THE 12TH ANNUAL

DOE Joint Genome Institute

Genomics of Energy & Environment Meeting

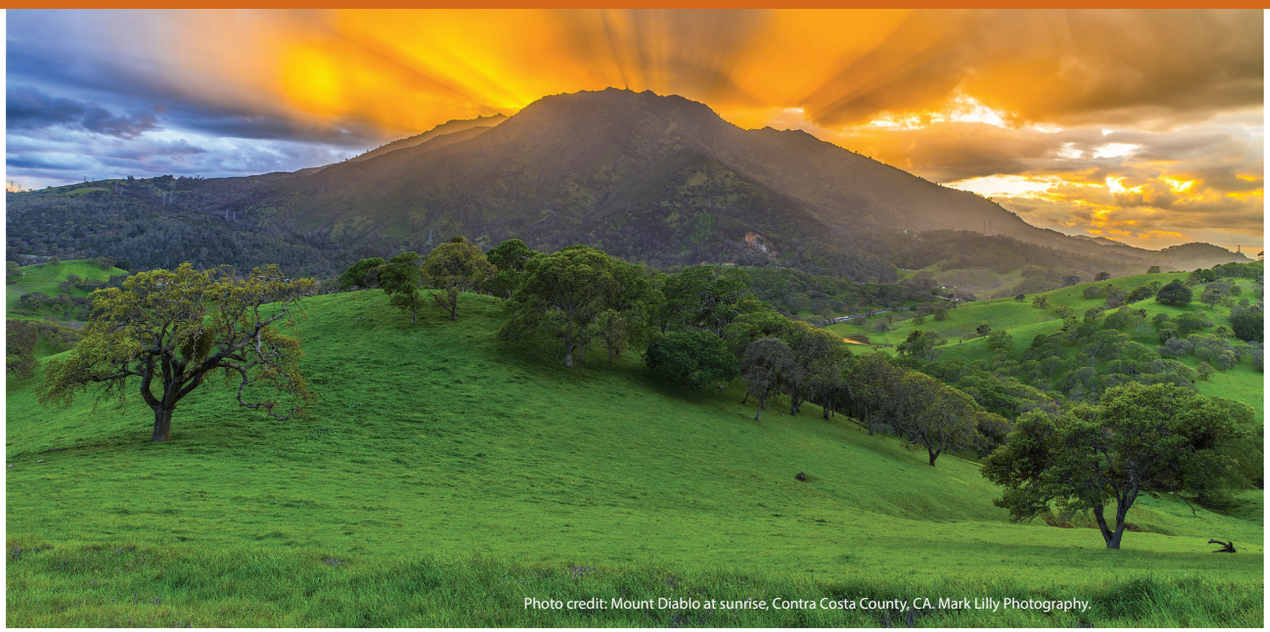

Photo credit: Mount Diablo at sunrise, Contra Costa County, CA. Mark Lilly Photography.

User Meeting Abstracts

March 20-23, 2017 | Walnut Creek Marriott | Walnut Creek, CA

JGI JOINT GENOME INSTITUTE
A DOE OFFICE OF SCIENCE USER FACILITY

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY
Office of Science

All information current as of March 16, 2017

JGI Contact: **Denise Yadon**
DOE Joint Genome Institute
dyadon@lbl.gov

The Joint Genome Institute is a user facility of
the Department of Energy Office of Science

DOE Joint Genome Institute: **www.jgi.doe.gov**
DOE Office of Science: **science.energy.gov**

Table of Contents

Speaker Presentations 1

Poster Presentations 10

Author Index 103

Attendees 110

Speaker Presentations

Abstracts alphabetical by speaker*

Unconscious Bias and Fungi: Studying Microorganisms in a Macro-Focused World

Adams, Cat* (catadams@berkeley.edu)

University of California, Berkeley, Plant and Microbial Biology, Berkeley, CA.

The scientific method is an organized process for experimentation, used to explore observations and answer questions about the natural world. This method of inquiry aims to be as objective as possible, but the humans conducting research are not immune to bias. Within scientific culture, stereotypes about what a 'typical' scientist looks like can result in unconscious social biases. Such social bias has been broadly demonstrated to affect diversity in STEM, which is dominated by white men relative to the general population. Fortunately, subsequent research has developed empirical methods to reliably reduce bias long-term. My organization, the Unconscious Bias Project, promotes both awareness of unconscious bias and its role in STEM diversity, and to provide evidence-based approaches to reduce bias and its effects.

Unconscious social bias is not the only bias plaguing scientists. Cognitive biases affect not only how we conduct scientific research, but how our research is received by both citizens and fellow researchers in different fields. I will briefly discuss unintentional cognitive biases, such as the continued influence effect, that may contribute to the relative obscurity of mycology compared to other biological disciplines. I will then revisit known solutions to reduce unconscious social bias with approaches to control cognitive bias. I will conclude with a list of suggested approaches for mycologists to prevent the formation of bias against fungi, and to reduce implicit bias against fungi and their study among our biological peers.

Whole-genus Association Analysis—Using Hundreds of *Aspergillus* Genomes for Linking Phenotype to Genotypes

Andersen, Mikael* (mr@bio.dtu.dk)

Technical University of Denmark, Lyngby, Denmark.

In the *Aspergillus* whole-genus sequencing project, we are sequencing a member of all ~350 species in the genus, focusing on type strains and/or the most representative isolates of the species. The goals of this project include the study of microbial speciation, diversity of primary metabolism, mapping of biofuel-relevant CAZymes, and elucidation of bioactive secondary metabolites. Currently, we have sequenced the genomes of 163 species, 61 are being sequenced, and 118 are being prepared for sequencing.

In this talk, I will give an overview of the status and goals of the project. Furthermore, I will present how we are analyzing the genomes, and show preliminary results from our analyses. In particular, I will present how we use the principles of association analysis – as known from human genetics – at a multi-species level to identify the genes responsible for specific phenotypes.

Genetic Networks: Systematic Mapping of Gene-Gene and Gene-Environment Interactions

Andrews, Brenda* (brenda.andrews@utoronto.ca) (KEYNOTE)

University of Toronto, Ontario, Canada.

How Has Nature Evolved the Enzymes Required for Life

Babbitt, Patsy* (babbitt@cgl.ucsf.edu), Professor

Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco.

The natural catalytic repertoire includes hundreds of functionally diverse enzyme superfamilies, each comprised of thousands of sequences and many different chemical reactions. Studies of some of these superfamilies suggest that they have evolved using "privileged" scaffolds, structural templates whose structural folds and active site architectures facilitate catalysis of partial reactions or other chemical capabilities common to many different homologous proteins. Evolutionary divergence of the ancestral forms of these scaffolds are thought to have lead to the variations in the topology, active site details, and reaction and substrate specificities used by contemporary superfamily members to catalyze varied chemical reactions. Using sequence similarity networks, we have computationally investigated on a global scale structure-function relationships in several of these superfamilies. Mapping functional and other types of information to similarity networks comprised of many thousands of sequences enables a large-scale view of functional trends from the context of their sequence and structural similarities. We describe what we have learned from investigation of a several of these superfamilies and provide examples for how the results can be applied to guide the choice of informative targets for biochemical and structural characterization, for prediction of functional features in unknowns, and to provide insight for enzyme engineering in the lab.

Notes from the Datapocalypse

Brown, C. Titus* (ctbrown@ucdavis.edu)

University of California, Davis, CA.

Since 2008, I've been struggling to analyze large JGI-generated sequencing data sets. My interactions with these data sets have informed my research direction and goals as much as anything else, and I'd like to share some lessons learned (even though I don't yet have solutions to some of the problems).

I'll talk about the problems of (effectively) infinite data, the challenges of analyzing organisms for which no reference exists, and some of the nifty tools that we're developing at the moment.

EvoNet: A Phylogenomic and Systems Biology Approach to Identify Genes Underlying Plant Survival in Marginal, Low-Nitrogen Soils

Coruzzi, Gloria* (gloria.coruzzi@nyu.edu)[1]; Gutiérrez, Rodrigo[2]; Varala, Kranthi[1,3], Stevenson, Dennis[4], DeSalle, Rob[5], McCombie, Richard W.[6], Ané, Jean-Michel[7;] Kaeppler, Heidi[7]

[1]New York University, NY. [2]Pontifica Universidad Católica de Chile, Santiago, Chile. [3]Purdue University, Indiana. [4]New York Botanical Garden, NY. [5]American Museum of Natural History, NY. [6]Cold Spring Harbor Laboratory, NY. [7]University of Wisconsin, Madison WI.

