THE 23rd INTERNATIONAL
LAKE ARROWHEAD MICROBIAL GENOMICS CONFERENCE

SCIENTIFIC PROGRAM

September 11-15, 2022
UCLA Conference Center
Lake Arrowhead, California
909 / 337-2478

SCIENTIFIC PROGRAM ORGANIZERS

Dr. Jeffrey H. Miller
University of California, Los Angeles, CA

Dr. Ashlee Earl
The Broad Institute of MIT and Harvard
Cambridge, MA

Dr. Jonathan Eisen
University of California, Davis, CA

Dr. Elisabeth Raleigh
New England Biolabs, Ipswich, MA

Dr. Elinne Becket
California State University, San Marcos,
San Marcos, CA

Dr. Gautam Dantas
Washington University School of Medicine in St. Louis
St. Louis, MO

Dr. Tanja Woyke
DOE Joint Genome Institute, Berkeley, CA

Dr. Rachel Dutton
Arcadia Science,
Berkeley, CA

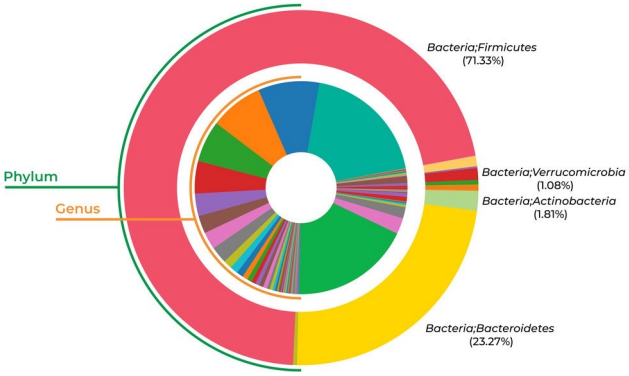
ZYMO RESEARCH
The Beauty of Science is to Make Things Simple

THE BIOCOLLECTIVE™

ZymoBIOMICS™ Fecal Reference with TruMatrix™ Technology

- **Real-Life Representation:** a microbiome quality control made from real human fecal material.
- **Consistent:** two million uniform preps available to act as a reliable, single point of reference for all microbiome researchers.
- **Well-Characterized:** deeply sequenced using targeted, metagenomic, and metatranscriptomic sequencing.

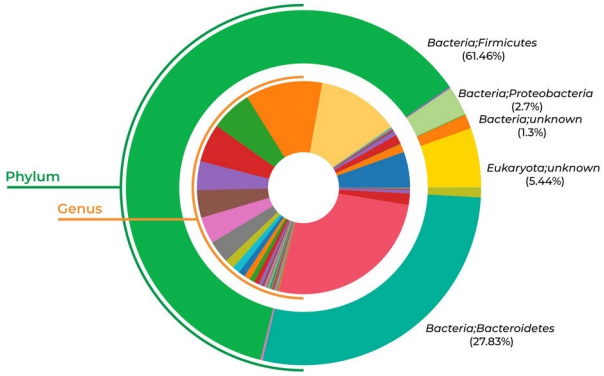
DNA Profile



High Microbial Diversity

The fecal reference contains hundreds of microbial species including bacteria fungi, and archaea, replicating the diversity of real stool.

RNA Profile



Compare Against a Validated Workflow

16S, metagenomic, and metatransciptomic raw data and analyses of the reference are available for download and data comparison.

Sign-up & Contribute to the New Database

The ZymoBIOMICS Fecal Reference Database is a public platform designed to host NGS characterization data generated specifically from the ZymoBIOMICS Fecal Reference with TruMatrix Technology. This database is a collection of raw sequencing data and analysis linked to applied methodology from various researchers of the fecal microbiome. To sign up, download, and upload visit https://www.fecalreferencedb.com/

Product	Cat. No.	Size
ZymoBIOMICS™ Fecal Reference with TruMatrix™ Technology	D6323	10 preps

info@zymoresearch.com www.zymoresearch.com Toll Free: (888) 882-9682

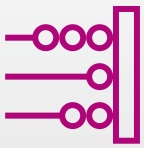
COME SEE A WHOLE NEW PACBIO

WHOLE GENOME SEQUENCING

Accurately assemble gold-standard microbial genomes

METAGENOMICS

Comprehensively characterize metagenomes – no assembly required



VIRAL POPULATIONS

Resolve complex populations, track viral evolution, and quantify minor variants

Scan to hear Melissa Smith, PhD (Univ of Louisville) share her experiences with the new HiFiViral kit for SARS-CoV-2 sequencing

PacBio

PHASE GENOMICS

ProxiMeta™ Metagenome Deconvolution Platform

Master the Microbiome

Obtain a complete picture of complex microbial communities

The ProxiMeta™ Metagenome Deconvolution Platform combines cost-effective proximity ligation data with short- or long-read shotgun sequencing data to assemble high-quality metagenomes, and accurately assign mobile genetic elements to their hosts. Venture beyond 16S-based techniques and conventional binning to discover more species, strains, genes, and gene functions.

Key features and benefits:

- Outperforms conventional binning methods to yield more complete and less contaminated metagenome-assembled fungal, bacterial, and viral genomes.
- Only technology that enables accurate host attribution accurate host attribution for phages, plasmids and anti-microbial resistance (AMR) genes.
- No microbial culturing or high molecular weight DNA extraction required.
- ProxiMeta bundles include user-friendly proximity ligation library prep and cloud-based analysis.

SCIENTIFIC PROGRAM

SUNDAY, SEPTEMBER 11

4:00-6:00 pm	Arrival and Check-in at Lake Arrowhead Conference Center
6:15-7:45 pm	Dinner (Dining Patio)
Opening of Meeting	(Pineview Room)
8:00-8:10	Jeffrey H. Miller University of California, Los Angeles **"Welcome: Introduction to Arrowhead"**
8:10-8:30	Jonathan Eisen University of California, Davis,, CA **"Arrowhead and Microbial Genomics; Update on the Tree of Life"**
8:30-9:00	Tanja Woyke Joint Genome Institue, Berkeley, CA **"Finding the Needle in the Haystack: What Targeted Sorting Approaches Can Reveal"**
9:00-9:30	Rachel Dutton Arcadia Science, Berkeley, CA **"Genetics and Genomics of the Simple Microbial Communities from Cheese"**
9:30 pm	Reception (Dining Patio)

MONDAY, SEPTEMBER 12

7:45-8:30 AM	Breakfast (Dining Patio)
Opening Session	(Pineview Room)
8:45-9:00 am	Introduction/Announcements
Session I	**Microbiomes, Microecology, Microbial Communities I**
Session Chair	Ashlee Earl, The Broad Institute of MIT and Harvard

9:00-9:30 am	Cesar de la Fuente University of Pennsylvania, Philadelphia, PA **"Artificial Intelligence Approaches for Antibiotic Discovery"**
9:30-10:00 am	Tami Lieberman Massachusetts Institute of Technology, Cambridge, MA **"De Novo Mutations in Human Microbiomes"**
10:00-10:30 am	Connie Rojas, University of California, Davis, Davis, CA **"Microbiome Variation and Function in Wild and Domestic Mammals"**
10:30-10:50 am	Break
10:50-11:20 am	Kjersti Marie Aagaard, Baylor College of Medicine Houston, TX **"Womb with a View: Importance of Maternal-Fetal Communication in Heritability and Adaptation"**
11:20-11:50 am	Hosein Mohimani, Carnegie Mellon University, Pittsburgh, PA **"Computational Methods for Natural Product Discovery from Microbial Communities by Integrating Mass Spectrometry and Metagenomic Mining"**
11:50-12:20 am	Harris Wang, Columbia University College of Physicians and Surgeons, New York, NY **"Spatial Metagenomics and Culturomics of Microbial Communities"**
12:30 pm	Lunch (Dining Patio) Free Time
4:00-6:00 pm	Poster Session I (Lakeview Room and Outdoors) Social/Mixer (Lakeview Room and Outdoors)
6:15-7:45 pm	Dinner (Dining Patio)
Session II	**Microecology, Microbial Communities II** (Pineview Room)
Session Chair	Rachel Dutton, Arcadia Science
7:45-8:15 pm	Cameron Thrash, University of Southern California, Los Angeles, CA **"Illuminating Bacterioplankton Evolution and Function Through Genomics"**

8:15-8:45 pm	Gina Chaput University of California Davis, Davis, CA **"Under the Sea: Exploring Plant-Microbe Interactions of Seagrass"**
8:45-9:05 pm	Break
9:05-9:35 pm	Jason Rothman University of California, Riverside, CA **"Longitudinal sequencing and variant detection of SARS-CoV-2 across Southern California wastewater"**
9:35	Reception (Dining Patio)

TUESDAY, SEPTEMBER 13

7:45-8:30 am	Breakfast (Dining Patio)
Session III	**Metagenomics/Pathogens/Antibiotics /Evolution**
Session Chair	Gautam Dantas, Washington University School of Medicine
9:00-9:30 am	Paul Turner Yale School of Medicine, New Haven, CT **"Predicting Evolutionary Genetics of Virus-bacteria Interactions in Patients Receiving Phage Therapy"**
9:30-10:00 am	Daria Van Tyne University of Pittsburgh, Pittsburgh, PA **"Functional Genomics and Phage Therapy for Antibiotic- Resistant Enterococcal Infections"**
10:00-10:30 am	Jeffrey E. Barrick, University of Texas at Austin, Austin, TX **"Unintentional Laboratory Evolution of Plasmids and Microbes"**
10:30-10:50 am	Break
10:50-11:20 am	Ilana Brito Cornell University, Ithaca, NY **"Mechanisms Underlying Microbiome Impacts Across Diseases"**

11:20-11:50 am	Alexandra Grote The Broad Institute of MIT and Harvard Cambridge, MA **"Using Genomics and Transcriptomics to Uncover Mechanisms of Bacterial Persistence"**
11:50-12:20 pm	Eric Alm Massachusetts Institute of Technology, Cambridge, MA **"Gene Transfer and Other Mechanisms of evolution in the Human Microbiome"**
12:30 pm	Lunch (Dining Room) Free Time
4:00-6:00 pm	Poster Session (Lakeview Room and Outdoors) Social/Mixer (Lakeview Room and Outdoors)
6:15-7:45 pm	Dinner (Dining Patio)
Session IV	**Genome Analysis**
Session Chair	Paul Orwin, University of the Pacific
7:45-8:15 pm	Allison J. Lopatkin University of Rochester, Rochester, NY **"Determinants of Plasmid Acquisition and Selection"**
8:15-8:45 pm	Katherine P. Lemon Baylor College of Medicine, Houston, TX **"In the Search for Beneficial Bacteria, Follow Your Nose!"**
8:45-9:00 pm	Break
9:00-9:20 pm	Annabelle Damerum Zymo Research Corporation, Irvine, CA **"Developing a Reproducible Microbial Lysis Methodology with Minimum Bias for Microbiome Profiling"**
9:20-9:40 pm	Ivan Liachko Phase Genomics, Seattle Washington "**Building the World's Largest Phage-Host Interaction Atlas using Proximity Ligation Technology"**
9:45 pm	Reception (Dining Patio)

WEDNESDAY, SEPTEMBER 14

7:45-8:30 am	Breakfast (Dining Patio)
Session V	**Analysis of Microbial Communities**
Session Chair	Katrine Whiteson, University of California Irvine
9:00 - 9:30 am	Gautam Dantas Washington University School of Medicine in St. Louis, St. Louis, MO **"Undertanding, Predicting, and Remediating Perturbations to Diverse Microbiomes"**
9:30-10:00 am	Rosie Alegado University of Hawaii at Manoa, Honolulu, HI **"Systems Microbial Ecology: Resistance and resilience Patterns of Subtropical Island Estuarine Microbial Assemblages in response to Seasonal Forcing"**
10:00-10:30 am	Xiaoxia (Nina) Lin, University of Michigan, Ann Arbor, MI **"Microdroplet enabled cultivation and dissection of bacterial communities"**
10:30-10:50 am	Break
10:50-11:20 am	Joanne B. Emerson University of California, Davis, CA **"Exploring the Soil Virosphere"**
11:20-11:50 am	Trent Northen Lawrence Berkeley Laboratory, Berkeley, CA **"Feeding My Friends: Examining Metabolite Exchange in the Rhizosphere"**
11:50-12:20 am	Carrie Ann Eckhart Oak Ridge National Laboratory, Oak Ridge, TN **"Development of Synthetic Biology Toolkits for Genotype to Phenotype evaluation in Non-Model Microbes"**
12:30 pm	Lunch (Dining Patio) Free Time

Session VI	**Human Genome and Microbiota Products and Elements, Synthetic Biology**
Session Chair	Elinne Becket, California State University, San Marcos
4:30-5:00 pm	Peter Weigele New England Biolabs, Ipswich, MA **"Genome Chemistry: DNA Hypermodification in Viruses of Bacteria"**
5:00-5:30 pm	Aindrila Mukhopadhyay Lawrence Berkeley Laboratories, Berkeley, CA **"Systems Biology Approaches to Scalable Bioproduction Hosts"**
5:30-6:00 pm	Ashlee Earl, The Broad Institute of MIT and Harvard, Cambridge, MA **"The Gut-bladder Axis of Recurrent UTIs"**
6:15-7:45 pm	Dinner (Dining Patio)
Session VII	**Genomics, Virus Evolution, Signaling** (Pineview Room)
Session Chair	Elinne Becket, California State University, San Marcos
8:00 -8:30 pm	Jean-Marie Volland, Joint Genome Institute, Berkeley, CA **"Crossing Bacterial Boundaries. Larger and More Complex than Ever"**
8:30-9:00 pm	Anthony Fodor University of North Carolina, Charlotte, NC **"Adventures with the Poisson Distribution: The Limitaitons and Surprising Power of Independence Assumptions in Microbiome Algorithm Design"**
9:00-9:20 pm	Jeremy E. Wilkinson Pacific Biosciences, Menlo Park, CA **"High MAG Recovery and Precision Species Profiling of a Pooled Human Fecal Reference Using PacBio HiFi Sequencing"**
9:30pm	Reception/Party (Dining Patio)

THURSDAY, SEPTEMBER 15

7:30-8:30 am	Breakfast (Dining Patio)
Session VIII	(Pineview Room)
Session Chair	Elisabeth Raleigh, New England Biolabs
9-9:20 am	Freeman Lan, University of Wisconsin, Madison, WI **"Ultrahigh-throughput single-cell genetic profiling of microbiomes"**
9:20-9:50 am	Asaf Levy, The Hebrew University of Jerusalem, Jerusalem, Israel **"Systematic Discovery and Characterization of Microbial Toxins"**
9:50-10:10 am	An-Ni Zhang, Massachusetts Institute of Technology, Cambridge, MA **"CRISPR spacer acquisition is a rare event in human gut microbiome"**
10:15-10:25 am	Closing remarks
10:25-10:50 am	Checkout
11:00 am	Departure of Bus for LAX
12:00 noon	Lunch
1:00 pm	Departure of Bus for LAX

Genome-wide search for host range mechanisms for generalist plant pathogens

Nathalie Aoun, Jonathan Beutler, Sam Holden, Yali Bai, Ariana Enriquez, Claudia Enriquez, Trung-Hieu Nguyen, Liz Jiang, Alexandra Olyushinets, Cloe Tom, and Tiffany Lowe-Power*

University of California, Davis; University of British Columbia

Background: The bacterial wilt pathogens in the Ralstonia solanacearum species complex are generalist plant pathogens that impact food security in the global tropics. Our central question is: what are the genetic factors that limit or expand phytopathogen host range? Based on global impact and well-documented epidemiological studies, we are developing the Ralstonia phylotype IIB-4 strains into a model clade to decipher host range mechanisms.

Methods: We collected more than 19 Ralstonia IIB-4 strains that were isolated from diverse locations/epidemics: Martinique Island, Colombia, Peru, Brazil, and ornamentals imported into Florida. We quantified virulence of each strain on four plant species that are natural hosts for some-but-not all IIB-4 Ralstonia: tomato cv. Moneymaker, melon cv. Sweet Granite, impatiens cv. Beacon Orange, and banana cv. Dwarf Cavendish. We used Illumina sequencing to assemble draft genomes of 12 new strains and inferred phylogeny of our 19 strains and other publicly available IIB-4 genomes. We used SCOARY to investigate whether presence/absence of any genes was associated with the level of virulence on the different host plants.

Results: Phylogenetic analysis revealed that there are at least 5 subclades of IIB-4 Ralstonia. The patterns of virulence correlated with phylogeny. SCOARY did not reveal obvious candidates.

Discussion: SCOARY works best with a larger dataset. We have sequenced more Ralstonia isolates from our collection and identified additional IIB-4 strains. Virulence assays on these new strains are ongoing. Additionally, because IIB-4 strains are reported to break genetic resistance of tomato cv. H7996, we are phenotyping all strains on this tomato line as well.

Investigating growth mechanisms required by xylem-invading bacteria under varying levels of plant defense responses

Nathalie Aoun1, Yali Bai1, Robyn Roberts2, Adam Deutscbauer3,4, Tiffany-Lowe Power1*

1 Dept Plant Pathology, UC Davis, Davis, California, USA
2 Plant Sciences, Colorado State University, Colorado, USA
3 Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA
4 Department of Plant and Microbial Biology, University of California, Berkeley, California, USA

* Corresponding author: tlowepower@ucdavis.edu

Background: Ralstonia spp. are the world's second most devastating phytobacteria, causing severe crop yield losses. Breeding for resistance to these bacteria is an important means of disease control. However, pathogen virulence evolves constantly, overcoming crop resistance. Here we investigate the genetic factors that benefit or hinder bacterial growth under varying levels of plant defense responses. We hypothesize that bacteria use common and unique genetic mechanisms to grow under varying levels of plant defense responses.

Methods: To explore mechanisms benefitting or hindering bacterial fitness inside the host, we are using a powerful functional genomics approach called random barcoded transposon mutant sequencing (RB-TnSeq). This approach can quantify the fitness of thousands of mutants during competitive growth in complex naturalistic environments. First, we screened tomato cultivars for their levels of quantitative disease response to Ralstonia sp. Second, we inoculated tomato cultivars with the RB-TnSeq library. Using Illumina sequencing, we quantified the change in the relative abundance of bacterial mutants before and after growth in eight different tomato cultivars. This quantification allows us to identify genetic factors that benefit or hinder bacterial growth in the selective environment inside the plant host.

Results: Analysis of the relative abundance of bacterial mutants in selective conditions reveals that Ralstonia sp. use common and unique genetic mechanisms under varying levels of plant defense pressure.

Discussion: Ongoing analysis of additional replicates of RB-TnSeq experiments in the Ralstonia-tomato model system will help us unravel genetic mechanisms that benefit or hinder bacterial growth under varying levels of plant defense responses. Our next step is to validate these mechanisms and understand their functions under the pressure of plant defense responses.

Global map of evolutionary dependencies between antibiotic resistance and virulence genes in E. coli

Eszter Ari 1,2,3, Enikő Kiss 2, Bálint Vásárhelyi 2,4, Gergely Fekete 2, Ágoston Hunya 2,3, Bálint Kintses 5,2,4,6, Balázs Papp 1,2

1 HCEMM-BRC Metabolic Systems Biology Research Group, Szeged, Hungary;
2 Synthetic and Systems Biology Unit, Institute of Biochemistry, Biological Research Centre, Eötvös Loránd Research Network (ELKH), Szeged, Hungary
3 Department of Genetics, ELTE Eötvös Loránd University, Budapest, Hungary;
4 National Laboratory of Biotechnology, Biological Research Centre, Eötvös Loránd Research Network (ELKH), Szeged, Hungary;
5 HCEMM-BRC Translational Microbiology Research Group, Szeged, Hungary
6 Department of Biochemistry and Molecular Biology, Faculty of Science and Informatics, University of Szeged, Szeged, Hungary.

Genes conferring antibiotic resistance or virulence phenotypes frequently undergo horizontal gene transfer in bacteria, contributing to the emergence of new multidrug resistant pathogenic variants. Mounting evidence indicates that pre-existing genome content variations influence the successful acquisition of such genes. However, the underlying evolutionary dependencies among specific genes, i.e. when one gene facilitates or hinders the acquisition of a second gene, remain poorly understood. Here we chart a high-resolution map of evolutionary dependencies between resistance and virulence genes by phylogenetic analysis of more than 20,000 Escherichia coli genomes. Our map reveals that (1) resistance genes generally facilitate each other's gain; (2) key virulence genes lack such a general pattern; and (3) contrary to some previous results, there is no overall negative dependency between the acquisitions of key virulence and resistance genes, indicating largely independent evolution between these two traits in E. coli. Strikingly, we found that the presence of efflux pump resistance genes in a genome strongly increases the chance of acquiring various other classes of resistance genes, making these efflux pumps an indicator of potentially emerging new multidrug resistant strains.

Pesticide stress induces rapid ecological changes in bacterioplankton species and gene composition without consistent evolutionary changes

Naíla Barbosa da Costa, Olga Maria Pérez-Carrascal, Marie-Pier Hébert, Vincent Fugère, Andrew Gonzalez, Gregor Fussmann, Jesse Shapiro

University of Montreal

Background: Evolutionary changes depend on shifts of heritable traits and are traditionally considered slow compared to ecological changes, related to shifts in species abundance. Nevertheless these two processes may overlap, particularly in bacteria, that exhibit short generation time and high mutation rates. Here, we studied the ecological and evolutionary responses of freshwater bacterial communities to pesticide contamination. We hypothesized that ecological responses, such as increase in abundance of resistant organisms, would be accompanied by similar evolutionary changes within populations, as ecologically successful organisms might have experienced adaptive evolution involving genome-wide selective sweeps.

