Director’s Corner

Using 3D Printer to Create Environmental DNA Research Tools

IGS – IGH Produce Cryptosporidium Gene Catalog

The Mobile Element Locator Tool (MELT) and the 1000 Genomes Project

Global City Sampling Day

IGS Launches New Core Providing Microbiome Analyses Services

New NIH BRAIN Initiative Cell Census Network Launched
Greetings Colleagues,

Recent visitors to IGS were maneuvering around boxes and rolls of bubble wrap as we packed up our entire operation for our move in early March to our new space in the School of Medicine’s Health Sciences Facility III (HSF III). We’ve been planning this move for several years and are delighted that we are now back on campus. We will be in close proximity to the Program in Personalized and Genomic Medicine (PPGM) which is also moving into HSF III, and we look forward to more face-to-face interactions with our colleagues at UMB. We will be hosting an open house in our new space this spring - look for an invitation in the not too distant future.

IGS faculty members continue to persevere, with over 60 grant submissions in FY17. Almost all of these represent collaborative efforts with colleagues here at UMB and elsewhere. In this issue, we are featuring several new research grants and publications - Matt Cannon’s work applying 3D printing to prototypes for field research (pg 3); Joana Silva’s work with IGH developing a Cryptosporidium Gene Catalog (pg 4); Scott Devine’s MELT tool developed as part of the 1000 Genomes Project (pg 5); Lynn’s Schriml’s Human Disease Ontology new NIH grant (pg 6) focusing on the classification of complex genetic diseases; Seth Ament/Owen White’s roles with the NIH BICCN, integrating informatics with brain research (pg 9).

We are also very excited to announce our new core service, the Microbiome Service Lab (MSL), directed by Mike Humphrys, which is providing microbiome analysis to UMB and worldwide investigators (pg 7). The MSL works closely with our two cutting-edge core facilities – the GRC, our sequencing core and IRC, our informatics core – for a comprehensive continuum of sequencing and analysis services. We are very pleased that the SOM is currently assessing the status of all of its Core facilities and we welcome any feedback about how we can do a better job in providing genomics expertise to interested investigators here at UMB.

Please don’t hesitate to contact any of us at IGS to discuss your needs and concerns.

Claire M. Fraser, PhD
Dean’s Endowed Professor in the School of Medicine Professor of Medicine, Microbiology and Immunology Director, Institute for Genome Sciences University of Maryland School of Medicine

Clare M. Fraser, PhD

University of Maryland School of Medicine
Institute for Genome Sciences
David Serre, Associate Professor in Microbiology and Immunology who joined IGS in 2016, is interested in applying genomic approaches to biological questions. Matthew Cannon, PhD, a research associate at IGS within Dr. Serre’s lab, is interested in the study of DNA from environmental samples (e.g. fresh or marine water, sediments and soils) to characterize biodiversity. The analysis of environmental DNA (eDNA) using next-generation sequencing is a powerful technique to identify and characterize organisms that leave traces of their DNA when they shed cells, hair, or when they decay.

One goal is to take samples deep underwater locations, where sampling methods are limited by the volume of water that can be collected and prone to contamination with surface water. Matt has wondered about ways to improve these sampling techniques to allow studies of deep oceanic environments which are traditionally very hard to explore. eDNA is particularly exciting for studying deep oceanic environments because a single sample can give a snapshot of the total biodiversity of a sampling site without direct organism sampling. Traditional methods such as collecting samples in trawl nets or expeditions using remotely operated vehicles are very expensive and can miss organisms that either cannot be captured by a net or avoid the lights of a rover. eDNA techniques could reduce the cost of characterizing the ecology of the oceans depths to a fraction of the cost of traditional methods.

To address the need for an inexpensive sampling device, Dr. Cannon designed a device to house a water filter and pump controlled by an arduino (a small computer about the size of a palm), to collect samples at any depth. This approach has the additional benefit that it allows the collection of much larger sample volume, limited only by filtering time. Dr. Cannon was able to use CAD software and 3D printing at the Health Sciences/Human Services Library Innovation Space to print the water sampling device design into a workable three-dimensional prototype.