"*EvoNet*" is a DOE-BER sustainability project that combines phylogenomic and network analyses to identify genes underlying plant adaptation to growth in marginal, low-N (nitrogen) soils. Our approach exploits the genomes of "extreme survivor" plants adapted to thrive in marginal, N-poor soils on the edge of the Atacama Desert of the Chilean Andes. It uses a phylogenomic pipeline we previously validated [1, 2], and a "paired species" sampling strategy, to identify genes that distinguish these "extreme survivors" in Chile from their related species in California (CA) that are not constrained by N. These "paired species" samplings represent multiple independent origins of the traits (adaptation to low-N and drought), and therefore offer diverse genomic backgrounds within which the "survival traits" repeatedly arose. The Chilean team – headed by R. Gutierrez – performed RNA-seq analysis of 32 "extreme survivor" species collected adjacent to the Atacama Desert. These 32 Chilean "survivors" are herbaceous species that span the main branches in flowering plants and include seven grass species of particular interest for biofuels. The RNA-seq data cover 80-90% of the genes in these genomes (judged by *BUSCO* analysis) and comprise ~20-30K genes/species. We next compared these Chilean "survivor" genomes to their closest related CA "sister species" for which deep-transcriptome is available from public sources including the 1KP project, using phylogenomic analysis. To construct this largest-ever phylogenomic tree, we enhanced the throughput of a phylogenomic pipeline we previously developed [1], now renamed "PhyloGeneious." The resulting phylogeny of 70+ genomes (20-30k orthologs/species) - built from ~1.7M input genes - comprises ~32,000 ortholog sets and ~2 million parsimony informative characters. This phylogenomic tree identified genes that provide Partition Branch Support (PBS) at 19 evolutionarily independent origins of low-N/Drought adaptation. 1,157 of these genes provide recurrent PBS support (>=3 independent origins) for the divergence of low-N adapted Chilean species from their available sister CA species. This set of 1,1,57 PBS genes is enriched in the processes Nitrogen compound metabolism, Photosynthesis and Energy (GO term analysis FDR <0.01). Crucially, this identified 19 transcription factors that provide recurring PBS support for low-N adaptation. We are functionally validating their role as master regulatory genes controlling NUE and drought-tolerance networks through CRISPER transgenic lines in the model monocot *Brachypodium*.

1. Lee E et. al. (2011), A functional phylogenomics view of the seed plants. *PLoS Genet* 7(12):e1002411.
2. Delaux et. al. (2014), Comparative phylogenomics uncovers the impact of symbiotic associations on host genome evolution. *PloS Genet* 10(7):e1004487.

The EvoNet grant DOE-BER DE-SC0014377 is supported by the Genomics Science program within the Office of Biological and Environment Research in the DOE Office of Science.

http://evonet.org

CETCH Me if You Can: A Synthetic Pathway for the Fixation of Carbon Dioxide

Erb, Tobias J.* (toerb@mpi-marburg.mpg.de) (KEYNOTE)

Max Planck Institute for Terrestrial Microbiology, D-35043 Marburg, Germany.

Carbon dioxide (CO_2) is a potent greenhouse gas that is a critical factor in global warming. At the same time atmospheric CO_2 is a cheap and readily carbon source. Yet, synthetic chemistry lacks suitable catalysts to functionalize the CO_2-molecule, emphasizing the need to understand and exploit the CO_2-mechanisms offered by Nature.

In my talk I will (1) discuss the evolution and limitation of naturally existing CO_2-fixing enzymes and pathways. I will (2) present strategies for the engineering and design of artificial CO_2-fixation reactions and pathways (*Peter et al., ACIE 2015*), and (3) outline how artificial pathways can be experimentally realized and further optimized (*Schwander et al., Science 2016*).

An example for such a synthetic CO_2-fixation pathway module is the CETCH cycle. The CETCH cycle is an *in vitro*-reaction network of 17 enzymes that was established with enzymes originating from nine different organisms of all three domains of life and optimized in several rounds by enzyme engineering and metabolic proofreading. In its version 5.4, the CETCH cycle converts CO_2 into organic molecules at a rate of 5 nanomoles of CO_2 per minute per milligram of protein. This is slightly faster than the Calvin cycle under comparable conditions and notably at 20% less energy per CO_2 fixed.

References:

Peter D, Schada von Borzyskowski L, Kiefer P, Christen P, Vorholt JA, Erb, TJ (2015) Screening and engineering the synthetic potential of carboxylating reductases from central metabolism and polyketide biosynthesis. *Angew Chem* 45:13457-61.

Schwander T, Schada von Borzyskowski L, Burgener S, Cortina NS, Erb TJ (2016) A synthetic pathway for the fixation of carbon dioxide *in vitro*. *Science* 354:900-4.

Manganese Oxidation and Mineralization by Ascomycete Fungi

Hansel, Colleen* (chansel@whoi.edu)

Woods Hole Oceanographic Institution.

Manganese (Mn) oxides are among the strongest sorbents and oxidants within the environment, controlling the fate and transport of numerous elements and the degradation of recalcitrant carbon. Both bacteria and fungi mediate the oxidation of Mn(II) to Mn(III/IV) oxides within various terrestrial and marine systems. While white-rot Basidiomycete fungi oxidize Mn(II) using laccases and manganese peroxidases in association with lignocellulose degradation, the mechanisms by which filamentous Ascomycete fungi oxidize Mn(II) and a physiological role for Mn(II) oxidation in these organisms remains poorly understood. Our previous work of a group of Ascomycete fungi identified transmembrane NADPH oxidases in the formation of Mn oxides via the extracellular production of the Mn(II) oxidant and reactive oxygen species superoxide. Here, using a combination of chemical and in-gel assays, bulk mass spectrometry, and iTRAQ proteomics, we demonstrate secretome-based

Mn(II) oxidation in three phylogenetically diverse Ascomycetes that is mechanistically distinct from the hyphal-associated Mn(II) oxidation processes. We show that the Mn(II) oxidative capacity of these fungi is dictated by species-specific secreted enzymes and varies with secretome age, and we observe direct, enzymatic Mn(II) oxidation in native PAGE gels in the absence of added reductants. Specifically, we identify the putative Mn(II)-oxidizing enzymes as bilirubin oxidase in *Stagonospora* sp. SRC1lsM3a, an FAD-binding oxidoreductase in *Pyrenochaeta* sp. DS3sAY3a, and GMC oxidoreductase in *Paraconiothyrium sporulosum* AP3s5-JAC2a. This study reveals the concurrence of hyphal- and secretome-based processes in fungal Mn(II) oxidation, expands the diversity of oxidoreductases involved in biogenic Mn(II) oxidation, and suggests that the ability of fungal secretomes to oxidize Mn(II) may be more widespread than previously thought.

Pieces of the Phytobiome: Multitrophic and Environmental Influences on Plant Health

Leach, Jan E.* (Jan.Leach@colostate.edu)

Colorado State University, Ft. Collins, CO.

Phenotypic responses of plants to biotic and abiotic stresses are frequently studied as the outcome of interactions between plants and one or two species of microbes or a single abiotic stress. However, in the phytobiome, plant health and productivity is impacted by simultaneous interactions among multiple organisms and the environment, and frequently, the responses to these interactions are distinct and would not be predicted from studying less complex systems. In this presentation, examples of multitrophic interactions among bacteria, insects and plants as well as interactions involving bacteria, plants and temperature will be used to demonstrate the difficulty in predicting outcomes from simultaneous stresses. Developing successful and sustainable crop improvement and management strategies for the future will benefit from studying these interactions as a system.

EMSL Capabilities for Functional Genomics of Plants and Microbes

Lipton, Mary S.* (mary.lipton@pnnl.gov)

Environmental Molecular Sciences Laboratory (EMSL), Pacific Northwest National Laboratory (PNNL), Richland, WA.

Room at the Bottom: Opportunities and Challenges in the Subsoil

Lynch, Jonathan* (jpl4@psu.edu)

Penn State University, University Park, PA 16802.