Methods: We designed a 8-week experiment to study the intersection of ecology, genetic and evolutionary responses of bacteria exposed to herbicide contamination. 1,000L freshwater mesocosms were treated with different concentrations of a glyphosate-based herbicide (GBH) and water samples were filtered at 11 timepoints for DNA extraction. To test ecological hypotheses, we assessed taxonomic diversity through 16S rRNA amplicon sequencing and functional diversity by measuring the use of organic carbon sources with EcoPlates. To test evolutionary hypotheses, we performed shotgun metagenomic sequencing, reconstructed metagenome-assembled genomes, annotated them for the presence of known functional genes and inferred single nucleotide variants within genomes.

Results and Discussion: Community composition shifted primarily in response to GBH treatments without affecting carbon substrate utilization, suggesting functional redundancy for this trait. GBH also selected (directly or indirectly) for an increase in antimicrobial resistance genes encoding efflux pumps, and species abundance after a high GBH pulse were better predicted based on efflux pumps than the presence of known glyphosate resistance mutations in the target enzyme. Finally, population genomic analyses showed that intra-specific diversity varied idiosyncratically across populations that reacted similarly to GBH, showing that ecological changes occurring over short timescales are not always accompanied with directional evolutionary changes.

Impact of antibiotic regimens for the treatment and prevention of travelers' diarrhea

KS Blake 1, DJ Schwartz 1,2,3, S Paruthiyil 1, B Wang 1, SD Isidean 4,5, CK Porter 5, MP Simons 5, G Dantas 1,3,6,7

1 The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO, USA
2 Department of Pediatrics, Division of Infectious Diseases, Washington University School of Medicine, St. Louis, MO, USA
3 Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA
4 Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD, USA
5 Naval Medical Research Center, Silver Spring, MD, USA
6 Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO, USA
7 Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO, USA

Background: International travel, travelers' diarrhea (TD), and antibiotic use disrupt the healthy gut microbiome and increase risk for colonization by antibiotic resistant organisms. Two recent clinical trials evaluated the efficacy of three single-dose antibiotics to treat (TrEAT TD), or once or twice daily prophylaxis to prevent (PREVENT TD) TD on deployed military personnel. While these trials showed the regimens are comparable at treating or preventing TD, we hypothesized they may have different adverse impacts on the gut microbiome and resistome.

Methods: To evaluate this, we performed shotgun metagenomic and whole genome sequencing on 424 fecal samples and 54 diarrheagenic E. coli isolates collected from 167 military personnel traveling from the US/UK to Kenya/Honduras.

Results: We observed no significant difference in microbiome richness or ARG abundance between TrEAT groups, but note that TrEAT subjects had low gut microbiome richnesses which remained low for up to 3 weeks post-diarrhea. For the PREVENT cohort, we found the twice-daily arm was associated with decreased microbiome richness and increased ARG abundance post-travel. However, E. coli relative abundance increased 3.8-fold in the placebo group. Lastly, we observed antibiotic-resistant E. coli persisting within subjects and spreading to other as far as 10 months apart, and identify non-pathotyped E. coli as potent reservoirs ARGs.

Discussion: In this study we demonstrate the risks and benefits of antibiotics for the treatment and prevention of TD. We show that different antibiotic treatments are comparable in their impact on the microbiome, but all result in a lower microbiome richness that does not recover. We also show that twice-daily antibiotic prophylaxis to prevent TD significantly disrupts the microbiome, but placebo had increased risk for acquisition of diarrheagenic E. coli. These findings can be used to inform treatment guidelines for future travelers.

Assessing Microbial DNA in Blood: Does Fraction Matter?

Kate R. Bowie, Jared Fischer, Erin M. Dahl, Lisa Karstens

Oregon Health & Science University; Cancer Early Detection Advanced Research Center; Division of Bioinformatics and Computational Biomedicine; Portland, OR USA

Introduction: Distinct bacterial populations have been discovered in the bloodstream of prostate cancer (PC) patients, which may be indicative of disease severity. The majority of studies have relied on plasma as a proxy for blood, even though other components of blood may also have microbial DNA. We sought to determine the microbial load and composition in each fraction of blood (plasma, buffy coat, red blood cells) in non-cancer compared to PC patients.

Methods: Blood was drawn from men with high-grade PC, low-grade PC, and without cancer (n=5 per group). Samples were processed in triplicate and aliquoted into plasma, buffy coat, red blood cell pellet (RBC) and whole blood prior to freezing. Negative controls and a mock microbial dilution series were processed alongside blood samples. Microbial DNA was extracted using the QiAMP DNA mini kit. Droplet-digital PCR was used to measure 16S rRNA copies and microbial DNA composition was determined using synthetic long read 16S rRNA gene sequencing (Loop Genomics).

Results: Plasma had the lowest copies of 16S rRNA genes with a median of 21.3 copies/μL (interquartile range [IQR] of 112.5 copies/μL), with RBCs having slightly more (median 38.7, IQR 135.9 copies/μL). Buffy coat followed RBCs (median 104.72, IQR 178.9 copies/μL), and as expected, whole blood had the highest (median 134.9, IQR 153.4 copies/μL). Compositionally, Actinobacteria was the most abundant phylum followed by Firmicutes and Proteobacteria. Initial sequencing demonstrated that whole blood had the most observed species (median 10, range 9-10) and plasma had the least (median 5, range 5-6).

Discussion: The results from our study indicate plasma contains the least amount of microbial DNA and fewest number of species, indicating a potential limitation of prior work. Future directions include evaluating the compositional differences in the fractions and identifying associations with disease status.

Genome wide reconstruction of gene family evolution in microbes using xenoGI and the DTLOR model

Eliot C. Bush, Joseph S. Wirth, Ran Libeskind-Hadas

Biology Department, Harvey Mudd College; Kravis Department of Integrated Sciences, Claremont McKenna College

Background: The evolution of gene families is a key part of how microbes evolve and adapt to new niches. Duplication, transfer, loss (DTL) based reconciliation algorithms are an important method for understanding gene family evolution. Such methods take a species tree and a gene tree as input and reconcile them to reconstruct the history of events such as duplications and losses. One deficiency of existing DTL methods is that they do not account for the fact that members of a gene family may enter the species tree multiple times.

Methods: We have developed a new DTL based model called DTLOR. In order to allow multiple entries into the species tree we introduce two new events, O (origin) and R (rearrangement). We also keep track of the syntentic location of gene families. A python implementation of this model is included in the xenoGI software package and can be applied to all the gene families in a given clade of microbes.

Results: We used simulations to assess the effectiveness of xenoGI at reconstructing the evolution of gene families. Precision and recall values suggest that effectiveness varies across events. For example, the highest precision values (in the 90s) are for origin and duplication events and the lowest are for transfer and rearrangement events (in the 30s). We will also present some example cases identified in enteric bacteria.

Discussion: Analysis of simulation output and application to real data suggests that DTLOR as implemented in xenoGI is effective at reconstructing gene family evolution in clades of closely related microbes.

Genomic Analyses of Longitudinal Mycobacterium abscessus Isolates in a Multi-Center Cohort Reveal Signatures of In-Host Parallel Adaptation

JooHee Choi 1, Eric C. Keen1,2, Meghan A. Wallace 2, Madsen Zimbric 3, Carlos R. Mejia-Chew 4, Shail B. Mehta 4, Thomas C. Bailey 4, Lindsay J. Caverly 3, Carey-Ann D Burnham 2,4,5,6 Gautam Dantas 1,2,6,7

1 The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine in St. Louis, St. Louis, MO, USA
2 Department of Pathology and Immunology, Washington University School of Medicine in St. Louis, St. Louis, MO, USA
3 Department of Pediatrics, University of Michigan Medical School, Ann Arbor, MI, USA
4 Department of Medicine, Washington University School of Medicine in St. Louis, St. Louis, MO, USA
5 Department of Pediatrics, Washington University School of Medicine in St. Louis, St. Louis, MO, USA
6 Department of Molecular Microbiology, Washington University School of Medicine in St. Louis, St. Louis, MO, USA
7 Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO, USA

Background: Nontuberculous mycobacteria (NTM) are ubiquitous in the environment and are increasingly causing opportunistic infections. Mycobacterium abscessus (MAB) is one of the major NTM lung pathogens which disproportionately affect patients with cystic fibrosis (CF). MAB often causes chronic lung infections spanning years with low rates of treatment success. Understanding in-host adaptive behavior of MAB will inform strategic treatment development.

Methods: We leveraged a cohort of 175 longitudinal isolates from 30 patients with MAB lung infection in two hospital centers to identify genomic correlates of in-host adaptation. Utilizing whole genome sequencing, we characterized the relatedness of isolates both within our cohort and in the broader global context of MAB genomes. We then further investigated genes undergoing parallel adaptation in the host lung environment. Finally, we test the phenotypic consequences of parallel mutations by conducting antibiotic resistance and mercury resistance assays.

Results: We found highly related isolate pairs (<10 single nucleotide polymorphisms) across hospital centers with low likelihood of transmission. We further annotate non-random parallel mutations in 23 genes, and demonstrate altered macrolide susceptibility co-occurring with a nonsynonymous whiB1

mutation. Finally, we highlight a 23kb mercury resistance plasmid whose loss in the host confers phenotypic susceptibility to organic and non-organic mercury compounds.

Discussion: Here we highlight genomic processes through which MAB is adapting to promote its own survival within the host. Many of these events occur in parallel across patients and hospital sites. In the absence of evidence of recent transmission, we suggest highly infectious strains of MAB exhibit low rates of mutation. Further, the within-lineage polymorphisms we observed have phenotypic effects, potentially benefiting fitness in the host, at the detriment of environmental survival. Our study provides novel insight into within-host behavior of MAB infections, and highlights evidence of parallel adaptation to a pathogen lifestyle.

Use of recombinant bacteria with unique tags as spike-in controls for the quantification of microbiome content

Ray-Yuan Chuang, Britany Tang, Monique Hunter, Stefan J. Green, Briana Benton, Juan Lopera, Leka Papazisi

ATCC, Manassas, VA 20110 University of Illinois at Chicago, Chicago, IN 60607

Advanced sequencing and bioinformatics technologies have revolutionized microbiome research in remarkable ways, opening up applications in diagnostics, therapeutics, and environmental sciences. Despite the promise of these technologies, the analysis of metagenomic data remains challenging due to the technical biases introduced throughout the metagenomics workflow—from sample preparation to bioinformatic analysis. Further, the natural complexity of microbial communities themselves has challenged microbiome researchers in their ability to make meaningful, quantifiable, reproducible, and comparable measurements across different laboratories. To help promote assay standardization and validation, ATCC has developed innovative spike-in standards for microbiome research. These controls are prepared as whole cell or nucleic acid mixtures comprising three genetically engineered bacterial strains (derived from Escherichia coli, Staphylococcus aureus, and Clostridium perfringens), each containing a unique synthetic DNA tag that can be detected and quantified in routine 16S rRNA gene amplicon and shotgun sequencing assays. To demonstrate the utility of these spike-in controls in microbiome research, we conducted studies where we mixed them with whole-cell or gDNA mock communities containing different bacterial strains at various ratios. The resulting data showed that the unique tags of all three bacteria were identifiable and quantifiable by shotgun and 16S rRNA amplicon sequencing. These proof-of-concept experiments support the utility of using spike-in controls with a unique 16S rRNA tag to monitor the full process from DNA extraction to data analysis of a microbiome workflow for both 16S rRNA and shotgun metagenomics assays.

Mutation rates and adaptive variation among the clinically dominant clusters of Mycobacterium abscessus

Nicoletta Commins 1, Mark R. Sullivan 2, Kerry McGowen 2, Evan Koch 1, Eric J. Rubin 2,3, Maha Farhat 1,4

1 Department of Biomedical Informatics, Harvard Medical School, Boston, MA, 02115, USA.
2 Department of Immunology and Infectious Diseases, Harvard T. H. Chan School of Public Health, Boston, MA, 02115, USA.
3 Department of Microbiology, Harvard Medical School, Boston, MA, 02115, USA.
4 Division of Pulmonary and Critical Care Medicine, Massachusetts General Hospital, Boston, MA, 02114, USA.

Mycobacterium abscessus (Mab) is a multi-drug resistant pathogen increasingly responsible for severe pulmonary infections. Analysis of whole genome sequences (WGS) of Mab demonstrates dense genetic clustering of clinical isolates collected from¬ disparate geographic locations. This has been interpreted as supporting patient-to-patient transmission, but epidemiological studies have contradicted this interpretation. Here we present evidence for a slowing of the Mab molecular clock rate coincident with the emergence of phylogenetic clusters. We find that clustered isolates are enriched in mutations affecting DNA repair machinery and have lower spontaneous mutation rates in vitro. We propose that Mab adaptation to the host environment through variation in DNA repair genes affects the organism’s mutation rate and that this manifests as phylogenetic clustering. These results inform our understanding of niche switching for facultative pathogens, and challenge the model of transmission as the major mode of dissemination of clinically dominant Mab clusters.

Identifying genes associated with aerobic and anaerobic growth in S. enterica and non-model E. coli strains

Jaime A. Cordova Jr 1, Eric C. Neeno-Eckwall 2, Jeremy D. Glasner 2, Nicole T. Perna 1, 2

1 Laboratory of Genetics, University of Wisconsin, Madison;
2 Center for Genomic Science Innovation, University of Wisconsin, Madison

The response to changing oxygen availability is well studied in Escherichia coli K-12, but the extent that this understanding can be applied to other facultative anaerobes is unclear. Here we investigate the genomic variation across 182 isolates of Escherichia and Salmonella and use a pan-genome association study to link genetic variations to growth rates in aerobic and anaerobic conditions. This study will give us a set of genes predicted to be responsible for the oxygen response in our strains. E. coli and Salmonella strains were grown in technical triplicate in MOPS minimal media with 20% glucose for at least ten hours. Optical density measurements were done every 16 minutes with continuous shaking. Genomes were sequenced with the NovaSeq6000 system with read lengths of 2 x 150bp. Ordered contigs were then annotated using custom scripts that combined results from various gene prediction tools. Orthologous genes were identified using OrthoMCL and BLASTP and the pangenome was constructed using the Roary software. A gene presence/absence table and the growth rates were input into the Scoary2 software to associate orthogenes to the growth rates in aerobic and anaerobic conditions. Clusters of orthologous genes were used to categorize the associated genes. About 23% of the genes predicted to be associated with growth rates in both conditions were metabolism-related; additionally, ~20% were genes of unknown function. Interestingly, we identified a pattern in some subclades with notable differences in their associated gene presence and absence compared to the rest of our strains. These results give us insight into genes predicted to be responsible for the oxygen response and will also help us select a subset of strains for further analysis, including using RNA-seq. The genes of unknown function with predicted association may now also be targets for further study.

Breaking the fourth wall on meta-analysis in microbiome data

Justine Debelius1, Megan Carnes2, Yingan Chen1, Amii M Kress1, Kristen McArthur1, Noel T Mueller1, Tiange Liu1, Grier Page2, Ekaterina Smirnova3, Julie D White1, Ni Zhao4, Lisa P Jacobson1

1 Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
2 Research Triangle Institute, Research Triangle Park, North Carolina, USA 3 Department of Biostatistics, Virginia Commonwealth University, Richmond, VA, USA
4 Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Background: Establishing a common understanding of specific terms is key to successful team science and an important process for transdisciplinary fields. Microbiome analysis draws on established disciplines including microbiology, epidemiology, computer science, ecology, statistics, genetics, oceanography, geology, psychology, and chemistry among others. As a result, the interpretation of specific terms may reflect a research group's primary disciplinary background. One example of this problem is the definition of the term, "meta-analysis".

Methods and Results: Within medical and epidemiologic research, a "meta-analysis" typically accompanies a "systematic review" and refers to a synthesis exercise that relies on a systematic, structured search of the literature, team-based extraction of study summary statistics, and risk-of-bias ratings to draw a conclusion based on all available evidence. However, within microbiome research, "meta-analysis" describes combination of individual-level data from multiple sources, where "different" may be defined as multiple geographic locations by a single research group; the use of a secondary convenience dataset to provide contextualization for primary results; or a systematic search process and data extraction.

Discussion: Given the ambiguity of this term, we invite the microbiome research community to comment on how they define meta-analysis. We hope this discussion becomes a catalyst to create a consistent definition across the research field.

Fully-automated library preparation for long read WGS technologies on a novel digital microfluidics system

Alexa DeMember BS, Adam Barner BS, Eugenia Carvajal BS, Ankita Chavan MS, Shivani Kumar BS, Cheng-Chang Lee MS, Eduardo Cervantes BS, Sridhar Mandali PhD, Mais Jebrail PhD, David Greenberg PhD, Fay Christodoulou PhD

Long-read sequencing has come to play a vital role in generating contiguous, quality genomes to be used for a multitude of applications such as detection of large structural variants, deep coverage of long repeat regions, microbial community studies, and de novo assemblies. Because they do not include PCR amplification, long read libraries are able to avoid a common source of base composition bias in sequencing data. Protocols from Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT) have been growing in popularity in academic and clinical settings, though the manual preparation of these libraries requires hours of benchtop work and can be complicated to complete. Miro Canvas is a digital microfluidics platform that offers a low-throughput, small-footprint automation solution for long and complex library preparation protocols. The Canvas is a true walkaway technology, minimizing the hands-on time required for library prep to reagent preparation and instrument setup. PacBio SMRTbell Express Template Prep 3.0 and ONT Genomic DNA by Ligation are fully automated on Miro Canvas, and can construct libraries from 1-5 μg of high molecular weight DNA input. Key metrics such as read length distribution, read quality, number of reads, read length N50, and variant detection are comparable between the Miro Canvas protocols and their manual counterparts. Additionally, the Canvas protocols include bead-based size selection, reduction in reagent volumes by up to 75%, and the small footprint of the system pairs well with portable ONT sequencers. Long read technologies on Miro Canvas consistently demonstrate high-quality results with the added benefit of simplified protocol setup in a fraction of the hands-on time required for the completion of benchtop protocols.

Gut Microbial Correlates of Serious Bacterial Infections in Febrile Term Infants

Anna DeVeaux, Carla Hall-Moore, David Schnadower, Nathan Kuppermann, Prashant Mahajan, Octavio Ramilo, Phil Tarr, Gautam Dantas, Drew Schwartz

1 Washington University School of Medicine, St Louis, MO
2 Cincinnati Children's Hospital, Cincinnati, OH
3 University of California Davis School of Medicine, Davis, CA
4 University of Michigan Medical School, Ann Arbor, MI
5 Nationwide Children's Hospital, Columbus, OH

Background: Infants in the first 60 days of life (DOL) are highly susceptible to serious bacterial infections (SBI), consisting of urinary tract infections (UTIs), bacteremia, and meningitis, yet they remain challenging to diagnose and prevent. The gut microbiome and its association with SBI in term infants remains understudied. We sought to determine if there are specific gut microbial correlates of SBI.

Methods: We studied a group of febrile term infants <60 DOL who presented to 14 US emergency departments and were enrolled in the Pediatric Emergency Care Applied Research Network (PECARN) Biosignatures II study. 40 infants with SBI and 80 infants without documented SBI were chosen, and shotgun sequencing was performed of the gut metagenome and of cultured isolates from the gut and extraintestinal site (blood, urine, cerebrospinal fluid). We used Metaphlan3 to profile taxonomic composition and MaAsLin2 to perform generalized linear mixed modeling. InStrain was used for detection of the extraintestinal pathogen within the gut metagenome.

Results: No significant differences were found in gut microbiome alpha or beta diversity and enterotype classification between groups. Escherichia coli abundance in the gut was greater for infants with UTI caused by E. coli as determined by MaAsLin2 (p-value<0.001). We recovered an isolate from the gut that matched the SBI isolate in 57.5% of cases (23/40), based on species identification and multi-locus sequence typing and serotyping. A strain isogenic to the pathogen was detected in the gut of two additional cases (population average nucleotide identity >99.99%, breadth>0.5) by InStrain.

Discussion: While community signatures of SBI were not apparent, analyzing both the gut metagenome and cultured isolates allowed detection of the pathogen in the gut in more than half of infants with SBI. These findings prompt consideration of new opportunities for surveillance to prevent SBIs among colonized, pre-symptomatic infants.

Intensive care unit water sources are persistently colonized with multi-drug resistant bacteria and are the site of extensive horizontal gene transfer of antibiotic resistance genes

Luke Diorio-Toth1, Meghan A. Wallace2, Christopher W. Farnsworth2, Bin Wang1,2, Danish Gul3, Jennie H. Kwon4, Saadia Andleeb2, Carey-Ann D. Burnham2,4,5,6, Gautam Dantas1,2,5,7

1 The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO, USA
2 Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA
3 Atta ur Rahman School of Applied Biosciences, National University of Sciences and Technology Islamabad, Pakistan
4 Department of Medicine, Washington University School of Medicine in St Louis, St Louis, MO, USA
5 Department of Molecular Microbiology, Washington University School of Medicine in St Louis, St Louis, MO, USA
6 Department of Pediatrics, Washington University School of Medicine in St Louis, St Louis, MO, USA
7 Department of Biomedical Engineering, Washington University School of Medicine in St Louis, St Louis, MO, USA

Background: Contamination of hospital sinks with microbial pathogens presents a serious threat to patients, but our understanding of colonization dynamics is largely based on outbreaks. In a previous survey of intensive care unit (ICU) surfaces, we observed extensive contamination by both common nosocomial pathogens and opportunistic, water-associated pathogens. This led us to hypothesize that water sources were acting as a reservoir for these environmental multidrug-resistant organisms (MDROs). Here, we investigate the colonization patterns of MDROs in ICU sinks and water from two hospitals in the United States and Pakistan collected over 27 months. We additionally collected samples from geographically-matched rooms in homes and shared office spaces.

Methods: We used selective culture and isolation to recover 822 bacterial isolates, which were subjected to Illumina sequencing and antimicrobial susceptibility testing. A subset of 60 were additionally sequenced using Oxford Nanopore long-read technology.