Currently, Dr. Cannon is testing the design, ensuring that the parts work well together. Each prototype can be 3D printed in a few hours, allowing rapid development. As the design improves, the combination of the sampling device with the genomics capabilities at IGS will enable scientific research not previously possible.

FOR MORE INFORMATION

http://www.medschool.umaryland.edu/profiles/David-Serre/
**IGS – IGH Produce Cryptosporidium Gene Catalog**

*Cryptosporidium* are protozoan parasites that cause debilitating diseases in developing countries. The Institute for Global Health (IGH) at SOM has been dedicated to diagnosing, treating and eradicating diseases of global impact, including those caused by *Cryptosporidium*. IGH’s work on vaccine-preventable infectious diseases has led to many internationally-recognized successes.

Yet, parasites are more complex organisms than viruses and bacteria, and biological and technical challenges have impeded traditional vaccinology approaches to identify targets for vaccines against *Cryptosporidium hominis*, the predominant species associated with human disease.

This challenging organism has led scientists at both Institutes - IGS and IGH - to team up for an innovative research collaboration using reverse vaccinology. For reverse vaccinology, an entire pathogenic genome can be screened using bioinformatics approaches to find genes predicted to encode good vaccine targets. The process has been well-documented with bacterial pathogens, but applying reverse vaccinology to parasitic pathogens was an innovative approach. The existence of genomic resources from multiple species in the *Cryptosporidium* genus enabled a stronger reverse vaccinology approach, based on comparative genomics.

Joana Silva, PhD, Associate Professor Microbiology and Immunology at SOM and IGS, worked with Jessica Kissinger, PhD, at the University of Georgia and Giovanni Widmer, PhD, at Tufts University, to sequence, assemble and annotate the genome of *C. hominis* isolate TU502_2012. Then, working with Myron Levine, MD, DTPH, and his team at IGH, the scientists organized the data in a *C. hominis* Gene Catalog (ChGC). ChGC is a searchable online database displaying all *C. hominis* TU502_2012 predicted genes and several of their attributes that are key to identifying candidate vaccine targets.

The Gene Catalog was described in a 2016 paper in *Database*, the Journal of Biological Databases and Curation. Database URL: [http://cryptogc.igs.umaryland.edu](http://cryptogc.igs.umaryland.edu)

This research was supported in part by the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIAID, NIH) sponsored Genome Sequencing Centers for Infectious Diseases, contract number HHSN272200900009C.

**FOR MORE INFORMATION**


[http://cryptogc.igs.umaryland.edu](http://cryptogc.igs.umaryland.edu)
The Mobile Element Locator Tool (MELT) and the 1000 Genomes Project

Scott E. Devine, PhD, Associate Professor of Medicine (Endocrinology), researches mobile element insertions in genomes. He and his team are part of the 1000 Genomes Project, a large-scale international initiative that has created the largest public catalogue of human variation and genotype data. It was the first project to sequence the genome of a large number of people to provide a comprehensive resource on human genetic variation.

Dr. Devine investigates human genome variation and studies how such variation affects human health. In November 2017, a publication about the tool that he and his team developed as part of the 1000 Genomes Project was published in the journal Genome Research. This tool, which is called the Mobile Element Locator Tool (MELT), efficiently discovers and annotates new mobile element insertions (or MEIs) using whole genome sequencing data. MEIs are essentially “jumping genes” that can move from one place to another in human genomes.

MEIs represent about 25% of all structural variants in human genomes and can be disruptive influences on human traits and diseases. Dr. Devine and his team developed the MELT to facilitate MEI discovery on a population scale in humans and other organisms. Their tool allows investigators to discover MEIs in human genomes with greater speed, scalability, specificity and sensitivity compared to previous MEI discovery tools.