Greater exploitation of subsoil resources by annual crops would afford multiple benefits, including greater water and N acquisition in most agroecosystems, and greater sequestration of atmospheric C.

Multiple root phenes under genetic control are associated with deeper roots, opening the possibility of breeding annual crops with root traits improving subsoil exploration. The 'Steep, cheap, and deep' ideotype for deeper rooting in maize, cintegrating architectural and anatomical root phenes and phene states, appears to be a useful paradigm. Comparable ideotypes are being developed for beans, rice and barley. Functional-structural modeling has been invaluable in analyzing and understanding the fitness landscape of root phenotypes. These phenes and issues merit research because of their potential value in developing more productive, sustainable, and resilient agricultural systems.

Community Genomes Uncover Pattern and Process in Freshwater Bacterial Ecology and Evolution

McMahon, Katherine (Trina)* (trina.mcmahon@wisc.edu)

University of Wisconsin-Madison, Madison, WI 53706, Departments of Bacteriology, and Civil and Environmental Engineering.

Freshwater bacterial communities underpin all biogeochemical cycles in lakes and are outstanding model systems with which to study ecological and evolutionary processes in microbes. We are partnered with the North Temperate Lakes Long Term Ecological Research site and have access to a rich time series of microbial community data extending back more than a decade. In this talk I will compare patterns of microbial community assembly and dynamics in two sharply contrasting lakes using 16S rRNA gene tag sequencing, explore population structure using single cell genomics and metagenomics, and propose a new genomics-based framework for thinking about traits within complex microbial communities.

The Evolution and Significance of Diverged Alleles in Diatoms

Mock, Thomas* (T.Mock@uea.ac.uk)

University of East Anglia, Norfolk, UK.

Natural selection is the hallmark of evolution. However, inference of selection in algal genomes has been a challenge but is key to provide unequivocal evidence of adaptive evolution under given environmental conditions. Here we explore the role of selection for the evolution of diatom genomes under variable environmental conditions. Diatoms often outcompete other phytoplankton species in variable environments, which indicates a significant level of adaptation to frequently changing environmental conditions. However, how selection has acted on the evolution of diatom genomes with respect to ecosystem variability has not been extensively studied yet but is important to understand why diatoms have become a dominant force in variable marine environments, which are considered the most productive marine ecosystems on Earth. To address that question, we sequenced a diatom genome from the Southern Ocean, which is considered to be a highly variable ocean due to strong seasonality in light, temperature and nutrients. Our study reveals that the genome of *Fragilariopsis cylindrus* contains highly diverged alleles that are differentially expressed depending on the environmental conditions and stresses imposed. Alleles with largest ratio of replacement over silent substitutions (largest dN/dS ratio) show the most pronounced condition-

dependent expression. This suggests that environmentally-induced diversifying selection drives the allelic differentiation. The highly diverged alleles with nucleotide divergence of up to 6% show nevertheless a signature of genetic recombination. Many of the diverged alleles encode proteins from conserved core and lineage-specific metabolism indicating the requirement to fundamentally adjust metabolism to cope with an extreme and variable environment. Homologs of diverged alleles account for 60% of all F. *cylindrus*-specific transcripts in natural communities, including the most highly abundant transcripts. Diverged alleles adapted to particular conditions are maintained in a vast gene pool and enable the population to respond to the highly variable environment of the surface Southern Ocean.

Genomics of Fungal-Bacterial Interactions

Pawlowska, Teresa E.* (tep8@cornell.edu)

Cornell University

Symbioses between fungi and bacteria are increasingly appreciated for their ubiquity as well as ecological and practical significance. Thus far, the majority of research efforts have been focused on symbioses of Mucoromycota with endosymbiotic beta-proteobacteria and Mollicutes, which are starting to yield novel insights into the mechanisms of partner co-evolution and symbiosis establishment. For example, we found that genome reduction in the '*Candidatus* Glomeribacter gigasporarum' nonessential mutualist of arbuscular mycorrhizal fungi (AMF, Glomeromycotina) is a non-degenerative process, representing a departure from degenerative genome contraction expected in heritable endobacteria. Similarly, another endosymbiont of AMF, '*Candidatus* Moeniiplasma glomeromycotorum', whose lifestyle remains unknown, appears to retain mechanisms responsible for recombination and genome plasticity, a pattern unexpected in heritable endobacteria. While the AMF-bacterial symbioses defy evolutionary expectations, they are recalcitrant to experimental manipulations aimed at unraveling the mechanisms of symbiosis establishment and maintenance. To gain such mechanistic insights, we focus on the heritable symbiosis between *Rhizopus microsporus* (*Rm*, Mucoromycotina) and *Burkholderia* endobacteria. In this study system, host isolates of *Rm* form a mutualism with their endobacteria, whereas the non-host isolates interact antagonistically with *Burkholderia* isolated from the host. Through transcriptional profiling and experimental manipulations of the partners, we inferred that bacteria use the same set of symbiosis factors, both known and novel, to engage the host and non-host. In contrast, host and non-host responses to bacteria differ substantially, with differences centered on cell wall remodeling and reactive oxygen species metabolism. Mutualism establishment in the host is mediated by changes in several signaling pathways, lipid metabolism, and cytoskeleton rearrangements. In addition to insights into the mechanisms of symbiosis establishment and avoidance, the *Rm-Burkholderia* system is an excellent model to study the evolution of symbioses. In particular, the *Rm-Burkholderia* mutualism appears to have evolved from an antagonism, a transition facilitated by the endosymbiont gaining control of its own vertical transmission through manipulation of host asexual and sexual reproduction. Collectively, functional and phylogenomic studies of fungal-bacterial symbioses contribute to the development of novel conceptual frameworks for understanding of symbioses.

Polyploidy and the Origins of Novelty: Impact of Duplication on Genomes and Network Evolution

Pires, J. Chris* (piresjc@missouri.edu)

University of Missouri.

Ancient whole genome duplications (WGDs) are ubiquitous throughout the evolutionary history of higher eukaryotic lineages. These events have been hypothesized to be the basis for major evolutionary transitions, including the origin of novel traits in large species radiations across plants, fungi, protozoa, and animals. Repeated rounds of WGDs, or polyploid events, have been best documented among the flowering plants. We have analyzed the impact of WGDs in *Arabidopsis, Brassica* and other members of the mustard family (Brassicaceae). Phylogenomic analyses of these WGDs show striking correlations of duplicate gene retention and novel traits. For example, the origin of two novel classes of chemical defenses, indole and met- derived glucosinolates (i.e. mustard oils), are associated with duplicated regulatory and biosynthetic pathways that arose via WGDs. Our analyses suggest that the origin of these novel defense compounds spurred an evolutionary arms-race with insect herbivores, resulting in massive co-radiations of both the host plant and predatory insects. The use of new systems biology and network approaches to discover co-evolved modules will be discussed.

Adaptation of Microbial Eukaryotes to Low Oxygen Conditions and the Gastrointestinal Tract

Roger, Andrew J.* (andrew.roger@dal.ca)

Centre for Comparative Genomics and Evolutionary Bioinformatics, Dalhousie University, Halifax, N.S.

All living eukaryotes are descended from a common ancestor that had mitochondria capable of aerobic respiration. However, a vast diversity of microbial eukaryote (protistan) lineages have secondarily adapted to living in low oxygen conditions; a number of these lineages include prevalent parasites inhabiting the gastrointestinal tract. Comparative genomic and transcriptomic approaches provide a window into the metabolic diversity of these protists and reveal the evolutionary innovations that have allowed them to thrive in these environments. I will showcase the extreme genomic diversity amongst the hyper-prevalent human gut parasites *Blastocystis* spp. and show how lateral gene transfer from bacteria, archaea and other eukaryotes has been critical in facilitating their adaptation to hypoxia and the gut environment.

Giant Viruses with an Unprecedented Translation System

Schulz, Frederik* (fschulz@lbl.gov)

DOE Joint Genome Institute, Walnut Creek, CA.

Setting a Safe Course for Gene Editing Research

Wegrzyn, Renee* (renee.wegrzyn@darpa.mil)

Defense Advanced Research Projects Agency (DARPA), Arlington, VA.