Results: ICU water sources in both countries had a high burden of Pseudomonas, Acinetobacter, and Stenotrophomonas spp. Genomic analyses revealed long-term colonization by Pseudomonas spp. and Serratia marcescens strains across multiple rooms for more than two years. Isolates recovered from ICU rooms were enriched in antibiotic resistance gene (ARG) abundance and diversity compared to home and work rooms. Nanopore sequencing uncovered

extensive cross-species horizontal gene transfer (HGT), predominately within Enterobacterales. Plasmids that harbor diverse and clinically-important ARGs are shared among multiple Enterobacterales and Acinetobacter species over a period of 19 months, and these plasmids confer phenotypic resistance to multiple beta-lactams, including carbapenems.

Discussion: In this environment, antibiotic resistance in Pseudomonas spp. is maintained by strain colonization, while HGT maintains AR elements within A. junii, A. johnsonii, and Enterobacterales independent of colonization. These results imply that even transient MDRO appearance in these reservoirs poses a risk to patients, and emphasizes the importance of proactive surveillance to prevent infections.

Longitudinal analysis of the microbiome and metabolome in the 5xfAD mouse model of Alzheimer's disease

Sage J. B. Dunham, Katelyn A. McNair, Eric D. Adams, Julio Avelar-Barragan, Stefania Forner, Mark Mapstone, Katrine L. Whiteson

University of California Irvine Department of Molecular Biology & Biochemistry; University of California Irvine Department of Computational Science; Institute for Memory Impairments and Neurological Disorders (UCI MIND); University of California Irvine Department of Neurology

Microbial exposures impact the onset and progression of Alzheimer’s disease (AD) through the mediation of inflammation, the exchange of small-molecules across the blood-brain barrier, and stimulation of the vagus nerve. A baseline understanding of AD model animal microbiomes is therefore vital for improving the efficacy of our animal models and improving AD research quality. Here we describe our evaluation of the microbiome and metabolome in longitudinal fecal, cecal, and plasma samples from the 5xfAD transgenic mouse model. We performed DNA extraction and shotgun Illumina sequencing on cecal and fecal samples from 5xfAD and wild-type B6J (WT) animals from 4–18 months of age. We also performed metabolomics on plasma and feces from a subset of the same animals. We observed significant sex, age, and cage-specific differences in the microbiome. Bacteria with significantly higher abundances in 5xfAD mice include multiple Alistipes spp., two Ligilactobacillus spp., and Lactobacillus sp. P38. Those with lower abundances included multiple Turicibacter spp., Lactobacillus johnsonii, and Romboutsia ilealis. In contrast to previous findings of depleted serotonin in persons with AD, plasma measurements revealed elevated serotonin in older 5xfAD animals relative to their WT littermates, although the differences were small in comparison to the large decrease in serotonin associated with aging across models. 5xfAD animals also exhibited significantly lower plasma concentrations of carnosine and the lysophospholipid lysoPC a C18:1. Correlations between the fecal microbiome, fecal metabolome, and plasma metabolome were also explored. Taken together, these findings strengthen the link between Turicibacter abundance and AD and provide a basis for further microbiome studies of murine models for AD.

To infinity and beyond: exploring viral diversity in extreme environments

Cassandra L. Ettinger, Morgan Saunders, Neptali Flores, Sudipto Paul, Claudia Coleine, Nicole Pietrasiak, Jason E. Stajich

University of California, Riverside; California State Polytechnic University, Pomona; San Jose State University; University of Tuscia, Italy; New Mexico State University; University of California, Riverside

Background: Despite their rampant abundance on the planet, relatively little is known about viruses in extreme environments, and whether viruses facilitate adaptation of their hosts to harsh conditions. In an effort to catalog viruses across extremes (e.g., temperature, salinity, aridity), we explored viral diversity in three datasets: (i) 50 metagenomes from terrestrial Cyanobacteria co-cultures isolated from hot desert environments, (ii) 191 metagenomes from Antarctic rocks from ice-free areas until recently thought incapable of supporting life, and (iii) 42 metagenomes and metatranscriptomes associated with seagrasses, the only existing marine plants.

Methods: We identified viral sequences in assembled metagenomes/metatranscriptomes using VirSorter1/VIBRANT or VirSorter2, assessed their quality using CheckV, grouped them into viral OTUs (vOTUs) at >95% similarity and then clustered vOTUs with RefSeq genomes using VContact2. We mapped reads to vOTUs to obtain relative abundance and predicted bacterial hosts using NCBI BLAST. Exploratory analysis was performed by undergraduate interns using KBase and R.

Results: We identified 814 viral sequences associated with the terrestrial Cyanobacterial co-cultures, over 100,000 sequences associated with endolithic Antarctic rock communities, and 8047 sequences associated with seagrass leaves, roots and associated sediment. However, of the viral sequences identified, only 72, 73 and 54 sequences, respectively, clustered with viral RefSeq genomes. Across all three datasets, the majority of viral sequences clustering with RefSeq genomes belonged to the order Caudovirales. Further within datasets, few vOTUs were shared across samples, and communities were largely dominated by taxonomically unclassified viral clusters, likely representing novel viral genera.

Discussion: Overall, these results are consistent with viruses in extreme environments being underrepresented in reference datasets. This foundational work expands knowledge of viral diversity in these habitats and provides catalogs of viral sequences for future work, such as investigating the role of viral auxiliary metabolic genes in adaptation of hosts to extreme environments.

Gain some, lose some: bacterial gene content shifts associated with growth temperature adaptation

Anne Farrell, Olga Zhaxybayeva

Dartmouth College

Life thrives across a wide span of temperatures, yet every organism is restricted to growing within a narrow optimum range. While some genomic features that contribute to adaptation to optimum growth temperatures (OGT) are known, specific sets of genes contributing to temperature preferences are unknown and the evolutionary processes leading to changes in OGT are poorly understood. The bacterial phylum Thermotogota is an excellent system for studying the evolution of OGT. It comprises mesophilic, thermophilic, and hyperthermophilic members that collectively grow between 20°C-90°C. Using genome-wide association studies, we identified a set of gene families associated with OGT in members of the Thermotogota phylum. Mapping family gains and losses on the species phylogeny revealed that families present in thermophiles with OGT ≥65°C were also present in the phylum's last common ancestor, whereas families found in species with OGT <65°C were gained later. Additionally, the gene families found in hyperthermophiles were lost on branches leading to Thermotogota with OGT <65°C. Additionally, the latter taxa frequently acquired genes horizontally, based on evidence from phylogenetic analyses. We hypothesize that adaptation to lower temperature necessitates both the acquisition of genes absent in thermophiles and loss of unnecessary genes. Our findings are in concordance with several analyses that suggest a non-hyperthermophilic nature of early microbes and a later evolution of both hyperthermophily and mesophily. Some of these gene families are physically and functionally associated with each other, with most of these gene networks contributing to metabolic functions. However, many of these genes encode proteins of unknown function, and their experimental characterization will advance our understanding how OGT diversity evolved in Thermotogota and throughout the tree of life.

The evolution of the binary toxin locus in Clostridioides difficile

Skye RS Fishbein1,2, Newcomer, Erin P1,2,3, Hink, Tiffany4, Reske, Kimberly A4, Cass, Candice4, Henderson, Jeffrey P4,6, Burnham, Carey-Ann D2,5,6, Dubberke, Erik R4, Dantas, Gautam1,2,3,6

1 The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St Louis, United States.
2 Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, United States.
3 Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, United States.
4 Division of Infectious Diseases, Washington University School of Medicine, St. Louis, United States.
5 Department of Medicine, Washington University School of Medicine in St Louis, St Louis, MO USA.
6 Department of Molecular Microbiology, Washington University School of Medicine, St Louis, United States.

Background: Clostridioides difficile (Cd) is a pervasive nosocomial pathogen and significant cause of healthcare-related diarrhea. Cd infection (CDI) is often provoked by microbiome perturbations (ie. antibiotics or chemotherapy), but Cd strain identity contributes to differences in colonization outcome. Cd genomes are highly mosaic, and the ramifications of this genetic diversity in context of disease severity remains poorly understood. Our previous data indicate that the presence of the binary toxin locus, an accessory locus purported to increase virulence, may be associated with CDI.

Methods: Diagnostic data from patients colonized with toxigenic Cd with a history of least one diarrheal stool was analyzed from three prospective/retrospective Cd studies at Barnes Jewish Hospital in St. Louis. To capture distinct presentation of disease, I examined 178 patient-isolate pairs from patients who had at least one diarrheal episode and whose stools were either enzyme immunoassay positive (EIA+) for toxin (n=71) or EIA negative (EIA-, n=107). Whole-genome sequencing and isolate genomics were used to examine correlates of disease severity.

Results: EIA+ patients were significantly more likely to be colonized with Cd strains containing cdtAB relative to EIA- patients. Using previously collected isolates and representative isolates from NCBI, phylogenetic analysis revealed that cdtAB is generally confined to ST11 and ST1 lineages of Cd, with diverse lineages containing cdtAB pseudogenes. Comprehensive examination of cdtA synteny indicates that cdtAB may be mobilized between isolates with a number of transcriptional regulators.

Discussion: The association of EIA status with the presence of the binary toxin locus confirms the previous literature that CdtAB may exacerbate Cd disease. Surprisingly, the genetic context of the cdtAB locus is not conserved across all isolates, indicating a mobilizable capacity of the locus. Given our previous understanding of toxin regulation in Cd, future work will emphasize the regulatory mechanisms underlying CdtAB protein expression.

Investigating wound microbiome composition in Type 2 Diabetic mice

Jaeden D. Flury1, Patricia A. Tulloch1, Brad T. Hunter1, Renee Ha1, Julie Jameson1*, and Elinne Becket1*

1 California State University San Marcos, San Marcos, CA
* Corresponding Authors

Chronic, nonhealing wounds affect many patients with Type 2 diabetes. Microorganisms are believed to play a significant role in the process of wound healing, as more diverse microbiomes induce faster healing. Various commensal and pathogenic bacteria including Staphylococcus, Corynebacterium, and Pseudomonas tend to colonize wounds once they are established on the host. However, how the microbiome influences chronic wounding is not fully understood. To better understand the role of microbes in diabetic wound healing, we performed 16S rRNA sequencing to compare the wound microbiomes of diabetic and nondiabetic mice. Each mouse was anesthetized and wounded with a double 2 mm biopsy punch, and those wounds were swabbed pre-wounding and days 2 and 7 post-wounding. The swabs then underwent DNA extraction, V3-V4 16S rRNA library preparation, and Illumina-based sequencing. ANOVA analyses on the change in wound diameter between the diabetic and nondiabetic mice groups revealed that the diabetic mice wounds healed significantly slower ($p=0.021$) than the control group. NMDS analyses on microbial community compositions revealed distinct community profile shifts between sampling days in nondiabetic mice wound ($p=0.01$) that were not observed in diabetic mice ($p=0.22$). Additionally, the taxonomic charts revealed unique profiles between the control and diabetic mice pre- and post-wounding. Collectively, those data suggest a correlation between wound healing and microbial composition.

Mapper: Fast and accurate sequence alignment via x-mers

Jeffry M Gaston1*, Eric J Alm2, An-Ni Zhang2*

1 Software Company, Cambridge, USA
2 Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, USA;
* Authors contributed equally to this study. &Corresponding author:
Address: Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, USA.

Read mapping quality determines the efficiency and accuracy of interpreting metagenomic data. Most sequence aligners base their search algorithms on splitting sequences into fixed-size pieces (k-mers). However, a known challenge of k-mer algorithms, the choice of kmer size, strongly affects the quality of sequence alignment. Here, we introduce Mapper, a new sequence aligner that addresses this challenge by dynamically incorporating k-mers of various sizes, which we refer to as x-mers. Mapper is 10-15 times faster than previous aligners, while producing fewer false positives and fewer false negatives. Mapper was tested on a diverse collection of microbial genomes and these advantages are consistent across species. In principle, any k-mer-based algorithm, such as similarity search, genome assembly, metagenomic annotation, and multiple sequence alignment, might benefit from incorporating x-mers.

Characterizing the impact of restoration on the Rana sierrae skin microbiome across restoration histories and sites

Sonia L. Ghose 1, Cassandra L. Ettinger 2, Jacqueline Colin 3, Vance T. Vredenburg 4, and Jonathan A. Eisen 1

1 Department of Evolution & Ecology, UC Davis, CA USA
2 Department of Microbiology & Plant Pathology, UC Riverside, CA USA
3 Applied Molecular Transport, San Francisco, CA USA
4 Department of Biology, San Francisco State University, CA USA

Background: The Sierra Nevada yellow-legged frog, Rana sierrae (Rs), has been driven close to extinction in part by the fungal pathogen Batrachochytrium dendrobatidis (Bd). Skin-associated microbes can inhibit Bd and microbial community structure can impact disease outcomes in hosts, indicating that the microbiome should be considered during restoration efforts.

Methods: Here, we leveraged on-going restoration efforts in the Desolation Wilderness to investigate how the Rs microbiome is impacted by history of restoration and restoration site. We sequenced bacterial 16S rRNA amplicons and performed qPCR for Bd from Rs skin swabs across one source site (naturally persisting frogs), one captive site (Zoo population derived from source), and two restoration sites (including frogs translocated from the source, reintroduced from captivity, and new recruits born at restoration sites).

Results: We tested hypotheses about the impact of restoration on the microbiome by comparing Shannon diversity, community structure (Weighted Unifrac; WU), taxonomic composition, and Bd prevalence and intensity at field sites. We found that reintroduced, translocated, and new recruit frog microbiomes did not differ significantly in any comparisons. While the two restoration sites did not differ from each other in most comparisons, WU was significantly different. In addition, captive frogs were different from all other restoration histories and sites in terms of Shannon diversity and WU. Source frogs also differed from one or more restored populations or restoration sites in all comparisons. Finally, all frog microbiomes were dominated by a single amplicon sequence variant, SV1 (Family Burkholderiaceae).

Discussion: Restoration site appears to be more important than history in determining the microbiome of reintroduced and translocated frogs, and both captive and source frogs were distinct from other groups. Moving forward, we will explore what factors at different sites are leading to microbiome differences. These findings may guide future conservation strategies that take Bd-frog-microbiome interactions into account.

Effect of population size on context-dependent trade-offs of antibiotic resistance

Indra Gonzalez Ojeda[1,2]; Cristina Herren[1,3]; Michael Baym[1]

1 Department of Biomedical Informatics and Laboratory of Systems Pharmacology, Harvard Medical School;
2 Biophysics Program, Harvard University;
3 Department of Marine and Environmental Sciences, Northeastern University

Evolutionary adaptations are generally associated with fitness costs, yet many studies on the costs of antibiotic resistance have failed to identify growth rate deficits in resistant bacteria. Recent work has shown that costs of resistance can manifest as a reduced tolerance of novel environmental conditions. Here, we consider the effects of population size on these context-dependent fitness costs. The relationship between population size and the environmental costs caused by resistance is unclear; it may be that a larger population size means that mutations with lower fitness costs are more likely to arise, or it may be that a larger population size allows for a higher degree of adaptation to specific environments. In this study, we characterize the effect of population size on the growth deficits incurred by chloramphenicol-resistant Escherichia coli in novel thermal conditions. We show that when resistance develops in a larger population, bacteria show greater growth deficits when compared to bacteria evolved in smaller populations. This result is consistent with the hypothesis that the observed difference is due to larger populations becoming more tightly adapted to the experimental conditions, causing a greater decrease in their thermal niche breadth; nonetheless, further studies are needed to confirm this hypothesis and integrate this observation into an evolutionary framework. The results of this study give insight into the interplay between two key ecological factors that could influence antibiotic resistance development. More broadly, further characterizing the fitness costs associated with resistance development can help predict scenarios where resistance will arise.

In vivo ligand-responsive gene expression in probiotic S. boulardii

Daniel J Haller*, Deniz Durmusoglu*, Ibrahim S Al'Abri*, Nathan Crook

Chemical and Biomolecular Engineerinering, North Carolina State University

Background: Saccharomyces boulardii (Sb) is a leading probiotic candidate for delivery of biotherapeutics to the mammalian gut. While Sb is a promising probiotic, precise control of gene expression in the species has not been achieved; such control can play a crucial role in maintaining the fitness and survival of the strain during colonization. Developing inducible gene expression systems that can be tuned via the addition of ligands could play an essential role for production and delivery of biotherapeutics.

Methods: Using CRISPR-Cas genome editing, we created an Sb strain that can utilize galactose and raffinose. Using various cloning techniques, we developed 5 ligand-responsive gene expression systems for Sb and characterized their titration curves with aerobic (yeGFP and mKate) and anaerobic (CaFbFP) reporters and measured the expression levels under aerobic and anaerobic conditions using different concentrations of the inducers.

Results: Galactose utilizing Sb allowed tunable gene expression (227-fold induction) when galactose is the sole carbon source with galactose promoter (pGAL1). Other four promoters that are activated in the presence of aTc (pTET), IPTG (pLAC), xylose (pXYL) or copper (pCUP1), achieved 16-, 88-, 10- and 3-fold induction, respectively in the presence of glucose. Under anaerobic conditions, pXYL, pLAC and p pGAL1 exhibited activation profiles similar to the aerobic conditions. Orthogonality assays confirmed that each promoter exhibited ligand-specific activation when tested with ligands other than its own.

Discussion: A challenge in the development of engineered probiotics is control over dosage. Ligand-responsive gene expression systems only activate transcription in the presence of a particular inducer and this regulation by the concentration of inducer enables custom dosage of a therapeutic. This work shows development of ligand responsive expression units for Sb, expanding the applicability of Sb in the gut microbiome.

Fast Average nucleotide identity (ANI) estimates are as good for delimiting Klebsiella species as ANI

Julie E Hernández-Salmerón, Gabriel Moreno-Hagelsieb

Wilfrid Laurier University

Klebsiella strains are among the most important multi-drug resistant human pathogens. The importance and ubiquity of these organisms call for quick and accurate methods for their classification. Genomic average Nucleotide Identity (ANI) is becoming a standard for bacterial species delimitation. Since complete ANI calculations can be too slow, faster estimates, based on sequence sampling, have been appearing in the literature. In this study we compared 1,189 Klebsiella genomes using measures calculated with Mash, Dashing, and DNA compositional signatures, all of which run in a fraction of the time required to obtain ANI. Receiver Operating Characteristic (ROC) curve analyses showed equal quality in species discrimination for ANI, Mash and Dashing, with Area Under the Curve (AUC) values above 0.99, followed by DNA signatures (AUC: 0.96). Accordingly, groups obtained at optimized cutoffs largely agree with species designation, with ANI, Mash and Dashing producing 15 species-like groups. Testing Mash to map species after adding draft Klebsiella genomes to the dataset, also showed excellent results (AUC above 0.99), suggesting that the whole 13,574 genome dataset could be divided into 26 species-like groups.

ENGINEERING AN INCUBATION ENVIRONMENT THAT MIMICS IN SITU CONDITIONS FOR IN VITRO COASTAL MICROBIOME STUDIES.

Brad T. Hunter 1, Jaeden D. Flury 1, Sebastian S. Cocioba 2, Matthew L. Cope-Arguello 1, Jared M. Helms 1, Kevin H. García 1, Gerardo Dominguez 1, Darcy A. A. Taniguchi 1*, and Elinne Becket 1*

1 Department of Biological Sciences, California State Univ., San Marcos, San Marcos, CA
2 Binomica Labs, Long Island City, NY 4301 22nd St fl 3rd studio 342, Long Island City, NY 11101, *Corresponding authors

Background: Coastal environments are dynamic and can vary widely on short- or long-term scales depending on location and weather. Incubation equipment that reflects these changes through programmable gradient light and temperature cycles would permit more precise in vitro coastal microbiome studies. Here, we present an open-source incubation environment that mimics in situ conditions for in vitro coastal microbiome studies using a modified shaking water bath that has fully customizable temperature and light gradients that can also mimic real-time field conditions.

Methods: To test how this build emulates real environment experimental conditions, we performed a 48-hour experiment comparing changes to coastal microbiomes at the Ellen Browning Scripps Memorial Pier in La Jolla, CA versus incubation in our build in the laboratory while mimicking live conditions at Scripps Pier. This was followed by metagenomic sequencing across conditions and time points to observe changes to microbial community profiles.

Results: NMDS plots from shotgun sequencing of microbiome samples taken at 0h, 24h, and 48h demonstrated the samples form distinct clusters based on time yet did not form separate clusters between conditions. Analyses of congruence revealed that clustering at each time point between conditions (in situ vs. in vitro) were not statistically different (p-value: 0.2 and 1 between conditions at 24h and 48h respectively). However, there was a significant difference in communities between time points sampled (p-value: 0.01).

Discussion: We developed an open-access build that can be pre-programmed to permit precise in-lab microbiome experiments in a controlled and accessible environment. Further experiments over extended time frames and comparing conventional laboratory testing methods are needed to further explore the efficacy of the programmed conditions, but these data serve as an initial proof of principle for a set of modifications that can be added to common laboratory equipment to bolster complex long-term microbial ecology studies.

Exploring microbial physiology in chronic human infection through metatranscriptomics

Carolyn B. Ibberson 1, 2 and Marvin Whiteley 1

1 School of Biological Sciences and Center for Microbial Dynamics and Infection, Georgia Institute of Technology; Emory-Children's Cystic Fibrosis Center, Atlanta, GA 30310
2 Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK 73019

Chronic bacterial infections including cystic fibrosis, chronic wounds, and osteomyelitis affect a significant portion of the population in developed countries and place a large burden on healthcare systems. These infections can last many weeks to years, despite numerous clinical interventions and are frequently recalcitrant to antibiotics. Chronic infections almost always harbor complex polymicrobial communities of commensal and environmental organisms and polymicrobial infections often result in higher bacterial burdens, increased tolerance to antibiotics, and more severe disease compared to single-species infections, a process termed "synergy". Although synergy between microbes has long been recognized, the underlying molecular mechanisms and their impacts on chronic wounds have remained difficult to elucidate. To begin to address this knowledge gap, we assessed 115 metatranscriptomes from human cystic fibrosis sputum and chronic wound infections, isolated from 3 continents and 5 clinics. We first determined the community composition using MetaPhlAn3 and found sputum is more diverse than chronic wounds with a mean of 11.7 and 6.6 species identified, respectively. Further, we found Gram-positive organisms dominated both communities, comprising 67.4% (+/-3.1 SEM) of the species identified in these communities. Further, we found these environments are likely hypoxic as strict anaerobes were found to comprise a mean of 21.7% (+/-2.6 SEM) of the communities in these chronic infection sites. Ongoing analyses are identifying functions critical community functions in each infection site using HUMAnN3, MetaPro, and SAMSA2. Collectively, this data will improve our understanding of microbe-microbe interactions and community functions in chronic infection.