Scalability is increasingly becoming important as MELT is now being applied to projects involving many thousands of human genomes. For example, Devine’s group is using MELT to discover MEIs in over 100,000 human genomes as part of the Top MED project and several other large consortia projects. MELT also has been downloaded by over 400 laboratories worldwide and is being used broadly for MEI discovery in projects involving neurological disorders including autism and schizophrenia, developmental disorders, cancers, and many other diseases.

Eugene Gardner, a recent graduate from the Molecular Medicine PhD program, was the lead developer of MELT. The Graduate Program in Life Sciences (GPLS) recently recognized Eugene’s work as particularly outstanding by awarding him the 2017 GPILS PhD Thesis Award for his work with MELT. He is currently a post-doctoral fellow at the prestigious Wellcome Sanger Institute in England.

These studies have been funded by awards from the National Human Genome Research Institute (NHGRI R01 HG002898) and the National Cancer Institute (NCI R01 CA166661).

FOR MORE INFORMATION

https://www.genome.gov/27528684/1000-genomes-project/
https://genome.cshlp.org/content/27/11/1916.long
http://www.medschool.umaryland.edu/profiles/Devine-Scott/
Global City Sampling Day

The mission of GCSD is to collect samples from mass-transit systems to understand the movement of microbes across transportation systems and is part of the MetaSUB consortium project.

In June 2017, Lynn Schriml, PhD, Associate Professor, Epidemiology & Public Health, and Emmanuel Mongodin, PhD, Assistant Professor, Microbiology & Immunology, and their lab teams, led a Global City Sampling Day (GCSD) in Baltimore. Mayor Catherine Pugh participated as a citizen scientist, swabbing samples at the Charles Center Metro Station. The mission of GCSD is to collect samples from mass-transit systems to understand the movement of microbes across transportation systems and is part of the MetaSUB consortium project.

The 2017’s project builds on the MetaSUB consortium’s 2016 effort, where 54 cities across six continents and thirty-two countries collected DNA samples from subway cars (poles, seats, floors) and subway station surfaces (turnstiles, ticket machines, railings). These surfaces represent high traffic areas often touched by subway riders which provides insights to the movement of microbes around cities. The group’s previous studies have shown that microbes in subway cars represent a stable community of organisms (viruses, bacteria, and fungi) that shift slightly with the movement of riders, varies across different surfaces (plastic vs. fabric) and is most similar to those found in the soil of local parks.

“The purpose of the MetaSub project is to understand the movement of microbes across our transportation systems, and to compare the types of microbes found in coastal versus inland cities, to compare older subway systems with newer systems, and locally to determine how the Baltimore metro subway’s microbial community changes from year to year,” said Lynn Schriml.

The MetaSUB (Metagenomics and Metadesign of Subways and Urban Biomes) is supported by leading grants in public health such as the Bill and Melinda Gates Foundation, the NIH and the World Quant Foundation.

FOR MORE INFORMATION

http://metasub.org/
http://www.medschool.umaryland.edu/profiles/Schriml-Lynn/
http://www.medschool.umaryland.edu/profiles/Mongodin-Emmanuel/
IGS Launches New Core Providing Microbiome Analyses Services

In January 2018, IGS launched a new core – the Microbiome Services Laboratory (MSL) – that provides a range of wet-lab and computational microbiome analyses for investigators at UMB and world-wide. The microbiome, the assemblage of microbes living in and on our body or in the environment, has been shown to play critical role in human health and diseases. The link between these communities of microbes and our health is the focus of a growing number of health-related research initiatives some of which are ongoing at IGS/UMSOM/UMB. New insights are emerging that are rapidly demonstrating that disruption of the fine-tuned balance that exist between the microbiota and its host is associated with a wide range of diseases and conditions. These include atopic dermatitis, asthma/allergies, autism, autoimmune diseases, depression and anxiety, diabetes, hardening of the arteries, inflammatory bowel diseases, metabolic disorders, sexually transmitted infections, including HIV, preterm birth and other adverse obstetric outcomes and obesity among many others. The field of microbiome research has grown exponentially over the past 10 years, and the demand for analytical services to characterize the composition and function of the microbiome is high in academia and industry.