The emergence of advanced genome editing tools has created the ability to modify genetic material in a manner that is precise, rapid, and broadly accessible. These editing tools have not only enabled significant advancements in genetic research, including manipulation of previously inaccessible genomes, but have also set the groundwork for transformative applications. These capabilities have the potential to impact biosafety and biosecurity and thus require a novel focus on biological capabilities that facilitate the safe pursuit of genome editing applications while providing tools and methodologies to mitigate risk. DARPA has implemented a "safety first" approach to the development of gene editors and derivative tools with the goal of fostering responsible innovation while mitigating the risk of unintended consequences. In addition to enhancing biosafety and biosecurity, DARPA is also focused on building a scalable, integrated, rapid design and prototyping infrastructure for the facile engineering of biology. This infrastructure will enable the generation of molecules and materials with myriad applications in a manner that will enhance the prototyping process through decreasing cost while increasing throughput.

Investigating Plant Terpene Metabolic Diversity for Biotechnology Applications

Zerbe, Philipp* (pzerbe@ucdavis.edu)

University of California at Davis, Department of Plant Biology, CA.

Terpenoids form the largest and most diverse class of plant specialized metabolites with wide-ranging functions in plant development and ecological adaptation. This chemical diversity also provides a rich and largely untapped resource for bioproduct manufacture and the improvement of crop vigor and yield. However, a broader industrial application of plant-derived terpenoids remains limited by their narrow taxonomic distribution, low abundance and complex diversity in nature. To deepen our knowledge of specialized terpenoid metabolism and accelerate the development of biotechnology applications, we established deep transcriptome resources for more than a dozen plant species that produce terpenoid metabolites of established or potential economic importance. We developed an efficient gene discovery platform, integrating metabolite and transcript profiling with functional enzyme characterization through co-expression in microbial and plant-based platforms that enable effective enzyme cross-validation. This approach revealed numerous novel terpene synthases and cytochrome P450 monooxygenases as key enzymes in terpenoid metabolism, including previously hidden terpenoid pathways with roles in biotic and abiotic stress resilience. Across several food, bioenergy and medicinal plant species, we identified multi-enzyme terpene synthase families that form part of dynamic modular pathways, where catalytically distinct enzymes can function in different combinations to enhance metabolic diversity. Following nature's lead, we develop proof-of-concept microbial and plant-based platforms for producing diterpenoids via co-expression of functionally distinct terpenoid pathway genes.

http://zerbelab.weebly.com/

Poster Presentations

Posters alphabetical by presenting author*

From Genome Content to Substrate Conversion: A Case of Orthology Inside the Fungal Kingdom

Aguilar Pontes, Maria Victoria* (v.aguilar@westerdijkinstitute.nl)[1]; Wiebenga, Ad[1]; Snel, Berend[2]; de Vries, Ronald P.[1]

[1]Fungal Physiology, Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology, Utrecht University, The Netherlands. [2]Fungal Physiology, Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology, Utrecht University, The Netherlands. [3]Theoretical Biology and Bioinformatics, Department of Biology, Science Faculty, Utrecht University, The Netherlands. [4]Fungal Physiology, Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology, Utrecht University, The Netherlands.

Fungi are found in all natural and man-made biotopes and can use highly diverse carbon sources. Plant biomass is the major source of carbon in nature for most fungi. Most natural carbon sources are polymeric and cannot be taken up into the cell directly. This means that fungi need to produce the right set of extracellular and metabolic enzymes that matches the available carbon sources in order to break down the polymers and utilized them to grow. However, little is known about the exact set of enzymes employed to degrade the different carbon sources. This complicates the development of new and more efficient plant biomass degrading enzyme cocktails for industrial purposes.

The ascomycete *Aspergillus niger* is one of the most utilized fungi in industry. Due to these applications, several carbon metabolic pathways have been studied in detail, which allowed us to identify the genes involved in the conversion of the different plant biomass derived monosaccharides. Based on this set of genes, we compared the results of different orthology inference methods in order to identify the presence of genes from a number of catabolic pathways in selected fungi, from the Ascomycota and the Basidiomycota, and validate the results using experimental growth data.

Exploring Fungal Dark Matter Using Single-Cell Genomics

Ahrendt, Steven R.* (sahrendt0@lbl.gov)[1, 2]; Quandt, Alisha C.[3]; Ciobanu, Doina[1]; Clum, Alicia[1]; Salamov, Asaf[1]; Cheng, Jan-Fang[1]; Woyke, Tanja[1]; James, Timothy Y.[3]; Grigoriev, Igor[1, 2]

[1]DOE Joint Genome Institute, Walnut Creek, CA. [2]Plant and Microbial Biology, University of California, Berkeley, CA. [3]Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI.

Current estimates suggest that only approximately 100,000 of the estimated 5 million fungal species worldwide have been described, and the overwhelming majority of those fall within the Dikarya. Thus, the diversity of the non-Dikarya lineages has been poorly explored. Our lacking insight into the metabolic potential of these fungi is a reflection of their absence from sequence databases, limiting our ability for meaningful comparative analyses and lifestyle predictions. Exacerbating this issue is the fact that a substantial fraction of these organisms is largely uncultivated, challenging genomics

exploration, despite being detected in environmental PCR surveys as a significant fraction of a community. The difficulties inherent in exploring the genetic make-up of this "fungal dark matter" can be overcome using single-cell sequencing to reconstruct genomes of uncultivated organisms directly isolated from the environment. Environmental fungi, particularly among the zoosporic and other early-diverging fungal lineages, make exceptional targets for single-cell genomic techniques; however, as much of the current single-cell genomic work focuses on mammalian, bacterial, and archaeal systems, there is a pressing need to adopt these protocols for fungi. Here we developed and tested these single-cell methods to sequence the genomes of seven species which cannot be grown in pure culture, including the first representatives of the previously unsequenced Zoopagomycotina subphylum. We show that although there is a large variation in gene space recovery from each single cell (ranging from 6 - 88%), combining multiple cell libraries can increases this recovery to around 90%. Phylogenomic analyses allowed us to place previously unsampled lineages within the fungal tree of life, even when considering partially complete genomes derived from single individual cells. Additionally we explored gene family expansions to identify patterns consistent with lifestyle aspects of these biotrophic organisms.

Community RNA Reveals Changes in Soil Biodiversity and Ecosystem Function in a Long-Term Soil Warming Chronosequence *(Short Talk)*

Alteio, Lauren V.* (lalteio@umass.edu)[1]; Rodriguez, William[1]; Billings, Andrew[1]; van Diepen, Linda[2]; DeAngelis, Kristen[1]; Frey, Serita[3]; Melillo, Jerry[4]; Blanchard, Jeffrey[1]

[1]University of Massachusetts Amherst, Amherst, MA. [2]University of Wyoming, Laramie, WY. [3]University of New Hampshire, Durham, NH. [4]Marine Biological Laboratory, Woods Hole, MA.

Introduction: The Harvard Forest site is home to three experimental soil warming sites heated continuously to 5° C above ambient temperature. When taken together, the three soil warming experiments at the Harvard Forest make up a unique, dynamic time sequence, or chronosequence, based on a common perturbation in a single ecosystem type.

Objectives: To relate changes in CO_2 emissions to changes soil community composition and physiology using a chronosequence of warming experiments.

Materials & Methods: Four soil samples were taken from control and warmed plots in each of the three experiments. Total RNA was extracted from the soil samples and submitted to the Department of Energy's Joint Genome Institute (JGI) for sequencing. Protein-coding genes were annotated to taxonomy using the DIAMOND algorithm, an advanced blastX approach, and annotated to functional categories using the SEED, KEGG and InterPRO databases. The resulting data was imported in R for statistical analyses.

Results: The metatranscriptomic sequencing results included a diverse representation of the soil organismal community including viruses, archaea, bacteria, protists, fungi, invertebrates and plants, a view that we would expect to see in a canonical soil food web illustration. Warming has resulted in a significant decrease in eukaryotic transcripts relative to bacteria, with relative losses in transcript abundance in classes of ascomycetes, basidomycetes and arthropods and increases in bacterial taxa. Cellular and ecosystem functions have been effected with respect to warming, including those related to protein stability, selfish genetic elements, toxin resistance, and biogeochemical cycling (C, N, P, S, Fe).

