Rasko DA: **Defining the phylogenomics of Shigella species: a pathway to diagnostics.** *Journal of clinical microbiology* 2015, **53**(3):951-960. PMID.

Shiu J, Piazuelo MB, Ding H, Czinn SJ, Drakes ML, Banerjee A, Basappa N, Kobayashi KS, Fricke WF, Blanchard TG: **Gastric LTi cells promote lymphoid follicle formation but are limited by IRAK-M and do not alter microbial growth.** *Mucosal immunology* 2015. PMID.

Silva JC, Egan A, Arze C, Spouge JL, Harris DG: **A New Method for Estimating Species Age Supports the Coexistence of Malaria Parasites and Their Mammalian Hosts.** *Molecular biology and evolution* 2015. PMID.

Takala-Harrison S, Jacob CG, Arze C, Cummings MP, Silva JC, Dondorp AM, Fukuda MM, Hien TT, Mayxay M, Noedl H, Nosten F, Kyaw MP, Nhien NT, Imwong M, Bethell D, Se Y, Lon C, Tyner SD, Saunders DL, Ariey F, Mercereau-Puijalon O, Menard D, Newton PN, Khanthavong M, Hongvanthong B, Starzengruber P, Fuehrer HP, Swoboda P, Khan WA, Phyo AP, Nyunt MM, Nyunt MH, Brown TS, Adams M, Pepin CS, Bailey J, Tan JC, Ferdig MT, Clark TG, Miotto O, MacInnis B, Kwiatkowski DP, White NJ, Ringwald P, Plowe CV: **Independent Emergence of Artemisinin Resistance Mutations Among Plasmodium falciparum in Southeast Asia.** *The Journal of infectious diseases* 2015, **211**(5):670-679. PMID 4334802.

Tallon LJ, Liu X, Bennuru S, Chibucos MC, Godinez A, Ott S, Zhao X, Sadzewicz L, Fraser CM, Nutman TB, Dunning Hotopp JC: **Single molecule sequencing and genome assembly of a clinical specimen of Loa loa, the causative agent of loiasis.** *BMC genomics* 2014, **15**:788. PMID 4175631.

Vasicek EM, Berkow EL, Bruno VM, Mitchell AP, Wiederhold NP, Barker KS, Rogers PD: **Disruption of the transcriptional regulator Cas5 results in enhanced killing of Candida albicans by Fluconazole**. *Antimicrobial agents and chemotherapy* 2014, **58**(11):6807-6818. PMID 4249418.

Wang W, Feng B, Xiao J, Xia Z, Zhou X, Li P, Zhang W, Wang Y, Moller BL, Zhang P, Luo MC, Xiao G, Liu J, Yang J, Chen S, Rabinowicz PD, Chen X, Zhang HB, Ceballos H, Lou Q, Zou M, Carvalho LJ, Zeng C, Xia J, Sun S, Fu Y, Wang H, Lu C, Ruan M, Zhou S, Wu Z, Liu H, Kannangara RM, Jorgensen K, Neale RL, Bonde M, Heinz N, Zhu W, Wang S, Zhang Y, Pan K, Wen M, Ma PA, Li Z, Hu M, Liao W, Hu W, Zhang S, Pei J, Guo A, Guo J, Zhang J, Zhang Z, Ye J, Ou W, Ma Y, Liu X, Tallon LJ, Galens K, Ott S, Huang J, Xue J, An F, Yao Q, Lu X, Fregene M, Lopez-Lavalle LA, Wu J, You FM, Chen M, Hu S, Wu G, Zhong S, Ling P, Chen Y, Wang Q, Liu G, Liu B, Li K, Peng M: **Cassava genome from a wild ancestor to cultivated varieties.** *Nature communications* 2014, **5**:5110. PMID 4214410.

Yeruva L, Myers GS, Spencer N, Creasy HH, Adams NE, Maurelli AT, McChesney GR, Cleves MA, Ravel J, Bowlin A, Rank RG: **Early microRNA expression profile as a prognostic biomarker for the development of pelvic inflammatory disease in a mouse model of chlamydial genital infection**. *mBio* 2014, **5**(3):e01241-01214. PMID 4073489.