Over the past 10 years, Jacques Ravel, PhD, Professor, Microbiology & Immunology, Associate Director, Genomics, has become a globally recognized expert in the human microbiome, focusing on its impact on women’s health. Initially, his laboratory developed and refined microbiome analyses strategies for his own research and that of his collaborators. However, over the years, an increasing number of investigators approached him to provide similar analyses as a service. As the requests for microbiome analyses increased, particularly in the past year, Dr. Ravel and his team were able to transition the services to a separate UMB core - the Microbiome Services Laboratory (MSL). The MSL can provide a range of services to any investigator – whether it's someone at UMB or anywhere in the world. Mike Humphrys, who has led the development of the different workflows as Dr. Ravel’s laboratory manager for the past three years, is the MSL Program Director. Mike Humphrys has extensive laboratory technical and management expertise. In his role, as MSL Program Director, he will manage the MSL operations and supervise a staff of 4 research technicians.

MSL services includes DNA/RNA extractions from biological or environmental samples, 16S rRNA gene amplicon sequencing, metagenomic and metatranscriptomic analyses, and

continued on pg 8
associated bioinformatics analyses. All of the services can be ordered as complete study support or à la carte. Microbiome analyses at the MSL are integrated with the other two cores at IGS – the Genomic Resource Center (GRC) and the Informatics Resource Center (IRC) - which have been a part of IGS since the Institute launched in 2007.

“Our wet bench services are quite flexible. We can analyze DNA that is already extracted or perform the complete analysis from samples to data, and provide a complete survey of all of the different microorganisms (bacteria or fungi) present in a sample. We can also perform customized microbiome analyses depending on investigators needs. People can contact us for a consultation.” explained Mike Humphrys.

“At IGS, because of the GRC and IRC, we have comprehensive expertise. The MSL relies on the latest state-of-the art high-throughput robotic equipment required for reproducibility and high precision data generation and analyses. The team brings over 10 years of experience in this field to support customers in both laboratory and bioinformatics/statistical analyses. Further, since at IGS the cores are well integrated, we offer a breadth of services and analyses not readily found elsewhere,” said Dr. Ravel.
In Fall 2017, NIH awarded funding for a major new initiative to characterize different cell types in the mammalian brain, the largest component of the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative. The BRAIN Initiative Cell Census Network (BICCN), a network of integrated centers, collaborating laboratories, and data resources, supports 11 grants projects totaling about $50 million annually over five years. All data generated by these projects will be shared with the research community.

As part of the BRAIN Initiative, researchers at IGS will develop a data repository that is known as the Neuroscience Multi-Omic Archive or NeMO Archive, specifically engineered for the storage and dissemination of 'omics' data from the BRAIN initiative and related brain research projects. The repository will eventually provide a catalog of omics data generated from tens of millions of cells in the mammalian brain, together with information such as cells’ morphology and physiology, and connections between these cell types and corresponding changes that occur in human diseases. Several IGS faculty and staff are involved in the project, including Owen White (PI), Seth Ament, Anup Mahurkar, Michelle Giglio, and Lynn Schriml.

Researchers from leading institutions in brain research will be generating vast amounts of disparate data types. Having led the Human Microbiome Project (HMP) data coordination and analysis, and other large-scale data collection projects, Owen White and the IGS Informatics Resource Center (IRC) team have developed expertise assembling different types of data from multiple institutions and organizing it into usable formats. They will apply their expertise as part of the BICCN initiative where data collection and dissemination will be a vital component to its success.

Seth Ament is a neuroscientist at IGS and in the Department of Psychiatry at UMSOM. He has expertise in integrating diverse genomic datasets to understand brain function and diseases.