Conclusions: Continued elevated CO_2 emissions coupled with progressive changes in soil community structure and function with long-term experimental warming suggest the effects of rising global temperatures are unlikely to be ephemeral and will produce complex ecological feedbacks.

Understanding Microbial Influences on Plant Response on a Genome-Wide Scale

Amos, Kirtley B.* (bkamos3@g.uky.edu)[1]; Sanchez Barrios, Andrea[1]; Lundberg, Derek[2]; De Lorenzo, Laura[2]; DeBolt, Seth[1]

[1]Department of Horticulture and Plant Physiology/Biochemestry/Molecular Biology Program, University of Kentucky Lexington, Kentucky 40546, USA. [2]Department of Plant and Soil Sciences, University of Kentucky, Lexington, KY, 40546-0312, USA. [3]Department of Molecular Biology, Max Planck Institute for Developmental Biology, Tübingen 72076, Germany.

The addition of microbial inoculums that influence plant traits, both agronomical and genetic, driving increased cellular expansion and transcript abundance, is the focus of study. Early changes within the rhizosphere could lead to a systematic reselection of endophytic bacterial communities, drastically influencing aerial tissue. Based on our interest in cell expansion, we observed *Nicotiana benthamiana* treated with monocultures of bacterial inoculums that promote or restrict growth. Application of three bacterial strains (*Bacillus cereus, Bacillus* sp., and *Micrococcus* sp.) isolated from tomato and switchgrass applied to *N. benthamiana* seeds as amendments presented different phenotypes. Statistically significant agronomical characteristics increasing cell expansion were associated with B. *cereus* treatments and included plant height, leaf length, number of leaves and flowers (Sanchez Barrios et al., unpublished). Measurements of mRNA abundance using genome-scale approaches (Poly(A)-Tag mRNA-sequencing, PAT-seq) revealed genes involved in stress responses were predominantly downregulated for two growth promoting treatments (*B. cereus* and *B.* sp.) and upregulated within one (*M.* sp.). Due to possible hormonal influences driven by bacteria and N. benthamiana having a draft genome, we decided measure plant hormones. Data showed gibberellin and auxin metabolites were at higher concentrations after B. cereus treatment in *N. benthamiana* compared to *M.* sp. and control, supporting PAT-seq analysis and increased cellular expansion. To further support our investigations we became interested in unveiling genetic interactions between the monocultures and *N. benthamiana*, therefore, we mapped PAT-seq data from *N. benthamiana* to genomes of the bacterial species applied. These studies provide insight into the interactions between plants and microbes and the potential for alternative approaches to traditional farming practices through the pre-treatment of seeds via bacterial coatings.

Assembly, Screening and Functional Characterization of Glucaric Acid Pathway in Model Organisms *E. Coli* and *S. Cerevisiae*

Anderson, Lisa A.* (laanders@mit.edu)[1]; Guay, Lisa M.[1]; Hicks, Michael A.[1]; Young, Eric M.[2, 3]; Sandoval, Luis R.[1]; Voigt, Christopher A.[2, 3]; Prather, Kristala L. J.[1]

[1]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.
[2]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.
[3]MIT-Broad Foundry, The Broad Institute of Harvard and MIT, Cambridge, MA 02142.

Glucaric acid is one of the top value-added chemicals from biomass according to the 2004 report from the U.S. Department of Energy. The Prather Lab has developed a biosynthetic pathway in *E. coli* and *S. cerevisiae* to generate glucaric acid from glucose in four steps, producing titers of up to 2.5 g/L from 10 g/L glucose. The goal of this proposal is to optimize a glucaric acid production with two methods: 1) homologous enzyme identification and screening; and 2) alternative promoter testing with variations in gene dosage. Homologs for pathway enzymes were identified from their respective Pfam protein families and selected for evaluation using sequence similarity networks to maximize enzyme diversity while maintaining key catalytic residues. We previously identified that myo-inositol oxygenase (MIOX) and myo-inositol-1-phosphate synthase (MIPS) are rate-limiting due to the interplay between growth and production. Assembly and screening of 31 homologs of MIOX, which converts myo-inositol to D-glucuronic acid, is in progress and we have begun screening promoter and terminator pairs to balance expression and production during growth. Preliminary screening of 31 homologs of MIPS, which converts glucose-6-phosphate to myo-inositol and represents the branch point from glycolysis has begun in *E. coli.* Initial results suggest at least one homolog produces significantly higher titers of myo-inositol from glucose. With a set of synthesized constructs containing both a library of promoters and enzyme homologs, permutations of these pathway components can be efficiently evaluated and optimized.

Lignin Degrading Enzymes from the Jurassic to Date

Ayuso-Fernandez, Ivan*; Martinez, Angel T. (atmartinez@cib.csic.es); Ruiz-Dueñas, Francisco J.

Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, E-28040, Madrid, Spain.

Floudas et al. (*Science* 336: 1715) established that lignin-degrading fungi appeared at the end of Carboniferous period associated with the production of the first ligninolytic peroxidases. Here, the subsequent evolution of these enzymes in Polyporales, where most wood-rotting fungi are included, is experimentally recreated using genomic information (from JGI). With this purpose, we analyzed the evolutionary pathway leading to the most efficient lignin-degrading peroxidases existing today, being characteristic of the order Polyporales, using a robust phylogenetic tree obtained by maximum likelihood methods. After ancestral sequence reconstruction from 113 peroxidase genes of ten sequenced genomes using the PAML software, the main enzyme intermediates (corresponding to different nodes in the above evolutionary pathway) were resurrected and characterized in the laboratory (after *Escherichia coli* expression and *in vitro* activation). Biochemical characteristics (including activity on phenolic/nonphenolic aromatics, Mn2+ and dyes, and stability against temperature and pH) were analyzed together with the predicted sequences and molecular structures, to understand how these enzymes acquired the ability to degrade lignin and how this ability changed with time. The most probable first peroxidase in Polyporales was predicted as a manganese peroxidase (Mn3+ oxidizing phenolic lignin) that did not change substantially until the appearance of an exposed tryptophan (responsible for nonphenolic lignin oxidation) originating an ancestral versatile peroxidase. Later, a quick evolution, with loss of the Mn2+-binding site, generated the first lignin peroxidase that evolved to the extant form by improving the catalytic efficiency on nonphenolic aromatics. Interestingly, higher stability at acidic pH, which strongly increases the oxidizing power of heme peroxidases, was observed paralleling the appearance of the exposed catalytic tryptophan. We show how the change in peroxidase catalytic activities meant an evolutionary exploration for more efficient ways of lignin degradation by fungi, a key step for carbon recycling in land ecosystems. The study also provided stable and highly-evolvable ancestral enzymes with a biotechnological potential in biomass-based economy.

Hidden diversity in the Oomycete Genus *Olpidiopsis* Is a Global Threat to Red Algal Cultivation

Badis, Yacine*[1]; Gachon, Claire M. M. (Claire.Gachon@sams.ac.uk)[1]; Klochkova, Tatyana A.[2]; Strittmatter, Martina[1, 3]; Garvetto, Andrea[1]; Murúa, Pedro[1]; Ostrowski, Martin[4]; Tringe, Susannah G.[5]; Prochnik, Simon E.[5]: Sanderson, Craig[6] Kim, Gwang Hoon[7]

[1]The Scottish Association for Marine Science, Scottish Marine Institute, Oban, Argyll PA37 1QA, United Kingdom. [2]Kamchatka State Technical University, Petropavlovsk-Kamchatsky, 683003, Russia. [3]Roscoff Biological Station, Place Georges Teissier, 29680 Roscoff, France. [4]Department of Biochemistry and Biomolecular Science, Faculty of Science and Engineering, Macquarie University, Sydney, NSW, Australia. [5]US DOE Joint Genome Institute, 2800 Mitchell Dr., Walnut Creek, CA 94598, USA. [6]School of Zoology, University of Tasmania, Hobart, Australia. 7Department of Biology, Kongju National University, Kongju, Korea.