Do you see what I see? Improving color accessibility and organization of microbiome data with the microshades R package

Lisa Karstens, Kate R. Bowie, Erin M. Dahl, Emory Neer, Eric T. Leung

Oregon Health & Science University; Division of Bioinformatics and Computational Biomedicine; Cancer Early Detection Advanced Research Center; Portland, OR USA

Background: Color Vision Deficiency (CVD), commonly known as colorblindness, affects 1 in 12 men and 1 in 200 women - approximately 300 million people worldwide. Individuals with CVD do not experience complete loss of color vision, but have reduced ability to distinguish different colors. When creating scientific figures, it is important to consider that individuals with CVD may not perceive all colors as intended. While there are several CVD friendly color palettes available, they are often insufficient for visualizing microbiome data.

Methods: We developed an R package, microshades, to overcome CVD accessibility for microbiome datasets. Microshades includes CVD accessible color palettes and data organization functions. To construct the palettes, hue (type of color), chroma (colorfulness), and luminance (brightness) were adjusted for optimal visual distinction and CVD accessibility. All shades were tested with a CVD simulator (cvdemulator) for accessibility. Data organization functions include grouping data by taxonomic ranking, sorting the data vertically and horizontally, and restructuring the plot legends.

Results: Each microshades color palette contains six hues with five sequential variations of chroma and luminance per hue, for a total of 30 available colors per palette. The microshades_cvd_palettes colors are universally CVD accessible to individuals with the three most common types of CVD (Deuteranope, Protanope, and Tritanope). The individual hues of the microshades_palettes colors are CVD friendly, but when used in conjunction with multiple hues, may not be universally accessible to all forms of CVD.

Discussion: The microshades R package is a visualization tool for microbiome researchers. The package contains two CVD accessible palettes, along with several organization features. The microshades package can be used in conjunction with common microbiome R packages, such as phyloseq, to enhance microbiome data visualization.

Ultrahigh-throughput single-cell genetic profiling of microbiomes

Freeman Lan1, Jason Saba1,2, Katie Krauska1, Tyler Ross1, Robert Landick1,2, Ophelia Venturelli1

Department of Biochemistry, University of Wisconsin, Madison1 Department of Bacteriology, University of Wisconsin, Madison2

Background: Single cell genetic heterogeneity in microbiomes underlies many important phenomena, such as evolution, antimicrobial tolerance and resistance, and the dynamics of mobile genetic elements. However, despite the importance of understanding the single cell heterogeneity in microbiomes, tools for single cell genetic profiling of microbes are not yet widely available.

Methods: We have developed a robust, generalizable, and widely accessible method for ultrahigh-throughput single-cell genetic analysis of microbes. Using microfluidic devices, microbial cells are encapsulated into pico-liter droplets, lysed, then targeted regions of the genomes are amplified through a multiplex PCR reaction that simultaneously amplifies regions of interest and connects them with cell-specific barcodes. Libraries are then run on an Illumina sequencer producing uniquely barcoded reads corresponding to amplicons of each cell. This method is designed to be simple, robust, and accessible to most academic laboratories with minimal microfluidics expertise.

Results: To show the broad applicability of this method, we demonstrate three vastly different applications of single cell sequencing on human gut microbes. 1. We capture the single-cell genomic variation within the capsular polysaccharide operons of a single colony of Bacteroides fragilis; 2. we profile the distribution of antibiotic resistance genes among a 25-member gut microbial community; and 3. we monitor the transmission dynamics of 3 naturally occurring self-transmissible plasmids within a community of Enterobacteriaceae.

Discussion: Single cell genetic profiling is a powerful way to study microbiomes, however the field currently lacks robust and widely applicable tools for doing so. This generalizable and accessible tool for single-cell genetic profiling of microbiomes opens up unprecedented opportunities for research in this field.

Impact of polymicrobial interactions on single-cell transcriptomic heterogeneity

Gina Lewin, Anayancy Ramos, Marvin Whiteley

Center for Global Health and Diseases and Department of Pathology, Case Western Reserve University, Cleveland, OH; Center for Microbial Dynamics and Infection and School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA; Emory-Children's Cystic Fibrosis Center, Atlanta, GA

Most microbes reside within complex communities, where microbe-microbe interactions influence their physiology and fitness. In these spatially structured communities, it is hypothesized that subpopulations of cells with altered transcriptional profiles have outsized roles. However, measurements of microbial interactions are almost exclusively performed at the population level. Here, we measure the impact of spatial structure and microbial interactions on transcriptomic heterogeneity using single-cell RNA-sequencing (scRNA-seq). Specifically, we identified the transcriptome of single cells and small aggregates of the oral pathogen, Aggregatibacter actinomycetemcomitans (Aa) grown in biofilms in mono-culture and in co-culture with the cross-feeding and cross-respiring oral commensal Streptococcus gordonii (Sg). Our dataset contains thousands of cells, with an average of ~100 unique protein-coding reads per cell or aggregate. Analysis of these samples is revealing how co-culture in biofilms impacts the extent of transcriptional heterogeneity and the identity of variably expressed genes. These data will be fundamental to our understanding of subpopulations in microbial communities.

Parallel changes in fitness effects and gene essentiality over 50,000 generations of evolution

Anurag Limdi, Sian V. Owen, Cristina Herren, Richard E. Lenski, Michael Baym

Anurag Limdi: Harvard Medical School; Sian V. Owen: Harvard Medical School; Cristina Herren: Harvard Medical School; Richard E. Lenski: Michigan State University; Michael Baym: Harvard Medical School

Background: Over evolutionary time, bacteria face changing environments, which may require different sets of genes for survival. As they adapt to a specific constant environment, some genes are modified and lost, which can increase fitness while also modulating the effects of further gene losses. However, whether evolutionary specialization leads to systematic changes in robustness to gene loss is largely unexplored.

Methods: Here, we compare the effects of insertion mutations in Escherichia coli between ancestral and 12 independently derived strains after 50,000 generations of growth in a simple, uniform environment using transposon insertion mutagenesis and sequencing (TnSeq)

Results: We find that epistasis between insertion mutations and genetic background is common, but the overall distribution of fitness effects is largely unchanged. In particular, we see systematic changes in gene essentiality, with more genes becoming essential over evolution than vice versa. The resulting changes often occurred in parallel across the independently evolving populations. A few of the changes in gene essentiality are associated with large structural variations, but most are not.

Discussion: Taken together, our results demonstrate that gene essentiality is a dynamic, evolvable property, and they suggest that changes in gene essentiality are a result of natural selection in this long-term evolution experiment, rather than a mere byproduct of structural changes.

Evaluation of addiction module function of chromosomal Toxin-Antitoxin gene pairs from Variovorax paradoxus EPS

Annie Lin, Trisha Dinh, Paul Orwin

University of the Pacific

Toxin-antitoxin (TA) systems are pairs of co-expressed genes found in many genomic contexts. They were originally identified as “addiction modules” in natural plasmids, but have subsequently been implicated in many microbial processes. This addiction is often based on the differential stability of the gene products, leading to post-segregational killing (PSK) in cells that lose the plasmid. We used the TASmania genome annotation for Variovorax paradoxus EPS to identify eight clear TA system loci in the genome. Gibson assembly was used to clone these loci into the broad host range vector pBBR8-GFPuv, replacing the fluorescent protein with the TA system structural genes under arabinose control. The pemIK genes from the plasmid pBBR5pemIKpBAD were used as a positive control for addiction phenotype. The plasmid sequences were verified using the whole plasmid sequencing service from Plasmidsaurus. Passage experiments with and without antibiotic selection were performed using arabinose to control expression of the TA system loci. After passaging, the cultures were plated on selective and non-selective media to determine the ratio of cells continuing to harbor the plasmid. Our hypothesis was that plasmid curing would be evident with no selection or induction, and that an active TA system would “addict” the cells to the plasmid, preventing loss. After 10d of passage in the stabilization assay we saw evidence for plasmid stabilization in the control construct as well as 2/8 Variovorax gene cassettes. We also saw evidence of toxicity in 3/8 constructs, indicating that the antitoxin efficacy as well as toxin function may be limited by host background. Samples from each culture were preserved for further analysis using Nanopore sequencing. Analysis of the variation in the plasmid population will shed further light on the selective process. This experimental scheme will provide a systematic strategy for identifying TA system genes in newly sequenced genomes.

Genetic diversity of commensal *Blastocystis* gut protists reveals strain-specific changes in host-interfacing pathways

Abigail L. Lind[1], Ami S. Bhatt[2], Katherine S. Pollard[1]

1 Gladstone Institute of Data Science and Biotechnology , San Francisco, CA
2 Department of Genetics, Stanford University, Stanford, CA

The human gut microbiome is a microbial ecosystem containing bacteria, archaea, viruses, and microbial eukaryotes. The most common human gut microbial eukaryote is the commensal protist *Blastocystis*, with an estimated prevalence in industrialized countries at 25%. While the presence of *Blastocystis* is associated with a significant reduction in pro-inflammatory bacteria and with reduced gut inflammation, little is understood about its role in the gut microbiome and its fundamental biology. Genetic evidence suggests that *Blastocystis* comprise a group of genetically diverse subtypes, but we lack high quality genomic data for these subtypes and understanding of their functional similarities and differences. Here, we cultivate 6 *Blastocystis* strains spanning the genetic diversity of the genus and generate contiguous, annotated genomes using a combination of long-read DNA sequencing, Hi-C, and RNA-seq. These genomes range in size from 14-25 Mb and have protein-coding genes with unusual features, including a frequent lack of canonical stop codons and a regular intron length of exactly 30 base pairs. Through comparison with the genomes of closely related stramenopiles, we find a pattern of genome reduction and gene duplication in *Blastocystis*, as well as genomic organization patterns that likely arose during the transitions from a free-living lifestyle to an obligate within-host lifestyle and in transitions between host species. We find substantial strain and subtype-specific gene duplications, including those of likely host-interfacing genes such as those involved in cell-cell adhesion and cell surface glycan production. Together, these genomes and our analyses reveal the adaptations *Blastocystis* has undergone to thrive in the gut microbiome. These results identify substantial biological variability between subtypes of *Blastocystis* which are likely to drive differences in interactions with other gut microbiota and the host.

Epidemiology of Plasmid Lineages Mediating the Spread of Extended-Spectrum Beta-Lactamases among Clinical Escherichia coli

Bejan Mahmud, Meghan A. Wallace, Kimberly A. Reske, Kelly Alvarado, Carol E. Muenks, David A. Rasmussen, Carey-Ann D. Burnham, Cristina Lanzas, Erik R. Dubberke, Gautam Dantas

Washington University School of Medicine, St. Louis, MO, USA; North Carolina State University, Raleigh, NC, USA

The prevalence of extended-spectrum beta-lactamases (ESBLs) among clinical isolates of Escherichia coli has been increasing, with this spread driven by ESBL-encoding plasmids. However, the epidemiology of ESBL-disseminating plasmids remains understudied, obscuring the roles of individual plasmid lineages in ESBL spread. To address this, we performed an in-depth genomic investigation of 149 clinical ESBL-like E. coli isolates from a tertiary care hospital. We obtained high-quality assemblies for 446 plasmids, revealing an extensive map of plasmid sharing that crosses time, space, and bacterial sequence type boundaries. We provide further support for plasmid-mediated spread of ESBLs but demonstrate that some ESBL genes rely on dissemination through plasmids more than the others. Through a sequence-based network, we identified specific plasmid lineages that are responsible for the dissemination of major ESBLs. Notably, we demonstrate that IncF plasmids separate into two distinct lineages that are enriched for different ESBLs and occupy distinct host ranges. Our work provides a detailed picture of plasmid-mediated spread of ESBLs, demonstrating the extensive sequence diversity within identified lineages, while highlighting the genetic elements that underlie the persistence of these plasmids within the clinical E. coli population.

Insights to Pathogenic Potential: Genomic Analysis of Escherichia coli from Asymptomatic and Symptomatic Patients

Grace Morales1,2, Seth Reasoner1,2 ,Benjamin Abelson2, Maria Hadjifrangiskou1,2,3, Jonathan Schmitz1,2,

1 Department of Pathology, Microbiology, and Immunology, Vanderbilt University, Nashville, Tennessee, USA
2 Department of Urology, Vanderbilt University Medical Center, Nashville, Tennessee, USA
3 Vanderbilt Institute for Infection, Immunology, and Inflammation, Vanderbilt University, Nashville, Tennessee, USA

Background: Urinary tract infections (UTI) afflict over 50% of women over the course of their lifetime. While uropathogenic Escherichia coli (UPEC) causes >75% of reported UTIs, UPEC lacks a specific molecular signature. In the clinical setting this poses a significant dilemma when patients present with asymptomatic bacteriuria (ASB), in which significant numbers of E. coli is found in the urine of patients, but without the associated symptoms of UTI. Both ASB and UTI-causing E. coli carry similar types of virulence factors, and – to date – no genomic signature exists to tell these two E. coli types apart.

Methods: In this study, we sequenced approximately 800 E. coli strains from the urine of symptomatic and asymptomatic patients. We utilized a custom bioinformatic pipeline to characterize strains outside of the traditional virulence factors, hypothesizing that a signature lies not in the carriage of distinct virulence genes, but rather in discrete polymorphisms that change tropism and pathogenic potential. Included in this analysis are applications of microbial Genome-Wide Association studies as well as machine learning methodology to detect potential loci associated with the pathotype in an unbiased manner.

Results: We show urine-isolated strains of E. coli maintain an open pangenome. These strains additionally carry a variety of virulence associated factors with no correlation of the development of symptoms. We queried mobile genetic elements, and found that plasmids, prophage, and other transposable elements also do not correlate with symptoms. We continue explore other more subtle genomic features that might constitute a defining feature of UPEC.

Discussion: Previous work has shown that strains causing UTI remain phenomenological, and there is no molecular signature to differentiate these strains from other extraintestinal E. coli strains. This work will more robustly define true uropathogenic strains with the aim of identifying a molecular signature to differentiate strains with pathogenic potential.

Taxonomic and Antibiotic Resistance Changes to Coastal Microbiomes in Response to Rainstorm Runoff

Elizabeth Murguia1, Christian Lopezguerra1, Mackenzie V. Pylant1, Ebony Stretch1, Dennis Zanesco1, Matthew Cope-Arguello1, Uriel Rivera1, Anahita Rahimi1, Nicholas Allsing2, Scott T. Kelley2, Elinne Becket1

1 California State University San Marcos
2 San Diego State University

Antibiotic resistance (AR) is a global healthcare issue driven by the overuse of antibiotics in clinical, agriculture, and aquaculture applications. Urban and agricultural runoff introduce antibiotic-resistant bacteria and antibiotic contamination to recipient environments. Antibiotics change microbial community compositions in favor of resistant species and can trigger the exchange of DNA carrying antibiotic resistance within a given community. Mapping the changes in microbial community composition and AR gene abundance in response to rainstorm runoff has yet to be elucidated. We therefore analyzed the taxonomic and AR gene abundance changes to coastal microbiomes throughout seasonal rainstorms. Sampling at the Batiquitos lagoon outlet in Carlsbad, California occurred over 14 days; before, during, and after the first two rainstorms of the 2019-2020 season. Coastal water was captured on-site on a 0.22 µm mixed cellulose ester (MCE) membrane filter. We performed total DNA isolation and shotgun library preparation on the isolated microbiomes followed by 2 x 150 base paired-end sequencing. Microbial composition and AR gene identification was performed on the resulting metagenomes, to determine a time course profile of relative microbial abundance and antibiotic resistance profiles. Additionally, we performed meta-SourceTracker analysis on the time course metagenomes to investigate proportions of exogenous and endogenous microbial community members throughout and following rainstorms, as well as to explore possible sources of exogenous taxa. We observed an overall bimodal increase in alpha diversity and AR gene counts in the 24-72-hour period following each rainstorm. Taxonomic changes are reflected by a relative depletion of Cyanobacteria and relative increase in Proteobacteria and Bacteroidetes. Increases in Proteobacteria appear to be predominantly marine eutrophication-associated microbes while Bacteroidetes increases were predominantly freshwater- and soil-associated microbes commonly implicated in fish and human disease. The microbial community profile returned to a pre-storm composition after approximately six days contrary to the three-day recovery time commonly referenced.

Longitudinal transmission patterns of diverse Clostridioides difficile lineages in a hospital: lessons and concerns.

Newcomer, Erin P1,2,3, Fishbein, Skye RS1,2, Hink, Tiffany4, Reske, Kimberly A4, Cass, Candice5, Dubberke, Erik R4, Dantas, Gautam1,2,3,6.

1 The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St Louis, United States.
2 Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, United States.
3 Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, United States.
4 Division of Infectious Diseases, Washington University School of Medicine, St. Louis, United States.
5 Department of Medicine, Washington University School of Medicine in St Louis, St Louis, MO USA.
6 Department of Molecular Microbiology, Washington University School of Medicine, St Louis, United States.

Background: Current methods for preventing Clostridioides difficile (Cd) infection (CDI) in healthcare environments include antibiotic stewardship programs and placing patients with CDI on contact precautions. However, a minority (19-30%) of CDI cases are verifiably transmitted from previous CDI cases, while asymptomatic Cd carriers (ACCs) outnumber patients with CDI 10-fold and can shed pathogenic Cd spores. These data suggest that ACCs could be an important vector in Cd transmission, yet little is known about the environmental contamination and resulting risk of CDI.

Methods: A prospective study, spanning 6 months in the hematopoietic cell transplant and leukemia wards at Barnes-Jewish Hospital in St. Louis, was conducted to investigate Cd transmission dynamics. Patients were screened for Cd carriage on admission and weekly, and environmental swabs from bedrails, sinks, and keyboards were collected at each timepoint, in addition to remnant stool specimens from patients tested for Cd during routine care. Multiplex PCR for toxin genes and whole-genome sequencing were conducted on cultured Cd isolates.

Results: This study collected 659 stool specimens and 495 rectal swabs from 384 patients and 400 environmental specimens, identifying a total of 22 CDI (EIA+) cases during the study period. In-silico MLST identified significant surface contamination by strains not carried by the patient occupying the room. Whole-genome SNP analysis revealed 5 lineages of Cd that spread between patients and involved ACCs carrying toxigenic strains.

Discussion: The involvement of ACCs in transmission networks associated with CDI indicates that ACCs may play a significant role in Cd transmission. Networks were found across diverse strains not typically considered hypervirulent, suggesting that strains that can asymptomatically colonize and cause CDI have a survival advantage in wards where only CDI patients are put on contact precautions. Future work will investigate genetic determinants of transmission associated with networks spanning multiple patients.

Metabolic profile of the prevalent human gut bacterium Eggerthella lenta

Cecilia Noecker1, Juan Sanchez2, Jordan E. Bisanz1, Veronica Escalante1, Margaret Alexander1, Kai Trepka1, Almut Heinken3, Yuanyuan Liu4, Dylan Dodd 4, Ines Thiele3, Brian DeFelice 2, and Peter J. Turnbaugh1,2*

1 Department of Microbiology & Immunology, University of California, San Francisco, San Francisco, CA 94143
2 Chan-Zuckerberg Biohub, San Francisco, CA 94158
3 School of Medicine, National University of Ireland, Galway, Ireland
4 Department of Pathology, Stanford University, Stanford, CA 94305
* Present address: Department of Biochemistry and Molecular Biology, Pennsylvania State University, State College, PA

Background: Human gut microbes perform diverse metabolic functions with consequences for host health. Yet the metabolic strategies relied upon for growth by many commensal gut microbes remain unclear. The prevalent and disease-linked Actinobacterium Eggerthella lenta does not metabolize sugars and instead displays several unusual metabolic traits: arginine dependence, a putative acetogenesis pathway, and numerous enzymes involved in xenobiotic metabolism and anaerobic respiration.

Methods: To obtain a comprehensive view of the metabolic network of E. lenta, we generated several complementary resources: novel chemically defined culture media, metabolomics profiles of strain isolates in vitro and in vivo, and a curated genome-scale metabolic reconstruction.

Results: We analyzed E. lenta growth and metabolism in defined media, finding that it uses arginine and acetate as energy and carbon sources in tandem and is auxotrophic for tryptophan and riboflavin. Untargeted metabolomics of culture supernatants revealed diverse metabolic byproducts, many of which were strain-variable and could be linked to accessory genes. E. lenta growth was strongly promoted by acetate, which we found to be incorporated into cell wall and nucleotide biosynthesis using stable isotope analysis. These characteristics were largely recapitulated in silico by flux balance analysis of an updated metabolic model of the E. lenta type strain. Finally, we compared the metabolic footprint of E. lenta in vitro and in gnotobiotic mice, finding consistent metabolite signatures across environments. This comparison led us to identify and characterize the host signaling metabolite agmatine as an alternative energy source for E. lenta that can replace the effect of arginine.

Discussion: These results elucidate the distinctive metabolic niche filled by E. lenta in the gut ecosystem and suggest potential novel impacts of this organism on human hosts and other community members. Our approach is a generalizable strategy to gain mechanistic and predictive insights into the metabolism of understudied microbes.