“We know a lot regarding the diversity of neurons at the level of individual cells but the challenges associated with categorizing that information properly into structural and functional relationships and understanding how each of those cells arise during development remain” said Dr. Ament. “We are interested in elucidating how particular cell types develop in healthy brains and become damaged or associated with risks for diseases of the brain.”

continued on pg 10
BICCN is structured into eight data generating projects. IGS is leading the genomics data archive and the University of Pittsburgh is leading the imaging data archive. National leaders in the field of brain research are all working together to harmonize the data formats.

“We are prioritizing particular projects,” explains Dr. Ament, “so we can understand the variety of data sets and the contributions of different cell types to psychiatric and neurodevelopmental disorders. One of our greatest challenges is the amount of data we are now able to produce. Even for experienced computational biologists, it strains our capabilities to integrate and analyze these big datasets. Owen White and his group will play a central role on this project by facilitating the analysis of these data.”

“One thing we do very well at IGS is develop more scalable computational approaches that allow us to easily look at large data sets of the size of BICCN. We do that in the context of cloud-based web-accessible computational environments, using visualization and analytical tools,” said Dr. White. “We have software engineering capabilities that other centers don’t have, that allow us developing tools to collect data and integrating multiple datasets quickly to visualize patterns across datasets.”

Ultimately, the goal with BICCN and the NeMO Archive is to build resources that make this data accessible and to maximize community utilization, to ultimately better understand brain function.

“`We have software engineering capabilities that other centers don’t have, that allow us developing tools to collect data and we can integrate multiple datasets quickly to visualize patterns across datasets.”

– Owen White, PhD

FOR MORE INFORMATION

http://www.medschool.umaryland.edu/profiles/White-Owen/
http://www.medschool.umaryland.edu/profiles/Seth-Ament/
Update from the Genomics Resource Center (GRC)

As we begin 2018, we are both reflecting on our first 10 years of operation and looking forward to another exciting year for the Genomics Resource Center (GRC) and the application of genomic technologies to human health. Since our inception as part of IGS in 2007, we have generated hundreds of trillions of basepairs of genomic sequence data for more than 50,000 samples on 15 sequencing and genomic analysis platforms. The landscape of genomic technologies continues to evolve rapidly and we remain dedicated to our mission of developing, evaluating, and implementing a suite of cutting-edge platforms to serve the needs of our investigators and collaborators.

Over the past year, demand for long-read sequencing platforms has increased dramatically. We have established the PacBio Sequel as the primary long-read platform for large genome sequencing and full-length cDNA (Iso-Seq) projects at the GRC. Successfully sequenced genomes in 2017 include multiple species of fruit trees, fish, butterflies, and numerous fungal and parasite genomes. Similarly, demand for our NanoString nCounter platform has grown significantly. With up to 800 unique probes, the platform is ideal for molecular counting of DNA, RNA and proteins. We have tested and implemented assays for miRNA, mRNA, neuropathology panels, inflammation panels, pan-cancer panels, and multiple custom designed panels.

Our newest technology is single-cell transcriptomics. We have recently acquired the 10X Genomics Chromium platform which enables 5’ and 3’ mRNA sequencing from hundreds to thousands of cells per sample. In addition, through our collaboration with Dr. Marcelo Sztein’s lab at CVD, we are testing the BD FACSMelody single-cell sorter for targeted single-cell sequencing assays.

We are excited to settle into our new laboratory in HSF III. The recently completed facility was custom designed to enable efficient and expanding genomic capabilities, including clinical sequencing in a CLIA environment. Look for an announcement about our open house coming soon!

Q&A

I’d like to try single-cell sequencing, how do I get started?

Contact us! Because single-cell sequencing is a delicate assay, we customize the project plan and coordinate logistics carefully with each investigator. An initial project consultation is crucial to a successful experiment.

Please check out our blog or email us!