Seaweed cultivation is the fastest-growing of all aquaculture sectors, with an annual growth rate of 8 % and a value in excess of $5bn. Emerging diseases are a threat to its sustained development. The oomycete genus *Olpidiopsis* contains pathogens of red seaweeds, including the most economically damaging disease in Pyropia (ex-*Porphyra*) farms in Asia. Here we identified three new *Olpidiopsis* species: *Olpidiopsis palmariae, O. muelleri,* and *O. polysiphoniae* spp. nov. A Scottish variety of *O. porphyrae,* a devastating pathogen only reported in Japanese seaweed farms, is also described as *O. porphyrae* var. scotiae. Importantly, two of the new species infected *Porphyra* and *Palmaria* sp, which are subject extensive farming trials in Europe and North America. To further assess the extent of undescribed *Olpidiopsis* diversity and the threat it might pose to aquaculture in different regions, we screened targeted metagenomes of *Porphyra umbilicalis* blades collected on the West and East Northern Atlantic shores (JGI proposal ID 946). Four different *Olpidiopsis* OTUs were detected on at least 5 out of 9 *Porphyra* blade metagenomes, illustrating the prevalence and diversity of those parasites in wild *Porphyra* populations. Finally, we extended our screening efforts to global metagenomic barcoding campaigns, revealing over 700 sequencing reads attributable to *Olpidiopsis* with a worldwide distribution. Close relatives of the Korean *O. pyropiae* are reported for the first time in Europe and the United States. In the light of our restricted sampling, our results highlight the diversity and abundance of *Olpidiopsis* worldwide. In the context of worsening impact of *Olpidiopsis* pathogens in Asia, this worldwide distribution should be treated as a serious threat to the global seaweed industry and wild red algal populations. Our data calls for more efforts towards the documentation of these pathogens, and for adequate biosecurity measures to be developed.

Microbes in Forest Soils: Tracing the Activity of Communities and Individual Taxa

Baldrian, Petr* (baldrian@biomed.cas.cz)

Laboratory of Environmental Microbiology, Institute of Microbiology of the CAS, Prague, Czech Republic.

In forest soils, microbes are important drivers of soil processes, because they mediate decomposition as well as nutrient transfer from primary producers into soil. Metatranscriptomics, metaproteomics and enzyme activity assays that can assess the dynamics of microbial processes and assign functions to higher taxa show that microbial activity undergoes seasonal cycling with a fungal dominance in summer and bacterial dominance in winter affecting multiple processes ranging from ectomycorrhizal

symbiosis to decomposition (Žifčáková 2016). The results characterize the summer season as a period with rapid microbial growth accompanied by decomposition of recalcitrant plant biopolymers likely induced by the priming effect of photosynthates delivered by plant roots. Winter appears to be a period of slow growth when reserve compounds such as starch, glycogen and trehalose are utilized. Methods focusing environmental metacommunities are, however, not able to identify individual microbial species, participating in the soil processes. To do that, individual microbial taxa need to be isolated and analyzed. Genomes of 20 dominant bacterial strains from the spruce forest soil were sequenced as well as genomes of ectomycorrhizal fungi, obtained from fruitbodies collected at the study sites. Mapping of metatranscriptomic reads on genomes of dominant bacteria showed, interestingly, that Acidobacteria are the likely major producers of decomposition-related enzymes that possess high counts of glycosyl hydrolases in their genomes, in contrast to Bacteroidetes and Proteobacteria. Transcript profiles of bacteria and fungi *in situ* differ among horizons as well as among seasons showing the effect of environmental conditions on transcription.

Žifčáková L, Větrovský T, Howe A, Baldrian P. 2016. Microbial activity in forest soil reflects the changes in ecosystem properties between summer and winter. Environmental Microbiology 18:288-301.

Omics-Enabled Discovery of a Novel Bacterial Enzyme Enabling First-Time Bio-Based Toluene Production

Beller, Harry R.* (HRBeller@lbl.gov)[1]; Rodrigues, Andria[1]; Zargar, Kamrun[1]; Saini, Avneesh[1]; Tringe, Susannah G.[2]; Keasling, Jay D.[1]; Petzold, Christopher J.[1]

[1]Joint BioEnergy Institute (JBEI), Emeryville, CA. [2]Joint Genome Institute (JGI), Walnut Creek, CA.

Toluene is an important petrochemical with a global market of 29 million tons per year; it would be desirable to offset the enormous volume of petroleum-derived toluene with biochemically produced toluene made from a renewable resource (such as cellulosic biomass) using engineered microbes. However, the enzymes necessary to catalyze biochemical toluene synthesis were heretofore unknown. Although anaerobic bacterial biosynthesis of toluene from phenylacetic acid was reported more than two decades ago, the biochemistry underlying this novel metabolism has never been elucidated. Here we report the omics-enabled discovery of a toluene synthase (phenylacetate decarboxylase) from an anaerobic, sewage-derived enrichment culture that quantitatively produces toluene from phenylacetate. The discovery process included metagenome sequencing of the culture (which included more than 340,000 protein-coding genes), development of a sensitive and specific GC/MS anaerobic assay to detect labeled toluene from labeled phenylacetate, anaerobic FPLC (fast protein liquid chromatography) of cell-free extracts of the culture, and differential metaproteomic analyses to identify proteins present in active (toluene-producing) FPLC fractions but absent in adjacent inactive FPLC fractions (i.e., toluene synthase candidates). Toluene synthase candidates included a novel glycyl radical enzyme (GRE) and its cognate activating enzyme [AE; a radical SAM (S-adenosyl-L-methionine) enzyme]. Recombinant, N-terminally tagged, codon-optimized versions of the GRE and AE genes were expressed in *E. coli* and purified under anaerobic conditions. After *in vitro* reconstitution of the AE to restore its [4Fe-4S] cluster, its activity was confirmed *in vitro* by measuring conversion of SAM to methionine. *In vitro* assays with the purified GRE, AE, and SAM were shown to successfully convert C13-labeled phenylacetate to C13-labeled toluene, whereas no toluene was produced in control assays lacking SAM. Thus, using an omics-enabled approach, we have discovered a novel glycyl radical enzyme (only 6 are currently known) that catalyzes decarboxylation of phenylacetate to form toluene.

This first-time identification of a phenylacetate decarboxylase will ultimately enable bio-based toluene production via engineered microbial hosts.

Microbial Ecology, Biogeochemistry, and Metagenomics of Methane Oxidation and Production in High Altitude Lakes

Beman, Michael J.* (mbeman@ucmerced.edu); Perez Coronel, Elisabet

University of California, Merced.

Freshwater lakes constitute a disproportionately large source of the important greenhouse gas methane (CH_4) to the atmosphere. Within lake water columns, microbial CH_4 oxidation reduces overall water-to-air flux by approximately 80%, such that small changes in microbial CH_4 oxidation produce large changes in overall CH_4 flux. However, the environmental factors that affect methane oxidation rates—as well as the genomic content and community composition of methanotrophs—are poorly understood. We examined methane oxidation and methanotroph communities along a freshwater lake elevation gradient in the Sierra Nevada, California that provides predictable differences in temperature, nutrient concentrations, and production with increasing elevation. Using biogeochemical measurements coupled with microbial community analyses and metagenome sequencing, we found active and diverse methanotroph communities that varied in space and time. CH_4 concentrations were elevated at high elevations, yet CH_4 oxidation was rapid and supported high methanotroph abundances. Metagenomics provided detailed insight into different segments of the methanotroph community—as well as other co-occurring organisms—along this gradient. In particular, our data capture the sporadic occurrence of methanogens within the oxygenated water column—an observation that has been made in other lakes and attributed to phytoplankton-methanogen interactions. We provide additional evidence for this idea based on metagenomic data. Altogether, our results indicate that CH_4 cycling in high altitude lakes is highly dynamic, with multiple, potentially interacting microbial groups regulating CH_4 cycling and flux.

Disco: High Performance Scalable Assembler for Integrated Assembly of Illumina and PacBio Reads *(Short Talk)*

Biswas, Abhishek* (biswasa@ornl.gov)[1, 2]; Chai-Crosskey, Juan[2]; Clum, Alicia[3]; Guo, Xuan[2]; Bushnell, Brian[3]; Yao, Qiuming[1]; Tringe, Susannah G.[3]; Hervey, IV, William J.[4]; Pan, Chongle[1, 2]

[1]Graduate School of Genome Science and Technology, University of Tennessee, Knoxville, TN. [2]Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN. [3]DOE Joint Genome Institute (JGI), Walnut Creek, CA. [4]Center for Bio/Molecular Science & Engineering, U. S. Naval Research Laboratory,Washington, DC.