Association of diet and antimicrobial resistance in healthy US adults

Andrew Oliver1, Zhengyao Xue1,2, Yirui Tang1,2, Blythe Durbin-Johnson3, Zeynep Alkan1, Diana H. Taft2, Jinxin Liu4, Ian Korf 3, Kevin D. Laugero1,5, Charles B. Stephensen1,5, David A. Mills2,6, Mary E. Kable1,5, Danielle G. Lemay1,3,5

1 USDA-ARS Western Human Nutrition Research Center, 430 West Health Sciences Dr., Davis, CA, 95616
2 Department of Food Science & Technology, University of California, One Shields Ave., Davis, CA, 95616
3 Genome Center, University of California, One Shields Ave., Davis, CA, 95616 4 Laboratory of Gastrointestinal Microbiology, College of Animal Science and Technology, Nanjing Agricultural University, Nanjing 210095, China
5 Department of Nutrition, University of California, Davis, One Shields Ave. Davis, CA, 95616
6 Department of Viticulture and Enology, Robert Mondavi Institute for Wine and Food Science, University of California, Davis, One Shields Ave., Davis, CA, 95616

Background: Antimicrobial resistance (AMR) represents a significant source of morbidity and mortality worldwide. Since resistance to antibiotics is encoded in the microbiome, interventions aimed at altering the taxonomic composition of the gut might allow us to prophylactically engineer microbiomes that harbor less antibiotic resistant genes (ARGs). Diet is one method of intervention, yet little is known about the association between diet and antimicrobial resistance. Our primary hypothesis was that increased intake of dietary fiber would be associated with reduced ARG abundance in human fecal metagenomes. Beyond our directed hypotheses, we utilized machine learning approaches on a variety of diet, physiological, and lifestyle features to assess whether the abundance of antibiotic genes is correlated with variables outside the scope of our directed hypotheses.

Methods: We examined diet using the food frequency questionnaire (FFQ, habitual diet) and 24-hour dietary recalls (ASA24) coupled with analysis of the microbiome using shotgun metagenome sequencing in 290 healthy adult participants of the USDA Nutritional Phenotyping Study. Additionally, we applied machine learning to examine 387 dietary, physiological, and lifestyle features for associations with antimicrobial resistance.

Results: We found that aminoglycosides were the most abundant and prevalent mechanism of AMR in these healthy adults and that aminoglycoside-o-phosphotransferases (aph3-dprime) negatively correlated with total calories and soluble fiber intake. Individuals in the lowest quartile of ARGs (Low-ARG)

consumed significantly more fiber in their diets compared to Medium- and High-ARG individuals, which was concomitant with increased abundances of obligate anaerobes in their gut microbiota.

Conclusions: Our results indicate individuals with lower abundances of antibiotic resistance genes consumed more diverse diets that were richer in fiber and limited in animal protein. We suspect that increased fiber likely drives the composition of the gut towards a more obligate anaerobe state, reducing footholds for facultative anaerobes, which are known harbors of inflammation and antibiotic resistance.

Ultra-deep Sequencing of Hadza Hunter-Gatherers Recovers Vanishing Gut Microbes

Matthew R. Olm, Bryan D. Merrill, Matthew M. Carter, Dylan Dahan, Surya Tripathi, Sean P. Spencer, Brian Yu, Sunit Jain, Norma Neff, Aashish R. Jha, Erica D. Sonnenburg, Justin L. Sonnenburg

Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA, USA; Chan Zuckerberg Biohub, San Francisco, CA, USA; Genetic Heritage Group, Program in Biology, New York University Abu Dhabi, Abu Dhabi, UAE; Center for Human Microbiome Studies, Stanford University School of Medicine, Stanford, CA, USA

The gut microbiome is a key modulator of immune and metabolic health. Human microbiome data is biased towards industrialized populations, providing limited understanding of the distinct and diverse non-industrialized microbiomes. Here, we performed ultra-deep metagenomic sequencing and strain cultivation on 351 fecal samples from the Hadza, hunter-gatherers in Tanzania, and comparative populations in Nepal and California. We recover 94,971 total genomes of bacteria, archaea, bacteriophages, and eukaryotes, 43% of which are absent from existing unified datasets. Analysis of in situ growth rates, genetic pN/pS signatures, high-resolution strain tracking, and 124 gut-resident species vanishing in industrialized populations reveals differentiating dynamics of the Hadza gut microbiome. Industrialized gut microbes are enriched in genes associated with oxidative stress, possibly a result of microbiome adaptation to inflammatory processes. This unparalleled view of the Hadza gut microbiome provides a valuable resource that expands our understanding of microbes capable of colonizing the human gut and clarifies the extensive perturbation brought on by the industrialized lifestyle.

Whole Genome sequencing and antibiotic resistance of isolates from the Alligator mississippiensis gut microbiome

Julissa Perez-Marron, Jenine Dong, Erica Meng, Paul Orwin

University of the Pacific

Multidrug resistant bacteria are a global threat to public health due in part to the lack of development of new antibiotics. One major avenue for the development of antibiotic resistance and the spread of resistant bacteria to humans is from agriculture. Although most studies of this process have focused on traditional farm animals (pigs, cows, chickens), the American alligator, Alligator mississippiensis is also raised on farms for food and leather. The gut microbiome of this animal is less well understood than its counterparts in agriculture. We cultured bacteria from the alligator colon, and identified 16 antibiotic resistant isolates using disk diffusion. Using 16S rRNA gene sequencing the 16 isolates were shown to fall into seven unique strains were identified among the selected isolates: Aeromonas sp., Citrobacter freundii, Citrobacter portuncales, Escheriera coli, Providencia revertii, Providencia sp., Pseudomonas aeruginosa and Pseudomonas sp. Through minimal inhibitory assays, it was determined that all strains were resistant to two or more of the seven antibiotics tested after 24 hours of exposure. All strains excluding Pseudomonas sp. exhibited resistance to both Ampicillin and Rifampin. Two of our isolates Pseudomonas sp. and Escherichia coli were resistant to six of the seven antibiotics tested. Genome assemblies were performed in Unicycler using Nanopore long reads and Illumina short read data with a hybrid approach. ABricate was used to identify the antibiotic resistance genes in the assembled genomes and to correlate the genes present with the observed phenotypes. The data presented here suggest that farmed alligators may be a significant and important reservoir of antibiotic resistant human pathogens.

Gut microbiome in endometriosis

Inmaculada Pérez-Prieto 1,2; Eva Vargas 3; Eduardo Salas-Espejo 1; Analuce Canha- Gouveia 1,4; Andres Salumets 5,6; Kreete Lüll 7,8; Elin Org 7; Signe Altmäe 1,2,9

1 Department of Biochemistry and Molecular Biology, Faculty of Sciences, University of Granada, Granada, Spain.
2 Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, Spain.
3 Systems Biology Unit, Department of Experimental Biology, Faculty of Experimental Sciences, University of Jaen, Jaen, Spain.
4 Department of Physiology, Faculty of Veterinary, University of Murcia, Campus Mare Nostrum, IMIB-Arrixaca, Murcia, Spain.
5 Department of Obstetrics and Gynaecology, Institute of Clinical Medicine, University of Tartu, Tartu, Estonia.
6 Competence Centre on Health Technologies, Tartu, Estonia. 7 Institute of Genomics, Estonian Genome Centre, University of Tartu, Tartu, Estonia.
8 Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia.
9 Division of Obstetrics and Gynecology, CLINTEC, Karolinska Institutet, Stockholm, Sweden.

Background: Endometriosis, defined as the presence of endometrial-like tissue outside of the uterine cavity, is one of the most common reproductive disorders affecting ~10% of women in reproductive age. There are different theories of the possible causes of endometriosis, however, its pathogenesis is not clear. Novel studies are indicating that the microbial composition may be involved in the etiology of endometriosis, nevertheless, the connection between microbes, its dysbiosis and the development of endometriosis remains unexplored. This study aims to analyze and compare the gut microbiome profile in women with and without endometriosis and determine microbial pathways potentially involved in the development of the disease.

Methods: This case-control study includes a subsample of 1000 women of the Estonian Biobank, a volunteer-based cohort of the Estonian adult population. 136 women with endometriosis and 864 control women who have not been diagnosed with endometriosis or any of its comorbidities were included. To characterize the microbiome composition and functionality, microbial DNA was extracted and sequenced by paired-end metagenomic shotgun sequencing on the Illumina Novaseq 6000 platform. Microbial functional pathways were annotated using the KEGG database (https://www.genome.jp/kegg/).

Expected results: The data processing is ongoing. The alpha- and beta-diversity and differential abundance analyses will be performed to assess the gut microbiome (species and KEGG orthologies [KO]) in both groups. The ANOVAlike Differential Expression tool (ALDEx2, v.1.18.0) will be used to identify species and KOs associated with endometriosis. Analysis will be adjusted for age, BMI, and stool consistency.

Discussion. To the best of our knowledge, this will be the biggest metagenome study in women with endometriosis. Our findings would provide better understanding of the endometriosis-associated microbiome in a large-scale cohort, which may lead to discovery of novel links between the microbiome and the etiology of endometriosis.

Building the World's Largest Phage-Host Interaction Atlas using Proximity Ligation Technology

Phase Genomics, Inc

Viruses, including bacteriophage and archaeal viruses, are the most abundant form of life on earth (10 31). They interact with all life and shape the global ecosystem through their impacts on community composition and horizontal gene transfer. However, phage-host relationships have proven challenging to identify without use of culture-based experiments to generate unambiguous evidence for a phage's presence in a given host. These experiments inherently require that all hosts are culturable, typically restricting the scope and microbial diversity that can be surveyed and limiting our understanding of potentially valuable phage-host relationships.

Proximity ligation sequencing is a powerful genomic method for associating viruses with their hosts directly in native microbial communities. Proximity ligation captures, in vivo, physical interactions between the host microbial genome and the genetic material of both lytic and lysogenic phage. Similar to culturing experiments, these linkages offer direct evidence that phage sequences were present within an intact host cell, thereby establishing a phage-host pair. However, unlike culturing experiments, proximity-ligation methods do not require the propagation of living bacterial cells and unlike single cell sequencing experiments, only capture phage-host interactions inside cells. The combination of intra-phage and phage-host signal enables us to simultaneously deconvolve viral genome bins (vMAGs) directly from metagenomes and to assign microbial hosts to large numbers of vMAGs without culturing.

Our application of this technology to hundreds of complex microbiome samples has yielded thousands of novel phage
and archaeal virus genomes with host assignments, as well as large numbers of new microbial genomes. Through
broad-scale application of proximity ligation sequencing, we are creating a global-scale database of highly diverse
phage-host interactions from samples from across the world. We will present published and unpublished work
highlighting the power of this approach in the field of metagenomic discovery.

“Microbes@Home”, a module to teach microbial diversity and bioinformatics to undergraduates.

Giorgia Pirino and Katherine L. Petrie

School of Biological Sciences, University of California, San Diego, LA Jolla, CA

The revised curriculum of the laboratory in microbiology at UCSD integrates a research component into the course, fostering critical thinking and a better understanding of the scientific process.
At the beginning of the pandemic, the course moved to a remote format, focusing more on data analysis. To keep a hands-on component and “student ownership”of experiments, we implemented the Microbes@Home module, in which students collected samples from surfaces in their household and shipped them to the lab, where 16S rDNA PCR amplicons were sequenced through the Illumina platform. Students then analyzed the microbial communities using the QIIME 2 pipeline via AWS, which provided a consistent computing environment for all students, regardless of their personal home computer.
Across 4 quarters, 221 samples were analyzed and over 7000 unique features were identified.
Each student analyzed the alpha and beta diversity of a subset of the samples and visualized the data through PCoA and taxonomy bar plots. Although students did not always find significant differences between surfaces in terms of alpha and beta diversity, they were usually able to discern distinct patterns in microbial composition at different sample sites via principle coordinates analysis.
The main challenge of the “Microbes@Home” module is the use of AWS for novices. If students have access to lab computers with QIIME installed, they do not need to use AWS, and the analysis can be scaled up to include more samples. The module can also be extended to the microbial communities of other environments accessible to undergraduate laboratories, such as soil samples and fermented food products.
The implementation of the Microbes@Home module engages students into the research process through hypothesis-driven experiments and discovery, while developing not only laboratory skills at the bench, but also bioinformatics skills at the command line.

Evaluation of taxonomic profiling methods for long-read shotgun metagenomic sequencing datasets

Daniel M. Portik 1, C. Titus Brown 2, and N. Tessa Pierce-Ward 2

1 Pacific Biosciences, 1305 O'Brien Dr, Menlo Park, California 93025 USA
2 Department of Population Health and Reproduction, University of California Davis, Davis, California USA

Background: Long-read shotgun metagenomic sequencing offers many advantages over short-read sequencing. Long reads contain higher information content which is useful for a variety of metagenomics analyses, including taxonomic profiling. New long-read specific tools for taxonomic profiling are being developed, yet there is a lack of consensus regarding their relative performance.

Methods: To address this, we performed a critical benchmarking study using both long-read and short-read methods. We applied these tools to several mock community datasets generated using PacBio HiFi or Oxford Nanopore Technology (ONT) sequencing. We evaluated method performance based on read utilization, detection metrics, and relative abundance estimates.

Results: We show that long-read methods generally outperformed short-read methods. Short-read methods (including Kraken2, Bracken, Centrifuge, and MetaPhlAn3) produced many false positives (particularly at lower abundances), required heavy filtering to achieve acceptable precision (at the cost of reduced recall), and produced inaccurate abundance estimates. By contrast, several long-read methods displayed near-perfect precision and suitable recall without any filtering required. These methods include BugSeq, MEGAN-LR using translation alignments (DIAMOND to NCBI nr) or nucleotide alignments (minimap2 to NCBI nt). Based on the abundance designs of the mock communities, we show that long-read methods detect all species down to the 0.1% abundance level with high precision in the PacBio HiFi datasets. Other methods (MetaMaps and MMseqs2) required filtering to reduce false positives and balance precision and recall. Read quality affected performance for methods relying on protein prediction or exact kmer matching, and these methods performed better with higher accuracy reads (PacBio HiFi reads). Finally, for a given mock community the long-read datasets produced better results than short-read datasets, demonstrating advantages for long-read metagenomic sequencing with regards to taxonomic profiling. Our critical assessment of methods provides recommendations for current research using long reads and establishes a baseline for future studies.

Enterococcus casseliflavus: Naturally vancomycin resistant motile species of Enterococcus

Janira Prichula1,2; Matheus M. S. M. Lima3; Adriana Seixas4; Michael S. Gilmore1,2

1 Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, MA 02114, USA; 2 Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA; 3 Bioinformatics Multidisciplinary Environment, Digital Metropolis Institute, Federal University of Rio Grande do Norte, Natal, RN, Brazil. 4 Department of Pharmacosciences, Federal University of Health Sciences of Porto Alegre, Porto Alegre, RS, Brazil.

Enterococcus casseliflavus is unusual among enterococci because it is motile, yellow-pigmented, and all isolates of the species harbor the VanC gene cluster, which reduces susceptibility to vancomycin, as a core chromosomal trait. The occurrence of this gene cluster long predates humans or human antibiotic use. Little is known of the nature of the environment to which this species adapted. To infer clues, we launched a study to systematically examine all available genomes in the NCBI database of E. casseliflavus (n=100) for insight into its ecology. Following quality curation 85 genomes were selected, being the largest group derived from human (29.1%), with the second largest group isolated from the environment (24.4%) but not a particular host. Average Nucleotide Identity (ANI) was determined for all 85 genomes. Surprisingly, of those, 82.35% (n= 70/85) possessed £ 95% ANI when compared to the prototype strain ATCC 12755 E. casseliflavus, with most of those belonging to the sister species E. flavescens based on >95% ANI with prototype strain ATCC 49996 E. flavescens. This reflects ongoing controversy about the relationship between E. casseliflavus and E. flavescens. Resistome analysis showed that all genomes harbor a vanC operon as expected. Interestingly, two genomes also harbor the vanD operon, which encodes resistance to high-level vancomycin. Additional antibiotic resistances were identified in 14 of the 85 genomes (12 of those 14 were derived from clinical or agricultural sites), including resistances conferred by ermB, tet(45), tet(C), tetM, and tetS, optrA, and poxtA genes. Pangenome analysis estimated a total of 18,938 genes, with 6.09% (1,155) being core to all strains. There is considerable confusion in the NCBI database with respect to species designation within the casseliflavus group, and the radiation of the species E. flavescens appears to be recent, and as for that within E. faecium, may be attributable to anthropogenic activity.

Isolating receptor dependent phages from the environment

Eleanor Rand1,2, Siân Owen1, Natalia Quiñones-Olvera1,2, Michael Baym1

1 Department of Biomedical Informatics and Laboratory of Systems Pharmacology, Harvard Medical School
2 Systems, Synthetic, and Quantitative Biology PhD Program, Harvard University

Bacteriophages (phages for short) are of growing interest both for their clinical uses as well as their evolutionary impacts. While there are a few well characterized phages that have been used across biology for decades, most phages in the environment remain uncharacterized. One of the most influential aspects used to determine a phage's use in the clinic or its ecological importance is the host cell component it requires as a receptor to begin infection. We developed an assay to quickly screen for and isolate environmental phages dependent on specific receptors. Our method expands upon the traditional plaque assay by plating multiple bacterial strains, each with one protein knockout and a unique fluorescent label, together on one plate with an environmental sample. Phages that display differential plaquing between the strains are likely to be dependent on a protein of interest. This method can be extended to any non-essential gene. Using this method, we have isolated phages dependent on proteins of interest in just a few days. Understanding the abundance and diversity of phages that interact with specific receptors could help inform the phages chosen for clinical use as well as contribute to our understanding of the ecological effects phages impose in the environment.

Isolating high quality nucleic acids for genome scanning and whole genome sequencing of “Haptophyte Genome Project”

Betsy Read, Hossein Khadivar

California State University San Marcos, Department of Biological Sciences

Haptophytes are a clade of marine phytoplankton. There are more than 300 extant species that range in size from between 2-40 μm. A number of haptophytes are coccolithophores that produces elegantly sculpted calcium carbonate cell coverings and significantly impact global carbon and sulfur cycling. Others exhibit mixotrophic tendencies, produce unusual very long fatty acids 〖(C〗_(36-42)) known as alkenones, form massive blooms, and synthesize harmful toxins. Although they extend deep into the tree of life, the evolutionary history of this important phylum of phytoplankton remains unresolved. For these many reasons the Department of Energy’s joint Genome Institute aims to sequence the genomes and transcriptomes of 35 members of the haptophyte phylum. The focus of this project was to grow 20 of these haptophytes and extract high quality nucleic acids for genome scanning to determine size and complexity, and eventually whole genome sequencing using PacBio and Illumina platforms. Haptophytes were grown in filtered-seawater supplemented with F/2, K/2 or K/8 vitamins and trace metals. DNA isolation was done during mid-late log phase. RNA was extracted to facilitate gene calling and was performed in the middle of the light and dark phases to maximize the number of genes expressed. Agarose gel electrophoresis was used to ascertain the integrity and size of the extracted DNA. High quality genomic DNA has been extracted from 20 species. Results from nucleic acid extractions and genome scanning will be presented, together with progress on RNA extractions.

Characterization of localized skin microbes recovered from patients undergoing immediate breast implant reconstruction to identify features correlated with surgical site infection development

*Olivia N. Rebeck 1, *Bejan Mahmud 1, David Zahrah 1, Meghan A. Wallace 2, Caitlin Johnson 2, Carey-Ann D. Burnham 2,3,4,5, Margaret A. Olsen 6, Gautam Dantas 1,2,3,7 * denotes equal contribution

1 The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO, USA.
2 Department of Pathology and Immunology, Washington University School of Medicine in St. Louis, MO, USA.
3 Department of Molecular Microbiology and Microbial Pathogenesis, Washington University School of Medicine in St. Louis, MO, USA.
4 Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA
5 Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA.
6 Division of Infectious Diseases, Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA. [7] Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO

Background: Breast cancer (BC) accounts for 15% of cancer-related deaths in women worldwide. Although mastectomy is a common and effective component of BC treatment, there is a high rate of surgical site infections (SSI) in patients undergoing mastectomy with immediate breast implant reconstruction (IBR). We hypothesize that the local skin microbiota is a reservoir for SSI-causing bacteria, such as Staphylococcus epidermidis and Cutibacterium acnes, and that differences in the community composition of the local skin microbiota can predispose a patient towards SSI development.

Methods: To test this hypothesis, we performed 16S rRNA gene sequencing on breast and axilla skin swabs and whole genome sequencing on isolates recovered from skin, implant, and seroma fluid samples of a cohort of patients undergoing mastectomy with IBR. This cohort consisted of 61 patients with risk factors for SSI development (e.g., obesity, previous SSIs) enrolled in two studies. Samples were collected across up to six timepoints (including pre- and post-mastectomy timepoints) from patients who developed infections, noninfectious wound complications, or no complications.

Results: To date, 192/432 metagenomic skin swabs have been sequenced. These metagenomes will undergo taxonomic identification via DADA2. We will analyze these sequencing data to assess differences in community composition. We have also sequenced 393 S. epidermidis and C. acnes isolates. We will analyze these sequencing data to track and compare isolates and characterize their microbial features, including virulence factors and antibiotic resistance genes. Together, we will use predictive modeling to determine which features of metagenomes and isolates are differentially associated with samples collected from patients who develop infections compared to samples collected from patients who did not.

Discussion: Robust characterization of these microbes can identify informative features that correlate to individual predisposition to SSI development, which can inform patient care by identifying patients at increased risk of SSI and improve antibiotic stewardship.

Plasmid evolution is guided by intracellular competition

Fernando Rossine, Michael Baym

Harvard Medical School

Bacterial plasmids carry historically variable loads of medically and ecologically relevant genes such as those responsible for antibiotic resistance or accessory metabolism. Yet, the evolutionary dynamics that drive this variation are poorly understood. Crucially, an arising plasmid variant must face challenges in two scales: bacteria hosting a new plasmid variant compete with bacteria hosting the ancestral type, but within each bacterium the new plasmid variant may directly compete for replication with the ancestral type. The extent to which intracellular competition for replication shapes plasmid evolutionary dynamics is unknown, as are the plasmid features that might favor its intracellular fixation. Here we propose that conflicts between plasmid replication and transcription limit the efficiency of plasmid replication, and therefore that low-transcription plasmids might be favored in the intracellular competition for replication. We develop a method to measure pairwise intracellular plasmid competition and show that plasmid transcription-replication conflicts can drive plasmid evolution regardless of cell-level costs or benefits. We start by setting up an experimental population in which each bacterium contains a well defined initial frequency of both plasmid variants. We achieve this by joining both plasmid variants into a single chimeric plasmid dimer that is then transformed into E. coli cells. Upon induction of a novel thermosensitive FLP-like recombinase, the dimers split into their component monomers, ensuring that each cell in the bacterial population contains both plasmid variants in equal frequencies. We find that when our plasmids differ in a few base pairs controlling the strength of a promoter, the weaker promoter plasmid increases in frequency as long as there is still co-occurrence of both plasmid types. More generally, these results indicate that intracellular competition might be a major deterrent to the plasmidial acquisition of new genes, as well as pose an alternative explanation for the longstanding observation of silenced plasmidial genes.

Longitudinal sequencing and variant detection of SARS-CoV-2 across Southern California wastewater

Jason A. Rothman; Katrine L. Whiteson

University of California, Irvine

Wastewater based epidemiology (WBE) is a useful method to detect pathogen prevalence and may serve to effectively monitor diseases at broad scales. WBE has been used throughout the COVID-19 pandemic to track disease burden through the quantification of SARS-CoV-2 RNA present in wastewater. Aside from case load estimation, WBE is being used to assay viral genomic diversity and the emergence of potential SARS-CoV-2 variants. Unlike sequencing SARS-CoV-2 in healthcare settings, WBE has been used to sequence virus without adding additional testing or supplies burden to the healthcare system.

Here, we present a study in which we sequenced RNA extracted from sewage influent samples obtained from eight wastewater treatment plants representing 16 million people in Southern California over April 2020 – August 2021. We sequenced SARS-CoV-2 with two methods: Illumina Respiratory Virus Enrichment and metatranscriptomic sequencing (N = 275), and QIAseq SARS-CoV-2 tiled amplicon sequencing (N = 96). We were able to build nearly full-length consensus genomes which largely mirrored emergent SARS-CoV-2 clades isolated from sequences derived from healthcare settings. Notably, in samples where RNA was too fragmented to assign clades, we identified single nucleotide variants (SNVs) of the virus that often corresponded to variants of interest and/or concern. Through our analyses, we also detected a large diversity of putatively novel SARS-CoV-2 SNVs, and SNVs of unknown potential function and prevalence. As the COVID-19 pandemic continues, and novel SARS-CoV-2 variants emerge, the ongoing monitoring of wastewater is an important tool in understanding local and population-level dynamics of the virus.

Our study shows the potential of WBE to detect SARS-CoV-2 variants throughout Southern California's wastewater and to track the diversity of viral SNVs and strains in urban and suburban locations. These results will aid in our ability to monitor the genomic potential of SARS-CoV-2 and help understand circulating SNVs to further combat COVID-19.

CRISPR-based killswitches for engineered microbes - stability and in vivo application

Austin G. Rottinghaus, Aura Ferreiro, Skye R.S. Fishbein, Gautam Dantas, Tae Seok Moon

Department of Energy, Environmental and Chemical Engineering, Washington University in St. Louis, St. Louis MO USA (AGR, TSM); The Edison Family Center for Genome Sciences & Systems Biology, Washington University in St. Louis, St. Louis MO USA (AF, SRSF, GD); Department of Biomedical Engineering, Washington University in St. Louis, St. Louis MO USA (AF, GD); Department of Pathology and Immunology, Washington University in St. Louis, St. Louis MO USA (SRSF, GD); Department of Molecular Microbiology, Washington University in St. Louis, St. Louis MO USA (GD); Division of Biology and Biomedical Sciences, Washington University in St. Louis, St. Louis MO USA (GD, TSM)

Background: Microbial biocontainment is an essential goal for engineering safe, next-generation living therapeutics. However, the genetic stability of biocontainment circuits such as genetically-encoded kill-switches is a challenge that must be addressed. Kill-switches are among the most difficult circuits to maintain due to the strong selection pressure they impart, often leading to evolution of escape mutant populations.

Methods: Here we engineer two CRISPR-based kill-switches in the probiotic Escherichia coli Nissle 1917, a single-input chemical-responsive switch and a 2-input chemical- and temperature-responsive switch. To maximize genetic stability, we employ parallel strategies including functional redundancy within the circuit, mitigation of the bacterial SOS response, antibiotic-independent plasmid maintenance, and innovatively, provision of intra-niche competition by a closely related strain. We test the efficacy of these kill-switch strains in vitro and in vivo, and characterize escape mutant populations through functional assays and sequencing of kill-switch circuit parts.

Results: We address two routes of DNA mutagenesis that contribute to kill-switch instability: i) stochastic inactivation of the kill-switch through introduction of functionally redundant Cas9 and gRNA expression cassettes, and ii) SOS-mediated DNA mutagenesis in response to Cas9-driven double-stranded genome breaks through knockouts of key SOS response elements. We further improve kill-switch stability through an antibiotic-independent plasmid maintenance system, and minimize the emergence of escape mutant populations through intra-niche competition provided by a related probiotic strain. Using these parallel strategies, we demonstrate that strains harboring these kill-switches can be selectively and efficiently killed inside the murine gut, or upon excretion in the case of the 2-input switch.

Discussion:
We developed CRISPR-based kill-switches in EcN towards a safe probiotic chassis for biomedical technologies, and which enable on-demand selective removal of engineered microbes from the gut. We explored diverse methods for improving the kill-switch stability and minimized key mechanisms of circuit inactivation, providing a template for future kill-switch development.

Comparative Genomics Of Diagnostic And Longitudinally-sampled Community-origin Staphylococcus Pseudintermedius Reveals The Absence Of Niche Adaptation But Widespread Strain Persistence Within Households

Sanjam S. Sawhney 1, Rhiannon Vargas 1, Meghan A. Wallace 1, Carol Muenks 1, Brian V. Lubbers2, Stephanie Fritz1, Carey-Ann D. Burnham*1, and Gautam Dantas*1

1 Washington University in St. Louis, St. Louis, MO
2 Kansas State University, Manhattan, KS; # authors contributed equally;
* corresponding authors

Background: Staphylococcus intermedius group (SIG) describes closely-related coagulase-positive staphylococcal species – Staphylococcus pseudintermedius (S.p.), delphini, and intermedius – associated with a range of host-species. Efforts investigating community MRSA spread have serendipitously identified widespread SIG colonization of household surfaces as well as human skin and fur of pet dogs and cats; additionally, clinical labs have reported increasing human infection resulting from SIG. In consideration of the growing host range and disease burden, we endeavored to interrogate the interrelatedness and potential niche adaptation of SIG across hosts and environments, and to assess the spatiotemporal dynamics of SIG.
Methods: We have assembled 572 SIG genomes collected from human diagnostic samples at the Barnes-Jewish Hospital (St. Louis, MO, n=181), animal diagnostic samples at Kansas State Veterinary Diagnostic Laboratory (n=100), and in community settings (St. Louis metropolitan area, n=291). All isolates were identified via MALDI-TOF (VITEK-MS) and underwent Illumina whole genome sequencing (WGS) for strain-level analysis.
Results: VITEK-MS had 100% sensitivity and 87.5% specificity at calling WGS-validated S.p. Remarkably, WGS did not identify niche adaptation between diagnostic and household S.p. by core, accessory, or gene family content, in contrast to reports on S. aureus; however, diagnostic S.p. had a higher ARG count ($p<0.001$) and isolates with the greatest ARG burden were often methicillin-resistant ($p<0.001$). Deeper analysis on strain persistence identified 72 strain-level clusters (≤45 pairwise SNPs at ≥98% coverage) comprising ~40% of isolates, primarily among intra-household isolates but also in diagnostic isolates from temporally-adjacent patients. Interestingly, a CRISPR-Cas operon and pathogenicity island were differentially present among diagnostic isolates of different strain backgrounds, while putative mobilizable elements differentiated certain isolates within a strain.
Conclusion: We argue against niche adaptation within S.p. but highlight widespread strain persistence within households, identifying genetic elements – often mobilizable – that differentiate isolates within the same strain and unite isolates of different strain backgrounds.

Preventing emergence of daptomycin resistance in Enterococcus faecium

Suelen Scarpa de Mello1, 2, Andrew G. Van Camp1, 2, Michael S. Gilmore 1, 2

1 Massachusetts Eye and Ear Infirmary, Boston, MA.
2 Harvard Medical School, Boston, MA.

Daptomycin is a last-line drug for treating multidrug-resistant E. faecium infections but is compromised by the emergence of resistant mutants. We discovered that defects in the gene lafB, which encodes a glycosyltransferase involved in lipoteichoic acid (LTA) biosynthesis, render E. faecium hypersusceptible to daptomycin. If lafB is dominant to other daptomycin resistance mutations, identifying a drug capable of blocking lafB would not only enhance daptomycin susceptibility by >10-fold but also block emergence of other daptomycin resistance mutations. To assess the dominance of a lafB mutation, we first generated an E. faecium strain with a clean deletion in lafB, which as expected was hypersusceptible to daptomycin. We next introduced mutant forms of liaR and cls known to confer increased daptomycin resistance to wild-type cells. However, they did not confer resistance in the lafB deletant, indicating that lafB mutation is in fact dominant to known daptomycin resistance mechanisms. Finally, we attempted to evolve daptomycin-resistant derivatives of the lafB-deleted strain. While controls evolved a 100-fold increase in resistance to daptomycin, little change was observed for the lafB-deleted strain. Little is known about the specific role of the disaccharide link between the diacylglycerol membrane anchor and polyphosphate domain of LTA, the formation of which is catalyzed by LafB, or why it varies. To identify variants related to lafB, we constructed a lafB gene tree. We observed that a subgroup of enterococci cluster with vagococci in possessing a variant form of the gene, indicating that these species likely produce an LTA with an altered disaccharide linkage. We are now comparing the properties of those cells to determine the potential contributions that this link makes to the physiology of the cell. We are also launching a small molecule screen for LafB inhibitors to identify compounds that sensitize E. faecium to daptomycin and prevent emergence of resistance.

Plant pathogenic Ralstonia species have multiple chemoreceptors for amino acids with overlapping, but divergent chemical specificities

Schomer R.A. 1, Avalos, J.K. 1,2, Lowe-Power, T.M. 2, Parales, RE. 1

1 University of California, Davis, Department of Microbiology & Molecular Genetics.
2 University of California, Davis, Department of Plant Pathology.

Background: Ralstonia solanacearum, a soil-borne plant pathogen, exploits motility and chemotaxis (directional movement along a chemical gradient) to locate host roots. Membrane-bound chemoreceptors drive chemotaxis and contain a conserved signal transduction domain and highly sequence divergent ligand binding domains (LBDs). We hypothesize that Ralstonia's repertoire of chemoreceptors and their respective chemical ligands drive plant host discovery and preference from the soil. Here, we target chemoreceptors that sense amino acids, a common component in root exudates.

Methods: To probe the chemotactic range of Ralstonia, we generated chimeras by fusing LBDs of 42 Ralstonia chemoreceptors with the transduction domain of a two-component system that controls gene transcription in E. coli. Using fluorescent reporter strains designed to measure transcriptional readout, we characterized the ligand ranges of LBDs from Ralstonia model strains with varying host ranges. A recently discovered LBD motif for amino acids allowed us to target predicted amino acid sensors.

Results: Across the genetically diverse Ralstonia solanacearum species complex, genomes contain two paralogous amino acid sensors, McpA and Rsc3307. We characterized the ligand range of McpA and Rsc3307 using the chimeric receptors. Interestingly, despite high sequence similarity between the Rsc3307 orthologs, the amino acid sensors tested from broad-host range Ralstonia solanacearum IBSBF1503 and narrow-host range Ralstonia syzygii PSI07 strains had variable ligand ranges. Using mcpA and rsc3307 mutants, their influence on the chemotactic range of Ralstonia and their effect on Ralstonia ability to infect plant hosts from the soil is being evaluated.

Discussion: To understand the role chemotaxis plays in host-microbe interaction, we must first identify which chemical signals bacteria detect. Our data indicate that subtle changes in LBD sequence can have significant influences on which chemicals a bacterium can detect. Variability in the chemical repertoire of a soil-borne plant pathogen could affect how effectively it senses host plants' roots or not.

Bloodstream Infection is Associated with Prior Gut Microbiome Colonization of the Same Strain in the Neonatal Intensive Care Unit

Drew J. Schwartz 1,2,3,4,5; N. Shalon 5; K. Wardenburg 5; C. Hall-Moore 1; B. Warner 1; P.I. Tarr 1,4; G. Dantas 4,5,6,7

1 Department of Pediatrics
2 Center for Women's Infectious Diseases Research
3 Department of Obstetrics and Gynecology
4 Department of Molecular Microbiology
5 The Edison Family Center for Genome Sciences and Systems Biology
6 Department of Pathology and Immunology
7 Department of Biomedical Engineering; Washington University School of Medicine

Background: Infants hospitalized in neonatal intensive care units (NICUs) have a high incidence of bacterial bloodstream infections (BSIs) leading to significant mortality. Common etiologies of BSI are pathogens that also colonize the infant gut. However, differentiating neonates who develop BSI with these organisms from those who are simply colonized remains challenging. We hypothesized that the same strain increases in abundance in the gut of affected neonates prior to BSI.

Methods: We performed shotgun metagenomic sequencing on 462 fecal samples from 19 infants with BSI and 37 controls in the same NICUs. We utilized MetaPhLan3 to infer taxonomy and InStrain to track the whole genome sequence of the BSI pathogen within the fecal metagenomes longitudinally. We used MaAsLin2 to determine significant associations between the microbiome and metadata.

Results: We found no differences in gut microbiome diversity between cases and controls. Infants who experience a BSI possess significantly higher relative abundance of the causative species in their gut relative to paired controls in the 2 weeks prior to bacteremia (median 21.4% vs. 0.4%, $p<0.05$). 10/19 infants produced at least 1 stool before BSI containing an organism with 0-1 genomic substitutions from the BSI isolate, establishing isogenicity between gut residing and invasive pathogen. The coverage (abundance) of the BSI isolate increased in the stool of affected neonates in the 10 days before infection to an average maximum of 0.48/1.0 (linear mixed effect model $p<0.05$). We observed the strongest associations between high gut microbiome abundance of E. faecalis, Klebsiella spp., E. coli, and S. marcescens of cases compared to all other samples ($p<0.01$).

Discussion: We show high relative abundance of the same strain that causes BSI in stool in hospitalized neonates. These data suggest the potential of gut microbial surveillance and management to identify blooms of gut pathogens to limit the incidence of bacteremia.

Arthropods are the Source of Genetically Diverse Enterococci

Bruna Sgardioli1 , Maria de Lurdes Enes Dapkevicius1, Paulo A.V. Borges1 , Ryan Whipple2, Abigail L. Manson3 , Ashlee M. Earl3, and Michael S. Gilmore2,4

1 University of the Azores, Angra do Herismo, Terceira, Portugal.
2 Harvard Medical School and Mass Eye and Ear, Boston, Massachusetts, USA.
3 Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.
4 Harvard Medical School, Boston, Massachusetts, USA.

Enterococci, leading causes of multidrug-resistant hospital infection, diverged from a common ancestor with vagococci (commonly found in the guts of fish) at the time animals colonized land and can be isolated from the guts of most land-dwelling animals – insects to humans. The first to colonize land were arthropods, and over 1,000,000 species of insects and invertebrates are believed to still exist. Preliminary work suggested that substantial enterococcal genetic diversity exists within insects and invertebrates. To test this, we collected insects and other invertebrates on the remote island of Terceira in the Azores. To maximize diversity and to detect the possible impact of human activity, insects were captured a) from the pristine wilderness at the top of the volcanic island, b) from agricultural pasturelands in the middle, and c) from the human-inhabited coastal region. From 125 insects, 125 presumptive enterococci were isolated and genomes sequenced. Twelve known enterococcal species were identified, the most common being E. faecalis, E. casseliflavus and E. mundtii. In addition, we obtained 13 with genomes more than 5% divergent (ANI) from a known species, indicative of novel species. Of candidate new species, 3 were also encountered in a parallel study in the lab. These 5 unique new species were obtained from Diptera, Annelid, Apidae, Syrphidae, Mythimna unipuncta, Popillia japonica, Forficula auricularia, Calosoma olivieri and Ocypus olens hosts. Given that we only have sampled 10’s of different insect species in total, we project that there are thousands of species of Enterococcus to be found possessing substantial genetic diversity. Enterococci are prolific genetic exchangers, and given the ubiquity of the species E. faecalis, which also colonizes and infects man, there is likely a very large pool of genes tuned to expression in the enterococcal background that pose a potential risk.

Hypermodification of dA by Mu Mom requires aminoacylated tRNA Gly

Rebekah M. B. Silva* 1, Anton Slyvka* 2, Yan-Jiun Lee 1, Chudi Guan 1, Sean Lund 1, Matthias Bochtler 2.3, and Peter R. Weigele

1 New England Biolabs
2 International Institute of Molecular and Cell Biology
3 Institute of Biochemistry and Biophysics, Polish Academy of Sciences

DNA hypermodification enzymes from bacteriophage install abundant and varied chemical groups on canonical bases that are protective against native host defenses (ex. Restriction endonucleases). Historically, much less has been understood about hypermodification of purines compared to pyrimidines. A classic example is the post-replicative hypermodification of adenine found in the phage Mu genome. The product of a single gene, mom (Modification of Mu), leads to an unusual carbamoylmethyl moiety found at the N6 position of adenine (6-NcmdA). Despite great effort, the group donor transferred by the Mom enzyme has eluded identification for over four decades. Mom has been predicted to belong to the Gcn5-Related N-Acetyltransferase superfamily (GNATs), most-commonly found to use acyl-CoA co-substrates for acylation of proteins, lipids, tRNA, and natural products. However, reconstitution of Mom activity in vitro has not been reported to date, even in the presence of putative co-substrates at high concentrations. Employing a suite of computational, analytical, biochemical, and biophysical tools, we have demonstrated that Gly-tRNA Gly is the co-substrate used by Mom to hypermodify dA, reconstituted the activity of Mom in vitro, and collected evidence for an on-base rearrangement of the glycyl group to yield the final structure of the Mu hypermodification. Using the predicted structure and an apo X-ray structure of Mom, we are gathering additional insights on the active site of Mom as well as the domains involved in co-substrate and substrate recognition. Furthermore, we have mined metagenome databases to find active Mom homologs and are currently aiming to harness the unique chemistry accomplished by Mom and Mom homologs to add new functionalities to DNA.

Genomic epidemiology and bacteriophage susceptibility over time of vancomycin-resistant Enterococcus faecium from bloodstream infections at a single hospital

Madison E. Stellfox, Hayley R. Nordstrom, Ryan K. Shields and Daria Van Tyne

University of Pittsburgh, Division of Infectious Diseases

Background: Enterococcus faecium is a gastrointestinal commensal and a difficult-to-treat nosocomial pathogen. Bacteriophage therapy is increasingly being studied as a potential treatment for E. faecium infections, and we recently treated a case of E. faecium bacteremia with two bacteriophages (9184 and Hi3). An important question regarding clinical use of bacteriophage therapy is scalability from personalized regimens to broader therapeutic tools. Therefore, we sequenced the genomes of 60 vancomycin-resistant E. faecium (VREfm) bloodstream isolates, contemporary with our clinical phage therapy experience, and characterized the lytic activity of both 9184 and Hi3 on all isolates.

Methods: 60 VREfm isolates from bloodstream infections in 58 unique patients were collected over two years and were sequenced on the Illumina platform. The earliest isolate was selected as a reference and was assembled with SPAdes and annotated with prokka. Variant calling was performed with snippy and a SNP-based, maximum likelihood phylogenic tree was constructed with RAxML. Pairwise SNP comparisons were used to assess genetic relatedness between isolates. Phage susceptibility of each isolate was determined in biologic duplicate with the soft agar overlay method.

Results: The 60 isolates spanned 12 different sequence types (ST), many of which are known to be hospital-associated. Using a 10 SNP cut-off, one third of the isolates were closely related to others in the dataset, suggesting patient-to-patient transmission. The probability of an isolate's phage susceptibility was stable over the study period. Approximately two-thirds of all isolates were susceptible to one or both phages, and susceptibility loosely clustered amongst isolates belonging to the same ST.

Discussion: Our findings suggest that a significant proportion of VREfm bloodstream infections are likely hospital-acquired. Differences in phage susceptibility between closely-related isolates will be investigated further. Overall, this study can serve as an initial step towards identifying genetic and phenotypic characteristics that could predict VREfm phage susceptibility.

Scalable pangenome-wide characterization of genetic variation in microbial species with Fusilli

Lucas R. van Dijk1,2, Abigail L. Manson1, Thomas Abeel2, Ashlee M. Earl1, Kiran V Garimella3

1 Infectious Disease & Microbiome Program, Broad Institute, Cambridge, MA 02142, USA
2 Delft Bioinformatics Lab, Delft University of Technology, Van Mourik Broekmanweg 6, 2628 XE Delft, The Netherlands
3 Data Sciences Platform, Broad Institute, Cambridge, MA 02142, USA

Identifying genetic variation among bacteria is critical for understanding their evolution, spread and functional potential. To date, most variant calling is done using a single reference, which often represents only a small fraction of a species pangenome, limiting the extent of variation that can be measured in an unrelated isolate. Graph-based approaches that accommodate additional species variation for interrogation are increasingly popular though most have limited applicability to bacteria. We developed Fusilli, a novel pangenome-wide variant calling tool designed specifically for diverse microbial species that combines multiple reference genomes in a Linked De Bruijn Graph and uses a multi-reference-assisted assembly algorithm to construct variant haplotypes. Fusilli works directly from raw reads, scales to hundreds of references and samples, and is able to identify more single nucleotide variants across a larger number of genes compared to single-reference tools, while also capturing large structural variants including unique genes. This tool will greatly improve our ability to analyze diverse bacterial genomes, including calling variants across an entire pangenome and rapidly detecting horizontally transferred and novel genes.