Disco is a scalable assembler developed to assemble billions of Illumina sequencing reads using an overlap layout graph algorithm. Disco implements a novel algorithm for using long PacBio reads to guide the assembly of unitigs generated from the overlap graph simplification of Illumina reads. Accurate seed unitigs are selected and extended along overlap paths in the graph that are verified by iterative mapping with the longer PacBio reads. The integration of PacBio read information while

simplifying the overlap graph improves the accuracy of the graph simplification algorithm and generates a more accurate and contiguous assembly. Disco is parallelized for computer clusters using a hybrid architecture that integrated shared-memory multi-threading, point-to-point message passing, and remote direct memory access. The computational performance of Disco was evaluated from single computers to high-performance-computing clusters. On single computers, Disco was more computationally efficient than SPAdes, ABySS, MEGAHIT, and MetaVelvet. On computer clusters, Disco also scaled better than Ray Meta, ABySS and HipMer. Disco is capable of assembling billions of Illumina reads generated from metagenome and plant genome sequencing and further improve assembly quality with PacBio reads. Disco is easy to use and freely available at http://disco.omicsbio.org.

Towards Lignin Degrading Enzyme Expression in *Yarrowia lipolytica*

Blenner, Mark A.* (blenner@clemson.edu); Rives, Dyllan

Clemson University, Dept. of Chemical & Biomolecular Engineering.

The overall goal of this project is to refactor the major lignin degradation pathways from various fungi and heterologously express these enzymes in a lipid producing yeast, *Yarrowia lipolytica*. The relevance of such a strain would be in consolidated bioprocessing or it could be used to convert lignin wastes into biofuel and biofuel precursors. The lignin degrading fungi from which these enzymes were first described are efficient at lignin catabolism; however, they lack robust genetic tools for tuning and controlling their expression, and they do not readily make biofuel or biofuel precursors. A further consequence of this lack of genetic tools is the inability to engineer these fungi to produce biofuels. These fungi act in concert to metabolize wood in nature, but in order to use a pure industrial microbe, such as *Y. lipolytica,* it will be important for us to be able to tune the expression of the six key enzymes we identified as critical for lignin degradation. As a starting point, these six enzymes will be expressed in *Y. lipolytica* and tested for functional activity. Initial expression from episomal vectors showed little activity. This poster discusses our recent evidence towards understanding the low activity from these enzymes, and the impact of media conditions on enzyme activity.

Phylogenetic and Functional Diversity

Brocchieri, Luciano* (lucianob@ufl.edu)[1,2]

[1]Department of Molecular Genetics and Microbiology, University of Florida, Gainesville, FL 32610. [2]Genetics Institute, University of Florida, Gainesville, FL 32610.

Phylogenetic diversity is commonly used as a proxy for functional diversity of ecological communities. In fact, measures of phylogenetic diversity operate on phylogenetic trees to estimate the distribution of frequency of features in a community and then calculate diversity as a function of these estimates. All current indexes of phylogenetic diversity assume a parsimonious model of evolution of phenotypic features (Camin-Sokal parsimony). We developed measures of phylogenetic diversity based on non-parsimonious models of evolution of features, including probabilistic models of protein-sequence evolution, and of gene-content evolution based on a birth and death model modified from Huson and Steel (2004). We calculated phylogenetic diversities for 3,446 phylogenetically-independent sets of sequences, from a phylogenetic tree of 14,727 bacterial and archaeal strains obtained from

alignments of conserved proteins. Furthermore, we calculated functional diversities for the same sets based on occurrence of 801 functional pathways. We found that phylogenetic diversities calculated based on parsimonious evolution correlated poorly with functional diversities, explaining from 6% to 28% of total variance in functional diversity. Phylogenetic diversities calculated based on probabilistic models of protein and gene evolution resulted instead in significantly improved correlations with functional diversity, explaining from 65% to 82% of variance in functional diversity. Our results indicate that phylogenetic diversity can be a significantly improved proxy for functional diversity when a non-parsimonious evolution of functional features is accounted for. Furthermore, they suggest that diversity in amino acid sequences is a useful marker of functional diversity in microbial communities.

Capturing the Complete Genome of *Botryococcus braunii* by Combining Illumina and PacBio Assemblies

Browne, Daniel R.* (dbrowne.up@gmail.com)[1]; Jenkins, Jerry[2]; Schmutz, Jeremy[2, 3]; Barry, Kerrie[3]; Koppisch, Andrew T.[4]; Devarenne, Timothy P.[1]

[1]Department of Biochemistry & Biophysics, Texas A&M University, College Station, Texas. [2]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA. [3]Joint Genome Institute, Walnut Creek, California, USA. [4]Department of Chemistry, Northern Arizona University, Flagstaff, Arizona, USA.

Assembling complex, eukaryotic genomes is still a tremendous challenge. Repetitive sequences and regions of extremely high or low GC content impede sequencing efforts. Moreover, the various DNA sequencers tend to have biases. Efforts to understand the sequence biases of the Illumina and PacBio instruments have shown that these technologies have different biases. Furthermore, sequence bias appears to vary with the underlying sequence complexity, which differs across species. The genome of the oil producing, colonial green microalga *Botryococcus braunii* has proved to be a very difficult sequencing and assembly challenge. Multiple sequencing strategies, including Illumina and PacBio, individually failed to yield a high-quality genome assembly. Here we present our finding that the Illumina and PacBio sequencing runs captured separate but partially overlapping sets of genomic loci. Building on this observation, we developed a de Bruijn graph-based strategy to merge the separate Illumina and PacBio assemblies, yielding a more complete and more contiguous draft genome assembly. This approach could be applied to other species with hard-to-assemble genomes and can be stacked on top of almost any genome assembly workflow.

Ionizing Radiation Resistance in Experimentally Evolved *Escherichia coli* Populations

Bruckbauer, Steven T.*; Cox, Michael M. (cox@biochem.wisc.edu)

Department of Biochemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706-1544.

All organisms have evolved defenses against harmful reactive oxygen species (ROS) and lethal DNA double strand breaks (DSBs) caused by ionizing radiation (IR). However, some organisms exhibit levels of resistance that are extraordinary. We have initiated a new effort to generate *E. coli* strains that are highly resistant to gamma ionizing radiation by directed evolution. The new effort uses a linear accelerator to administer IR, allowing for controlled, reproducible, and highly accurate conditions

throughout multiple evolution trials. The dose rate of this device is nearly four times that of the dose rate used in earlier trials, allowing for an unprecedented look at resistance to extreme doses of IR. After more than 50 cycles of irradiation followed by outgrowth of survivors, we are generating the most IR resistant *E. coli* populations ever produced in a laboratory. In all cases, the cells are exhibiting a major revamping of cellular nucleic acid metabolism. The mutations in the genes encoding RecA, RecBCD, RecN, components of the replication restart primosome, RNA polymerase, multiple helicases, and many other functions, are altering and streamlining cellular systems to effect very efficient repair of double strand DNA breaks. Genomic alterations that ameliorate the effects of reactive oxygen species and alter cell wall structure are also evident. An update of this project will be presented.

Bacterial and Archaeal Biogeochemically-Relevant Gene Abundance and Expression Vary Relative to Different Spatial and Environmental Scales in the Delaware Bay

Campbell, Barbara J.* (bcampb7@clemson.edu); Gholamian, Jason; Wang, Kai

Department of Biological Sciences, Clemson University, Clemson, SC.