Characterizing the soil microbial community composition of MtDef4 defensin Arabidopsis variants

Rhiannon Vargas1, Maggie Newberry 1, Kimberley Sukhum 1,2, Gautam Dantas 1,2,3,4

1 The Edison Family Center for Genome Sciences and Systems Biology, Washington University in St. Louis School of Medicine, St. Louis, MO USA
2 Department of Pathology and Immunology, Washington University in St Louis School of Medicine, St Louis, MO, USA
3 Department of Molecular Microbiology, Washington University in St Louis School of Medicine, St. Louis, MO, USA 4 Department of Biomedical Engineering, Washington University in St Louis, MO, USA

Background: Plant defensins such as MtDef4 have antimicrobial properties and have the potential to be implemented in agricultural crops to decrease pathogen burden. Little is known about how plants with the MtDef4 gene influence plant growth and belowground microbial communities. Subcellular localization of transgenic defensins is shown to be a determinant of its antimicrobial activity but they are not well characterized with MtDef4. We predicted that the microbial community assembly on the roots and in the soil of Arabidopsis thaliana expressing MtDef-4 extracellularly (Ec) or in the endoplasmic reticulum (ER) will be different from the microbial communities of Col-0 wild-type (WT) plants.

Methods: A longitudinal study design was conducted with three Arabidopsis genotypes - Ec, ER, WT – and soil as a control. Growth measurements were collected along with three belowground compartments for each plant – endophytic root, rhizosphere soil, and bulk soil – at time of harvest. Plants were randomly assigned a harvest day starting at day 0 (post-germination), 24 hours, and week 1, 2, or 3. Each factor had a replicate of 10 (n=200). 16S rRNA and ITS sequencing were conducted to characterize the bacterial and fungal communities respectively.

Results: MtDef4 plants grew at a faster rate than WT plants. Bacterial alpha diversity and plant biomass were positively correlated. The composition of bacterial communities changed temporally but was generally consistent between genotypes, whereas subtle beta diversity differences were observed. Two phyla of interest that differ between genotypes overtime are Verrucomicrobia and Acidobacteria. Preliminary ITS results suggest that there may differences in Basidiomycota and Ascomycota phyla in the rhizosphere between genotypes.

Discussion: Differences between genotypes are shown when comparing growth rates, beta diversity overtime, and certain microbial phyla. Future directions include investigating significant bacterial and fungal correlations. This study will inform fundamental questions regarding the biological impacts of the MtDef4 plant-microbiome overtime.

Progress Towards Standardizing Metagenomics- Application of Metagenomic Reference Materials to Develop a Reproducible Microbial Lysis Methodology with Minimum Bias

Michael M. Weinstein 1,2, David C. Danko 3,4, Elaine Wolfe 1, Shuiquan Tang 1, Karen G. Jarvis 5, Christopher J Grim 5, Venu Lagishetty 6, Jonathan P. Jacobs 6, Jason Arnold 7, Ryan Kemp 1, and Christopher Mason 4,8

1 Department of Bioinformatics, Zymo Research Corp., Irvine, CA, USA.
2 Molecular, Cell, and Developmental Biology/Quantitative and Computational Biosciences Institute, University of California, Los Angeles, CA, USA.
3 Tri-Institutional Computational Biology and Medicine Program, Weill Cornell Medicine of Cornell University, New York, NY, USA.
4 Institute for Computational Biomedicine, Department of Physiology and Biophysics, Weill Cornell Medicine of Cornell University, New York, NY, USA.
5 Center for Food Safety and Applied Nutrition, U.S. Food and Drug Administration, College Park, MD, USA. (6) Microbiome Core, Division of Digestive Diseases, Department of Medicine, University of California, Los Angeles, CA, USA.
7 Microbiome Core, School of Medicine, University of North Caroline, Chapel Hill, NC, USA.
8 The Feil Family Brain and Mind Research Institute, Weill Cornell Medicine of Cornell University, New York, NY, USA

Background: Citations involving microbiome analysis have undergone an exponential growth since approximately 2010 with little evidence of slowing down. Along with this rapid growth has come evidence of technical weaknesses causing a lack of reproducibility between different methods. This lack of reproducibility limits our ability to draw new conclusions from microbiome analyses and can call previous conclusions into question. To address these issues, leaders in the field have highlighted the importance of standardized protocols. However, the likelihood of a substantial portion of the microbiome research community standardizing around a single method is unlikely and will still leave unanswered the questions about the reality of a sample's nature.

Methods: To address these and similar challenges, we sought to compare the performance of many different microbiome lysis and extraction methods from both commercial and academic sources when extracting DNA from a mock microbial community standard of defined composition. In order to compare accuracy between different methods, we performed 16S rRNA gene sequencing as well as shotgun metagenomic analysis. Additionally, we present a new statistical measurement: The Measurement Integrity Quotient, providing a single, easy to understand number score that describes the accuracy of an observed composition from a known composition standard. Utilizing this method, a

comprehensive analysis of many variables potentially determining lysis efficiency were tested, including differences between thermal, enzymatic, and mechanical lysis.

Results/Discussion: Utilizing the data gathered from these analyses, we were able to determine which methods are likely to produce nearly complete microbial lysis. Additionally, we have developed and publicly released open-source bioinformatics packages for rapidly assessing the accuracy of a microbial extraction method relative to a defined composition microbial standard. The use of these data and analyses enables the selection of methods that generate minimal bias and facilitate benchmarking and quality control for many different microbiome pipelines.

Characterization of the gut metagenome in adults and children with asthma

Naomi G. Wilson, Leonard B. Bacharier M.D., Andrew L. Kau M.D. Ph.D.

Washington University School of Medicine in St Louis MO, Vanderbilt University Medical Center Nashville TN

Background: Asthma is a common allergic airway disease whose etiology remains elusive. The composition and function of the infant microbiota are linked to asthma risk, but the microbiota's role in older patients with established asthma is less clear.

Methods: Here, we performed metagenomic shotgun sequencing on human stool collected from patients with and without moderate-to-severe asthma to characterize the functional profiles of the gut microbiota in asthma beyond early childhood.

Results: Asthma contributes to beta diversity variation in gene content even when accounting for other important clinical demographics. Long-chain fatty acid metabolism gene-sets were associated with asthma and may be indicative of the involvement of these pathways' metabolites in airway smooth muscle and immune responses already associated with asthma. Antibiotic resistance gene (ARG) profiling revealed that people with asthma have a higher richness but a similar load of ARGs compared to healthy people. Further, we found an ARG enriched in asthma metagenomes that confers resistance to macrolides, which are often prescribed for asthma exacerbations, and most often encoded by Bacteroides species. Given our previous finding that the B. fragilis toxin, BFT, can alter lung inflammation in humanized gnotobiotic mice and is more prevalent in the asthma cohort compared to healthy, we calculated co-occurrence of the macrolide resistance marker and bft. We found that proportionally more MARS samples encode both genes compared to publicly available B. fragilis genomes, suggesting a possible relationship between B. fragilis toxin, antibiotic resistance in the asthmatic gut.

Discussion: Our results are the first to characterize the functional underpinnings of the gut microbiota in moderate-to-severe physician-diagnosed asthma and provide several avenues for future studies of gut-lung axis host-microbe interactions.

Automating microbial taxonomy workflows with PHANTASM: PHylogenomic ANalyses for the TAxonomy and Systematics of Microbes

Joseph S. Wirth, Eliot C. Bush

Harvey Mudd College

The nucleotide sequences of 16S ribosomal RNA (rRNA) genes have been used to inform the taxonomic placement of prokaryotes for several decades. Whole-genome approaches can better resolve evolutionary relationships of organisms, but these analyses often require computational proficiencies that are uncommon among microbiologists. PHANTASM is a new tool capable of automating these workflows. This tool was designed to work for a wide range of prokaryotes and is the first example of an automated reconciliation of NCBI's Taxonomy database with that of the List of Prokaryotic names with Standing in Nomenclature (LPSN). In this study, we describe the workflow of PHANTASM and provide several examples of results generated by it. The source code is freely-available on GitHub. In order to facilitate the ease-of-access for researchers, PHANTASM is also available as a Docker image. While other tools exist to facilitate starting points for these analyses, PHANTASM provides users with a greater degree of control and produces outputs that can be used to make publication-quality figures.

Translational Ramifications of Crowd-Sourced Genomics Data

David A Yarmosh, Juan G. Lopera, Nikhita P. Puthuveetil, Patrick Ford Combs, Amy L. Reese, Corina Tabron, Amanda E. Pierola, James Duncan, Samuel R. Greenfield, Robert marlow, Stephen King, Marco A. Riojas, John Bagnoli, Briana Benton, Jonathan L. Jacobs

American Type Culture Collection (ATCC), Manassas, Virginia, USA

Background: Public genomics databases serve a critical role in the life science research community. Despite existing guidelines which require metadata associated with a given genome assembly, other relevant data (e.g., sequencing platform, assembly method) are often incomplete or missing. Ultimately, this gap renders the assembly data itself questionable from the perspective of reliability, traceability, and accuracy. Previously, Yarmosh et al. illustrated the impact of poor data provenance by comparing several publicly available assemblies to assemblies which had complete traceability. While it was found that some public assemblies were labelled as derivatives of ATCC™ source material, there was a tendency toward fewer relative variants between these assemblies and their ATCC Standard Reference Genome (ASRG) counterparts. However, several of these assemblies still contained a large quantity of variants, including those inducing translational changes.

Methods: To better understand the consequences outlined in a previous study, the Prokaryotic Genome Annotation Pipeline was run on 190 public assemblies labeled as ATCC type material and their 127 corresponding ASRGs. Annotations from both sets were compared in terms of amino acid identity, gene count, gene identity, and the gain/loss of stop codons.

Results: Despite the claim of being assembled from type material, 25 of 190 public genomes contain premature stop codons and over 35,000 of these annotations have less than 50% reciprocal identity relative to their ASRGs.

Discussion: Public genomics databases are unable to curate their immense library of submitted assembly data to ensure the utmost quality. The de facto usage of these databases in modern research coupled with the repercussions of incomplete metadata underscores the urgent need for a more stringent curation process, such that future research and public health are not burdened by unreliable data. Toward this goal, ATCC's Enhanced Authentication Initiative aims to provide high quality reference genomic data directly from ATCC™ source material.

Comparison of long and short read metagenomic assemblers for recovery of low-abundance species

Sosie Yorki; Terrance Shea; Colin J. Worby; Abigail L. Manson; Ashlee M. Earl

Infectious Disease and Microbiome Platform, Broad Institute of MIT and Harvard

Recent technological and computational advances have made metagenomic assembly a viable approach to achieving high resolution views of complex microbial communities. In previous benchmarking studies, short-read (SR) metagenomic assemblers had the highest accuracy, long-read (LR) assemblers generated the longest sequences, and hybrid (HY) assemblers utilizing SR and LR balanced length and accuracy. However, until now, no benchmarking studies have specifically addressed the ability of assemblers to recover low-abundance organisms, which typify pathogens and other clinically relevant organisms in the gut. We generated synthetic long- and short-read metagenomic datasets by computationally spiking E. coli isolate reads into fecal backgrounds at fixed coverage levels, and compared the ability of different assemblers to generate E. coli scaffolds and recover antibiotic resistance genes. Then, we assessed how each assembler's performance changed when a competing E. coli strain was spiked in. Although SR assemblers recovered more resistance gene content with high accuracy, even at low coverages, the increased contiguity in LR assemblies allowed for greater association of E. coli genes with their taxonomic origin. HY assemblies identified resistance genes with high accuracy, but had lower contiguity than LR assemblies. LR and HY assemblers were able to recover most ARGs on E. coli-specific contigs and assemble 99% of an isolate plasmid, even with coverage as low as 4x. Our results highlight the trade-offs associated with each approach, and that the optimal tool is goal-dependent. For strain characterization and determining gene context, LR assembly is optimal, while for base accurate gene identification, SR assemblers outperform other options. HY assembly offers contiguity and base accuracy, though requires generating data on multiple platforms, and may suffer high misassembly rates when strain diversity exists.

Evolving the biofilm transcriptome of Mycobacterium tuberculosis

Madison A. Youngblom, Tracy M. Smith, Mohamed A. Mohamed, Caitlin S. Pepperell

Microbial Doctoral Training Program, University of Wisconsin-Madison; New York Genome Center; Department of Medicine, School of Medicine and Public Health, University of Wisconsin-Madison; Department of Medicine, School of Medicine and Public Health, University of Wisconsin-Madison

Background: Tuberculosis (TB), caused by Mycobacterium tuberculosis (M. tb), is a leading cause of death due to infectious disease. TB is not often thought of as a biofilm infection, however M. tb readily forms a pellicle biofilm at air-liquid interfaces. Often underrecognized is the role that M. tb biofilms play in infection, including their contribution to drug and immune tolerance.

Methods: Here we used experimental evolution and transcriptomics to investigate adaptation of M. tb to this complex phenotype. We first selected a sample of clinical isolates and passaged them under selective pressure to grow as a biofilm. We then performed RNA-sequencing on the ancestral and evolved isolates to compare transcriptomes both across different strains as well as across evolutionary time.

Results: We found significant diversity of genes differentially expressed during biofilm growth among our ancestral isolates, with very few differentially expressed genes (DEGs) shared across isolates. After applying selective pressure to grow as a biofilm however, we observed convergence towards a shared biofilm transcriptome. Even more interesting is the observation that gene expression patterns of our strains converged in a genetic-background dependent manner; strains from two different sub-lineages of L4 evolved distinct biofilm transcriptomes.

Discussion: Elucidating strategies used by M. tb to adapt to new environments will help us to continue develop treatments that overcome M. tb's persistence. We have shown here that even closely related strains of M. tb have diverse transcriptomic landscapes, and respond differently to selective pressure, which has important implications for how we approach treatment of infections with different strains. We have also shown that it is possible to evolve bacterial transcriptomes in the laboratory, opening a new avenue for studying bacterial adaptation in vitro. Overall, this work highlights the complexity of bacterial adaptation, as well as the significant role played by epistasis in evolution to new environments.

Investigating the effects of temperate prophage elements on antibiotic resistance evolution.

Amy Zamora, Siˆan V. Owen, & Michael Baym

Laboratory of Systems Pharmacology and Department of Biomedical Informatics, Harvard Medical School, Boston, MA.

Pathogenic bacteria commonly experience exposure to antibiotics deployed to kill or inhibit growth. This pressure drives bacterial evolutionary pathways by selecting for genomic traits that enhance resilience. Bacterial viruses (bacteriophages) also act as a selective force on bacterial genomes. Importantly, lysogenic bacteriophages integrate into host genomes upon infection and are ubiquitously present as ‘prophages’. Bacteriophage integration is stable, but activation of the SOS response pathway can cause these integrated viruses to become lytic. In this process, bacteriophages exit the genome, replicate, and lyse their host to release new progeny. Several DNA-damaging antibiotics (e.g., Ciprofloxacin) trigger lysis through this process, known as “prophage induction”. It has been demonstrated that prophages in S.typhimurium genomes can increase bacterial susceptibility to antibiotics. This increased sensitivity could be due to phage-mediated death as a result of prophage induction, preventing growth of natural resistance mutants. Although this work elucidated how lysogenic phages affect antibiotic response, the interaction between prophages and antibiotic resistance evolution has yet to be rigorously investigated. To study this relationship, Ciprofloxacin resistant mutants were evolved in isogenic prophage (+) and (-) strains of S.typhimurium and challenged against a gradient of Ciprofloxacin to observe resistance development and extent. These mutants were collected and sequenced to understand mutational diversity at varying antibiotic concentrations and prophage effects on the mutational evolution pathways taken by bacterial hosts. Results yield that although the distribution of Ciprofloxacin mutants remains similar for both strains of S.typhimurium, the overall quantity of mutants evolved is increased in prophage (-) strains. Sequencing results further elucidate the effect which this mobile genetic element has on evolutionary pathways to resistance. Such results further characterize the specific role inducible prophages play in both the direct mechanisms they leverage to interact with antibiotics and the evolutionary mechanisms they exploit to influence bacterial survival and antibiotic resistance.

CRISPR spacer acquisition is a rare event in human gut microbiome

An-Ni Zhang1, Jeffry M Gaston2, Shijie Zhao1, Xiaoqiong Gu3, Eric J Alm1*

1Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, USA;
2 Google, Cambridge, USA;
3 Antimicrobial Resistance Interdisciplinary Research Group, Singapore-MIT Alliance 9 for Research and Technology, 1 Create Way, Singapore 138602, Singapore
* Corresponding author:
Address: Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, USA.

Host-parasite (host-virus) interactions are important for all cellular life, for example, human immunity and SARS-COV2. In bacteria and archaea, CRISPR systems actively acquire spacers to ensure continued defense against phages, which have been reported to acquire a new spacer within a few hours or days in an experimental setting. However, spacer acquisition in natural environments has often been to slow to observe, and limited literature reports suggest a much slower rate (new spacers only acquired over months or years). This high variance highlights the needs to improve our understanding of host-parasite interactions in natural environment.
By investigating temporal WGS datasets and metagenomes of human gut microbiome in healthy individuals, we found that spacer acquisition was a rare event in human gut microbiome, with an average rate of 1 spacer per 2,142-5,000 cell divisions, i.e. over 7-8 years. This low rate reflects only a small proportion of phage (0.4%-0.6% infection rate) challenged the CRISPR systems in human gut microbiome. The rare spacer acquisition in CRISPR suggest that CRISPR might not be the primary risk of effective phage therapy for human microbiome. The results of this study may inform future efforts involving phage therapy and pandemic defense.

Transcriptome Analysis of planktonic and biofilm cultures from several strains of Variovorax

Christina Zhang, Kalea Bringas, Paul Orwin

University of the Pacific

Variovorax is a genus in the Betaproteobacteria that exhibits diverse morphological, physiologic, and metabolic characteristics. This genus is thought to be important in the formation of the rhizosphere community. Previous studies have identified the static biofilm transcriptome of V. paradoxus EPS. Recent work on finishing the genomes of many Variovorax isolates has shown that the genomes are frequently multi-partite. We now wish to examine the roles of genes on these different replicons in biofilm formation, and to evaluate this phenotype more broadly. We evaluated two distinct biofilm growth phases (static and colony biofilms) along with two planktonic growth phases (logarithmic and stationary). A major goal of our study is to understand the common genetic programming that is responsible for changes in gene expression across these different bacterial isolates. Triplicate biological samples were imaged, tracked for growth, and RNA was extracted for the strains 4MfCol3.1, 110B, MF350, NFACC26, ATCC 17713(NBRC 15149), EPS, and VAI-C in 2.5 g/L YE media. 20-30 million paired end reads (2x150bp) were acquired from each sample and aligned to the genome. Prokka annotations were used to identify genes in the genomes, unless they had already been annotated by PGAP. The comparisons were used to identify genes up and down regulated >2-fold when comparing alternative conditions. Based on PCoA clustering, we observed substantial variation between the forms of biofilms, as well as the planktonic cultures. We also observe that very large sets of genes (1000-3000) are altered in expression during biofilm formation in all strains tested. We have also identified genes in all different replicons with differential biofilm regulation, and have identified genes commonly upregulated in biofilms by different strains. This work will form the basis of further comparative studies in biofilm formation across this important genus.

Arrowhead Directory 2022

Kjersti Marie Aagaard
Baylor College of Medicine Houston
One Baylor Plaza
Houston, TX 77030
Email: aagaardt@bcm.edu

Esse Aigbokhan
Washington University in St Louis
4515 McKinley Avenue
St. Louis, MO 63110
Tel: (314) 719-9843
Email: a.esse@wustl.edu

Rosie Algado,
University of Hawaii at Manoa
1000 Pope Road
Honolulu, HI 96822
Tel: (808) 956-0565
E-mail: rosie.alegado@hawaii.edu

Eric J. Alm
Massachusetts Institute of Technology
NE47-379
Cambridge, MA 02139
Tel: (617) 253-2726
Email: ejalm@mit.edu

Nathalie Aoun
University of California, Davis
1 Shields Ave
Davis, CA 95616
Email: naoun@ucdavis.edu

Jenna Archambeau
American University
4400 Massachusetts Avenue, NW
Washington, DC 20016
Tel: (757) 262-9747
Email:
ja7473a@student.american.edu

Eszter Ari
HCEMM-BRC Metabolic Systems Research Group.
Dugonics tér 13.
Szeged 6720 Hungary
Tel: +36 30 405 1683
Email: arieszter@gmail.com

Ben Auch
Phase Genomics, Inc.
4000 Mason Road
Seattle, WA 98195
Email: ben@phasegenomics.com

Naila Barbosa da Costa
University of Montreal
1205 Dr. Penfield Avenue
Montreal, Quebec, Canada H3A 1B1
Tel: +15142438554
Email: nbcosta@gmail.com

Jeffrey E. Barrick,
University of Texas at Austin
2500 Speedway
Austin, TX 78712
Tel: (512) 471-3247
Email: jbarrick@cm.utexas.edu

Michael Baym
Harvard Medical School
Countway 303
10 Shattuck Street
Boston, MA 02115
Email: baym@hms.harvard.edu

Elinne Becket
California State University, San Marcos
333 South Twin Oaks Valley Road
San Marcos, CA 92096-0001
Tel: (760) 750-8239
Email: ebecket@csusm.edu

Claudia Berlinn
Zymo Research Corp.
17062 Murphy Ave
Irvine, CA 92614
Email: cberlinn@zymoresearch.com

Shawal Bhalli
University of California, Los Angeles
1602 Molecular Science Building
Los Angeles, CA 90095
Tel: (310) 825-8924
Email: shawalbhalli@g.ucla.edu

Noelle Bittner
PacBio
1305 O'Brien Drive
Menlo Park, CA 94025
Email: nbittner@pacb.com

Kevin Blake
Washington University in St. Louis
4515 McKinley Avenue
St. Louis, MO 63110
Tel: (518) 441-4376
Email: kevin.blake@wustl.edu

Kate Bowie
Oregon Health & Science University
3181 SW Sam Jackson Park Rd
Mailcode BICC
Portland, OR 97239
Email: bowiek@ohsu.edu

Kalea Bringas
University of the Pacific
3601 Pacific Ave
Stockton, CA 95211
Tel: (818) 642-0331
Email: k_bringas@u.pacific.edu

Ilana Brito
Cornell University
Meinig School of Biomedical
Engineering
Kimball Hall, Room 289
Tel: (607) 254-2938
Email: ibrito@cornell.edu

Eliot Bush
Harvey Mudd College
F.W. Olin Science Center
Room 1240
301 Platt Blvd.
Claremont, CA 91711
Tel: (909) 607.0653
Email: Eliot_Bush@hmc.edu

Gina Chaput
University of California, Davis
One Shields Ave.
Davis, CA, 95616
Email: ginamchaput@gmail.com

Alexander Chase
Southern Methodist University
Department of Earth Sciences
Heroy Hall, P.O. Box 750395
Dallas, Texas 75275-0395
Email: abchase@ucsd.edu

JooHee Choi
Washington University in St. Louis
4515 McKinley Avenue
St. Louis, MO 63110
Tel: (314) 480-2272
Email: joohee.choi@wustl.edu

Sebastian Cocioba
Binomica Labs
Tel: (347) 721- 5754
Email: Sebastian@binomicalabs.org

Nicoletta Commins
Harvard Medical School
Countway Library of Medicine
10 Shattuck Street, Office 307
Boston, MA 02115
Email: nacommins@g.harvard.edu

Jaime Cordova
University of Wisconsin Madison
426G Henry Mall Rm 4434
Genetics/Biotechnology VBuilding
Madison, WI 53706
Email: jaime.cordova@wisc.edu

Aline Curiel
Palomar College
1140 W Mission Rd
San Marcos, CA 92069
Tel: (760) 224-3004
Email:
acurielsilv1514@student.palomar.edu

Annabelle Damerum
Zymo Research Corp.
17062 Murphy Ave
Irvine, CA 92614
Email:
adamerum@zymoresearch.com

Gautam Dantas
Washington University School of Medicine
4515 McKinley Ave, Rm 5314
St Louis, MO 63110
Tel: (314) 362-7238
Email: dantas@wustl.edu

Justine Debelius
Johns Hopkins
615 N. Wolfe Street
Baltimore, MD 21205
Email: justine.debelius@ki.se

Cesar de la Fuente
University of Pennsylvania
3610 Hamilton Walk
Philadelphia, PA 19104
Email: cfuente@upenn.edu

Anna DeVeaux
Washington University in St. Louis
4515 McKinley Ave
St. Louis, MO 63110
Tel (248) 736-3502
Email: a.deveaux@wustl.edu

Trisha Dinh
University of the Pacific
3601 Pacific Ave.
Stockton, CA 95211
Email: t_dinh10@u.pacific.edu

Luke Diorio-Toth
Washington University in St. Louis
4515 McKinley Ave
St. Louis, MO 63110
Email: ldiorio-toth@wustl.edu

Sage Dunham
University of California, Irvine
3315 McGaugh Hall
Irvine, CA, 92697
Tel: (208) 512-3231
Email: sjdunham@uci.edu

Deniz Durmusoglu
NC State University
911 Partners Way
Raleigh, NC 27695
Email: ddurmus@ncsu.edu

Rachel J. Dutton
University of California, San Diego
AP&M Building, room 3872
9500 Gilman Dr.
La Jolla, CA 92093-0116
Tel: (858) 822-5359
Email: rjdutton@ucsd.edu

Ashlee Earl
Broad of Institute of MIT and Harvard
415 Main Street
Cambridge, MA 02142
Tel: (617) 714-7927
Email: aearl@broadinstitute.org

Carrie Ann Eckert
Oak Ridge National Laboratory
PO Box 2008, MS6060
Oak Ridge, TN 37831-6060
Tel: 303-726-6337
Email: eckertca@ornl.gov

Jonathan Eisen
UC Davis Genome Center
Davis, CA 95616
Tel: (510) 752-3498
Email: jaeisen@ucdavis.edu

Joanne B. Emerson
University of California, Davis
One Shields Ave.
Davis, CA 95616
Tel: (530) 752-9848
Email: jbemerson@ucdavis.edu

Cassie Ettinger
University of California, Riverside
900 University Ave
Riverside, CA 92521
Tel: (510) 982- 9212
Email: cassande@ucr.edu

Anne Farrell
Dartmouth College
Hanover, NH 03755
Email:
anne.a.farrell.GR@dartmouth.edu

Aura Ferreiro
Washington University in St. Louis
4515 McKinley Avenue
St. Louis, MO 63110
Tel: (530) 220-0841
Email: alferreiro@wustl.edu

Skye Fishbein
Washington University in St. Louis
4515 McKinley Avenue
St. Louis, MO 63110
Email: skye.f@wustl.edu

Jaeden Flury
Cal State San Marcos
333 South Twin Oaks Valley Road
San Marcos, CA 92096-0001
Tel: (760) 750-8239
Email: flury001@cougars.csusm.edu

Anthony Fodor
UNC Charlotte
9201 University City Blvd
Charlotte, NC 28223
Tel: (704) 687-8214
Email: afodor@uncc.edu

Kevin Garcia
Cal State San Marcos
333 South Twin Oaks Valley Road
San Marcos, CA 92096-0001
Tel: (760) 750-8239
Email:
garci1904@cougars.csusm.edu

Jeffry Gaston
MIT
NE47-MIT
Cambridge, MA 02139, USA
Email: gastoj3@gmail.com

Renae Geier
Armata Pharmaceuticals
4503 Glencoe Ave
Marina del Rey, California, 90292,
Tel: (608) 498-6383
Email: rgeier@armatapharma.com

Sonia Ghose
University of California, Davis
One Shields Ave.
Davis, CA 95616
Tel: (510) 816-2706
Email: slghose@ucdavis.edu

Michael Gilmore
Harvard Medical School
243 Charles Street
Boston MA 02114
Tel: (617) 573-3845
Email:
Michael_gilmore@meei.harvard.edu

Tijana Glavina del Rio
DOE Joint Genome Institute
Lawrence Berkeley National
Laboratory
Tel: (925) 296-5849
Email: TGlavina@lbl.gov

Indra Gonzalez Ojeda
Harvard Medical School
10 Shattuck Street
Boston, MA 02115
Tel: (787) 412- 8042
Email:
igonzalezojeda@g.harvard.edu

Olivia Gorushi
Washington University in St. Louis
4515 McKinley Avenue
St. Louis, MO 63110
Tel: 520-954-5875
Email:Olivia.g@wustl.edu

David Greenberg
Miroculus
458 Brannan St.
San Francisco, CA 94107
Tel: (831) 332-7511
Email: david@miroculus.com

Alexandra Grote
Broad of Institute of MIT and Harvard
415 Main Street
Cambridge, MA 02142
Email: agrote@broadinstitute.org

Miranda Harmon-Smith
Joint Genome Institute
2800 Mitchell Drive
Walnut Creek, CA 94598
Tel: (925) 997-2796
Email: mlharmon-smith@lbl.gov

Chris Herring
Enchi Corporation
Tel: (603) 276-0221
Email: chrisherringfish@gmail.com

Jared Holms
Indian Health Council
50100 Golsh Rd,
Valley Center, CA 92082
Tel: (760) 749-1410
Email: rrodriguez@indianhealth.com

Ping Hu
Procter & Gamble
1 P&G Plaza Cincinnati, OH 45202
Tel: (513) 349- 4298
Email: hu.p@pg.com

Brad Hunter
Cal State University, San Marcos
333 South Twin Oaks Valley Road
San Marcos, CA 92096-0001
Tel: (760) 750-8239
Email:
hunte051@cougars.csusm.edu

Carolyn Ibberson
University of Oklahoma
770 Van Vleet Oval
Norman, OK 73019
Tel: (405) 325- 4321
Email: carolyn.ibberson@ou.edu

Josh Jones
Cornell University
Tel: (323) 333- 9461
Email: jcj79@cornell.edu

Sandeep J. Joseph
CDC
Tel: (404) 578-6790
Email: lww9@cdc.gov

Guillaume Jospin
Animalbiome
400 29th Street Suite 502
Oakland, CA 94609
Email: guillaume@animalbiome.com

Lisa Karstens
Oregon Health & Science University
3181 SW Sam Jackson Park Rd
Mailcode BICC
Portland, OR 97239
Tel: (503) 418-8877
Email: karstens@ohsu

Hossein Khadivar
California State University San Marcos
333 South Twin Oaks Valley Road
San Marcos, CA 92096-0001
Tel: (760) 650-5502
Email: khadi002@cougars.csusm.edu

Freeman Lan
University of Wisconsin Madison
433 Babcock Drive
Madison, WI 53706
Tel: (858) 205-4897
Email: flan3@wisc.edu

Mackenzie Lee
University of California, Los Angeles
1602 Molecular Science Building
Los Angeles, CA 90095
Tel: (310) 825-8924
Email: mackenzielee01@g.ucla.edu

Katherine Paige Lemon,
Baylor College of Medicine
One Baylor Plaza
Houston, TX 77030
Email: Katherine.Lemon@bcm.edu

Asaf Levy
The Hebrew University of Jerusalem
P.O. Box 12
Rehovot 76100 Israel
Tel: +972-(0) 8-9489596
Email: alevy@mail.huji.ac.il

Gina Lewin
Case Western Reserve University
10900 Euclid Ave.
Cleveland, OH 44106
Tel: (609) 658-1301
Email: glewin3@gatech.edu

Ivan Liachko
Phase Genomics, Inc.
4000 Mason Road
Seattle, WA 98195
Tel: (206) 384-6228
Email: ivan@phasegenomics.com

Tami Lieberman
Massachusetts Institute of Technology
77 Massachusetts Ave.
Cambridge, MA 02139
Tel: (617) 253- 9757
Email: tami@mit.edu

Anurag Limdi
Harvard University
200 Longwood Ave,
Warren Alpert 112
Boston, MA 02115
Tel: (857) 999-5236
Email: alimdi@g.harvard.edu

Annie Lin
University of the Pacific
3601 Pacific Ave.
Stockton, CA 95211
Tel: (415) 794- 2269
Email: a_lin23@u.pacific.edu

Xiaoxia (Nina) Lin
University of Michigan
2800 Plymouth Road
Ann Arbor, MI 48109
Tel: (734) 647-8026
Email: ninalin@umich.edu

Kris Locken
Zymo Research Corp.
17062 Murphy Ave
Irvine, CA 92614
Email: klocken@zymoresearch.com

Allison J. Lopatkin
University of Rochester
Chemical Engineering
4301 Wegmans Hall
Rochester, NY 14629
Email: Allison.lopatkin@rochester.edu

Tiffany Lowe-Power
University of California, Davis
282 Hutchison Hall
One Shields Ave
Davis, CA 95616
Tel: (530) 754-1212
Email: tlowepower@ucdavis.edu

Bejan Mahmud
Washington University School of Medicine, St. Louis
4515 McKinley Avenue
St. Louis, MO 63110
Email: bmahmud@wustl.edu

Jeffrey H. Miller
University of California, Los Angeles
1602 Molecular Science Building
Los Angeles, CA 90095
Tel: (310) 825-8924
Email: jhmiller@microbio.ucla.edu

Hosein Mohimani
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
Tel: (412) 268- 6326
Email: hoseinm@andrew.cmu.edu

Grace Morales
Vanderbilt University Medical Center
1211 Medical Center Drive
Nashville, TN 37232
Email:
grace.h.morales@vanderbilt.edu

Gabriel Moreno-Hagelsieb
Wilfrid Laurier University
75 University Avenue West
Waterloo, Ontario, Canada N2L 3C5
Tel: (519) 884- 1970
Email: gmorenohagelsieb@wlu.ca

Aindrilla Mukhopadhyay
Lawrence Berkeley Laboratory
1 Cyclotron Road
Berkeley CA 94720
Tel: (510) 495-2628
Email: AMukhopadhyay@lbl.gov

Elizabeth Murguia
Cal State San Marcos
333 South Twin Oaks Valley Road
San Marcos, CA 92096-0001
Tel: (760) 750-8239
Email:
murgu014@cougars.csusm.edu

Erin Newcomer
Washington University in St. Louis
4515 McKinley Avenue
St. Louis, MO 63110
Tel: (636) 236-4168
Email: erin.newcomer@wustl.edu

Cecilia Noecker
University of California San Francisco
Tel: (612) 219-4054
Email: cecilia.noecker@ucsf.edu

Trent Northen
Lawrence Berkeley Laboratory
1 Cyclotron Road
Berkeley, CA 94720.
Phone: (925) 927-2827
Email: trnorthen@lbl.gov

Andrew Oliver
United States Department of Agriculture
Tel: (805) 857-1805
Email: andrew.oliver@usda.gov

Matthew Olm
Stanford University
Fairchild Science Building
Rm D307a
299 Campus Drive
Stanford, CA 94305-5124
Tel: (480) 213 – 6684
Email: mattolm@stanford.edu

Paul Orwin
University of the Pacific
3601 Pacific Ave.
Stockton, CA 95211
Tel: (909) 908-4220
Email: porwin@pacific.edu

Dave Pacheco
Invitae
Tel: (949) 439-3276
Email: dave.pacheco@invitae.com

Leka Papazisi
ATTC
10801 University Blvd
Manassas, Virginia, 20110
Email: lpapazisi@atcc.org

Julissa Perez-Marron
University of the Pacific
3601 Pacific Ave.
Stockton, CA 95211
Tel: (909) 659 8767
Email: j_perezmarron@u.pacific.edu

Inmaculada Pérez-Prieto
University of Granada
Avenida de la Fuente Nueva S/N
18071 Granada Spain
Email: nmapprieto@ugr.es

Giorgia Pirino
University of California, San Diego
9500 Gilman Drive #
La Jolla CA 92093
Tel: (949) 3949636
Email: gpirino@ucsd.edu

Daniel Portik
PacBio
1305 O'Brien Drive
Menlo Park, CA 94025
Email:
dportik@pacificbiosciences.com

Janira Prichula
Harvard Medical School
243 Charles Street
Boston MA 02114
Tel: (617) 994- 87043
Email: jprichula@meei.harvard.edu

Jerome Prusa
Washington University in St. Louis
4515 McKinley Avenue
St. Louis, MO 63110
Tel: (402) 669-4298
Email: prusa@wustl.edu

Elisabeth Raleigh
New England Biolabs
240 County Rd
Ipswich, MA 01938
Tel: (978) 380-7238
Email: Raleigh@neb.com

Eleanor Rand
Harvard Medical School
10 Shattuck Street
Boston, MA 02115
Email: eleanor_rand@g.harvard.edu

Connie Rojas
University of California, Davis
451 Health Sciences Drive
Davis, CA 95616
Email: carojas@ucdavis.edu

Fernando Rossine
Harvard Medical School
10 Shattuck Street
Boston, MA 02115
Email:
fernando_rossine@hms.harvard.edu

Jason Rothman
University of California, Irvine
3315 McGaugh Hall
Irvine, CA, 92697
Email: rothmanj@uci.edu

Jian Ryou
Washington University in St. Louis
4515 McKinley Avenue
St. Louis, MO 63110
Tel: (314) 224-2615
Email: j.ryou@wustl.edu

Sanjam Sawhney
Washington University in St. Louis
4515 McKinley Avenue
St. Louis, MO 63110
Email: ssawhney@wustl.edu

Suelen Scarpa de Mello
Harvard Medical School
243 Charles Street
Boston MA 02114
Tel: (857) 701- 2015
Email:
suelen_scarpademello@meei.harvard.edu

Rebecca Schomer
University of California, Davis
One Shields Ave
Davis, CA 95616
Email: raschomer@ucdavis.edu

Drew Schwartz
Washington University School of Medicine
4515 McKinley Avenue
St. Louis, MO 63110
Tel: (919) 452-8813
Email: schwartzd@wustl.edu

Bruna Sgardioli-Gilmore
University of the Azores
Angra do Herismo
Terceira, Portugal
Email: brunasgardioli@gmail.com

John Sherman
Zymo Research Corp.
17062 Murphy Ave
Irvine, CA 92614
Email:
jsherman@zymoresearch.com

Rebekah Silva
New England Biolabs
240 County Rd
Ipswich, MA 01938
Tel: (951) 543-7620
Email: rsilva@neb.com

Celia Souque
Harvard Medical School
10 Shattuck Street
Boston, MA 02115
Tel: (617) 201- 3045
Email:
celia_souque@hms.harvard.edu

Jason Stajich
University of California, Riverside
900 University Ave
Riverside, CA 92521
Tel: (951) 827-2363
Email: jason.stajich@ucr.edu

Madison Stellfox
University of Pittsburgh
1218 Scaife Hall
3550 Terrace Street
Pittsburgh, PA 15261
Tel: (412) 983-2005
Email: stellfoxme@upmc.edu

Cameron Thrash
University of Southern California
Department of Biological Sciences
3616 Trousdale Pkwy, AHF 209
Los Angeles, CA 90089
Email: thrash@usc.edu

Paul Turner
Yale School of Medicine
165 Prospect Street, PO Box 208106
New Haven, CT, 06520-8106
Tel: (203) 432 –5918
Email: paul.turner@yale.edu

Camilla Urbaniak
NASA Jet Propulsion Laboratory
Pasadena, CA 91109, US
Tel: (626) 840- 6992
Email:
camilla.urbaniak@jpl.nasa.gov

Lucas Van Dijk
Broad Institute of MIT and Harvard
415 Main Street
Cambridge, MA 02142
Email: lvandijk@broadinstitute.org

Daria Van Tyne
University of Pittsburgh
3550 Terrace Street
Pittsburgh, PA 15261
Tel: (412) 648-4210
Email: VANTYNE@pitt.edu

Rhiannon Vargas
Washington University in St. Louis,
4515 McKinley Ave
St. Louis, MO 63110
Email: rcvargas@wustl.edu

Samantha Vogel
Armata Pharmaceuticals
4503 Glencoe Ave, Marina del Rey,
California, 90292, United States
Tel: (661) 433-9308
Email: svogel@armatapharma.com

Jean-Marie Volland
Joint Genome Institute
2800 Mitchell Drive
Walnut Creek, CA 94598
Email: JVolland@lbl.gov

Christine Vu
University of California, Los Angeles
1602 Molecular Science Building
Los Angeles, CA 90095
Tel: (310) 825-8924
Email: christinevusc@gmail.com

Bruce Walker
MIT Ragon Institute
400 Technology Square
Cambridge, MA 02139
Tel: (857) 268-7073
Email: walker@mit.edu

Harris S. Wang
Columbia University Medical Center
3960 Broadway
New York, NY 10032
Tel: (212) 305-1697
Email: hw2429@columbia.edu

Yan Wang
University of California, Los Angeles
650 Charles E. Young Dr
Los Angeles, CA 90095
Tel: 3108258089
Email: wangyan@g.ucla.edu

Peter R. Weigele
New England Biolabs
240 County Rd
Ipswich, MA 01938
Tel: (978) 380 7304
Email: Weigele@neb.com

Alex Westman
PacBio
1305 O'Brien Drive
Menlo Park, CA 94025
Tel: (916) 943-6220
Email: awestman@pacb.com

Katrine Whiteson
University of California, Irvine
3236 McGaugh Hall
Irvine, CA 92697
Tel: (949) 824-9032
Email: katrine@uci.edu

Jeremy Wilkinson
PacBio
1305 O'Brien Drive
Menlo Park, CA 94025
Tel: (617) 334-5353
Email: jwilkinson@pacb.com

Naomi Wilson
Washington University School of Medicine in St. Louis
660 S. Euclid Ave.
St. Louis, Mo 63110
Tel: (585) 748-7477
Email: ngwilson@wustl.edu

Joseph Wirth
Harvey Mudd
1250 N. Dartmouth Ave.
Claremont, CA 91711
Tel: (909) 607- 5863
Email: jwirth@g.hmc.edu

Tanja Woyke
DOE Joint Genome Institute
2800 Mitchell Drive
Walnut Creek, CA 94598
Tel: (925) 296-5839
Email: twoyke@lbl.gov

Xue Yao-Peng
Washington University in St. Louis
4515 McKinley Avenue
St. Louis, MO 63110
Email: x.yaopeng@wustl.edu

David Yarmosh
American Type Culture Collection (ATCC)
10801 University Blvd
Manassas, Virginia, 20110
Manassas, Virginia, USA
Email: dyarmosh@atcc.org

Sosie Yorki
Broad Institute of MIT and Harvard
415 Main Street
Cambridge, MA 02142
Tel: (713) 382-7970
Email: syorki@broadinstitute.org

Madison Youngblom
University of Wisconsin-Madison
5325 Microbial Sciences Building
Tel: (608) 262-6167
Email: youngblom@wisc.edu

Amy Zamora
Harvard University
200 Longwood Ave
Boston, MA 02115
Tel: (484) 538-9148
Email: azamora@g.harvard.edu

An-Ni Zhang
MIT
NE47-MIT
Cambridge, MA 02139, USA
Email: anniz44@mit.edu

Christina Zhang
University of the Pacific
3601 Pacific Ave.
Stockton, CA 95211
Tel: (209) 751-7733
Email: c_zhang13@u.pacific.edu