Microbes mediate many important biogeochemical cycles, including nitrogen (N), sulfur (S) and phosphorus (PO4). Estuarine archaea and bacteria are known to vary with habitat, time, and physicochemical gradients. However, microspatial and seasonal variation of biogeochemically-relevant genes and gene expression are largely uncharacterized in these microbes. We determined the abundance and expression of genes mediating N, S and P cycling at microspatial (free-living (FL) and particle-attached (PA) fractions) and seasonal scales along the salinity gradient in the Delaware Bay. Bacterial and archaeal gene composition significantly varied with season and salinity but not by size fraction. Discrete patterns in transcript composition were observed microspatially and were less distinct with season and salinity. Major N metabolic pathways were more expressed in PA cells than in FL cells but their abundances did not significantly differ. Genes mediating ammonia oxidation were more abundant and expressed in summer than in spring, while genes mediating dissimilatory nitrate reduction and denitrification were more abundant and expressed in spring. Ubiquitously, N metabolic genes were more abundant and expressed at low salinity than at high salinity. PO4-transporter genes were more expressed in FL cells than in PA cells, while an alkaline phosphatase gene was more expressed in PA cells, suggesting free PO4 starvation in PA cells. Genes mediating assimilatory sulfate reduction (SR) were more expressed in PA cells than in FL cells, while FL cells contributed more transcripts for dissimilatory SR. Most differentially expressed genes mediating SR and SOX were expressed more in summer than in spring, though SOX genes were all more abundant in spring. Up to 70% of differentially expressed S metabolic genes were expressed more at high salinity than at low salinity. Conclusions: Covariation of gene abundance and expression with season and salinity reflected the genetic and functional adaption of bacteria and archaea along the Delaware Bay. Although less variation in biogeochemically-relevant genes was found between FL and PA cells, their expression was significantly different. Our results suggest functional redundancy in genetic potential across spatial gradients and seasons, but microenvironments control the functional activity of these microbes.

Microbial Communities Involved in Subsurface Nitrogen Cycling across the Upper Colorado River Basin and Potential Implications for Uranium Release

Cardarelli, Emily L.* (ecardare@stanford.edu)[1]; Bargar, John[2]; Francis, Christopher[1]

[1]Department of Earth System Science, Stanford University, Stanford, CA. [2]SLAC National Accelerator Laboratory, Menlo Park, CA.

Throughout the upper Colorado River Basin, uranium (U) persists as a legacy contaminant of former ore processing activities. While these sediments are generally nutrient-poor due to their silicoclastic composition, large inventories of nutrients and elevated solid-phase uranium levels exist [as low-solubility U(IV)] in fine-grained, organic-enriched sediments. Iron sulfide minerals are abundant within these layers, giving rise to the moniker "naturally reduced zones" (NRZs). There is concern that NRZs are acting as slow-release sources of uranium to the aquifer that could persist for hundreds of years. Nitrate, produced by the decomposition of organic matter coupled to nitrification, is suspected of playing a central role in NRZ-floodplain biogeochemistry by (i) driving heterotrophic carbon cycling in NRZs and (ii) oxidizing the stored U(IV) to relatively mobile U(VI), which makes possible its escape to the surrounding aquifer. Previous researchers have posited that nitrate and nitrite can oxidize U(IV). Nitrate/nitrite may diffuse into the reduced interiors of NRZs and become readily available for denitrification, the stepwise anaerobic reduction of nitrate/nitrite to dinitrogen gas. Denitrification may then be coupled to the oxidation of sediment-bound U(IV), forming mobile U(VI), allowing it to resupply uranium in the aquifer. Thus, N-cycling microbial communities are believed to be fundamentally important to the mobility and fate of organic carbon and uranium. However, very few studies have examined N-cycling communities within NRZ sediments or the subsurface.

Here we examine the overall microbial community composition as well as the diversity and abundance of key N-cycling genes within five DOE-Legacy Management sites (Rifle, Grand Junction, and Naturita, CO; Shiprock, NM, and Riverton, WY). These sites span a 900 km longitudinal transect and the samples retrieved extend to a depth of 10 m at each site throughout the upper CRB. We examine community changes, through depth and across the basin, with a combination of functional gene approaches, 'deep' 16S rRNA amplicon (iTag) sequencing, and extensive geochemical data. Our initial results suggest that the composition of Thaumarchaeota communities within the subsurface is driven by interactions with the watertable and the NRZs. There also appear to be distinct ecotypes associated with different core depths and geochemical conditions at each site. Overall, the geochemistry and community composition seems to be location-dependent, with Rifle, CO as the most distinct site within the region. This data has helped illuminate sites and depths where we will use metagenomic and metatranscriptomic sequencing to acquire valuable new insights into the phylogenetic, genomic, and functional/metabolic diversity of subsurface N-cycling microbial communities within the upper CRB.

Pathway Visualization Tools in Phytozome

Carlson, Joseph W.* (JWCarlson@lbl.gov); Davidson, Patrick; Goodstein, David; Rokhsar, Daniel

DOE Joint Genome Institute Walnut Creek, CA.

We have recently begun adding metabolic pathway data to the genomes hosted in Phytozome, the Joint Genome Institute's plant genomics portal.

Our initial set of assignments is based on our in-house implementation of the Pathologic (http://pathwaytools.org/) workflow, coupled with Enzyme E.C. assignments derived from E2P2 (https://dpb.carnegiescience.edu/labs/rhee-lab/software).

We provide access to these pathway assignments via PhytoMine (http://phytozome.jgi.doe.gov/phytomine/begin.do), Phytozome's implementation of the open source genomic data warehouse InterMine (http://www.intermine.org).

While InterMine's data model can be extended to accommodate pathway information, the InterMine UI does not include any default pathway visualization capabilities. To provide access to these assignments to our users, we designed a pathway visualization tool for InterMine, based on the d3 library (http://d3js.org/). This tool supports viewing reactions and enzyme assignments in both graphical and tabular format, as well as the overlay of RNA-Seq expression data (to assist, for example, with identifying conditions under which a particular pathway is over- or under-expressed).

The InterMine data model and visualization tool capabilities will be presented.

Comparative Genomic Analyses of Two Pyrophilous Fungi

Carver, Akiko A.* (acarver@lbl.gov)[1]; Stillman, Kyra[1]; Liu, Haley[1]; Branco, Sara[2]; Kuo, Alan J.[3]; Bruns, Tom[1]; Grigoriev, Igor[1, 3]

[1]Department of Plant and Microbial Biology, University of California, Berkeley, CA, USA. [2]Laboratoire Ecologie, Systématique et Evolution, Un Paris Sud, CNRS-UPS-AgroParisTech, Orsay, France. [3]DOE Joint Genome Institute, Walnut Creek, CA, USA.

Pyrophilous fungi are filamentous fungi that have historically been found fruiting exclusively on burned soil. Forest soil changes greatly after fire, with an enrichment of dead biomass, partially pyrolyzed carbon compounds and increased hydrophobicity. The fungal species that inhabit such soil comprise a diverse guild, though most are Ascomycota in the order Pezizales and most of these are in the family Pyronemataceae. To better understand the role of pyrophilous fungi in post-fire soil, we compare the genomes of two pyrophilous fungi of the Pyronemataceae family, *Pyronema omphalodes* and *Wilcoxina mikolae,* to 12 non-pyrophilous *Pezizales* species. Although *P. omphalodes* and *W. mikolae* have otherwise distinct lifestyles as saprotrophs or as ectomycorrhizae respectively, both are dominant in forest soil communities after fires. They share an expansion in a novel chitin-binding gene family that is unique amongst sequenced fungi and may be horizontally transferred, and both genomes contain glycoside hydrolases that are absent in other sequenced *Pezizomycetes* species. In addition, *P. omphalodes* has an expansion in heat shock proteins over other sequenced Ascomycetes, which may assist in its colonization of burned forest soil. The information from this study will increase our understanding of the lifestyles and metabolic processes of pyrophilous fungi and their role in post-fire carbon cycling.

Best Practices for Whole Genome Sequencing Using the Sequel System

Chakraborty, Shreyasee* (Schakraborty@pacificbiosciences.com); Lambert, Christine; Eng, Kevin; Baybayan, Primo; Hatas, Emily

Pacific Biosciences, 1305 O'Brien Drive, Menlo Park, CA 94025.

Plant and animal whole genome sequencing has proven to be challenging particularly due to size, high density of repetitive elements and various levels of ploidy. The Sequel™ System delivers long read lengths which allow more complete and accurate contiguous assemblies of larger and more complex genomes. With the latest Sequel chemistry v2.0, usable bases of 5 – 10 Gb per SMRT© Cell can be achieved with reduced input SMRTbell™ libraries (as low as 5pM). Read lengths averaging 13 kb or greater can be routinely achieved, with the longest reads approaching 50 kb. Furthermore, 50% of usable bases are greater than 15 kb read length.

Here, we present the best practices for achieving long reads for whole genome shotgun sequencing of complex plant and animal genomes. Guidelines for constructing large insert SMRTbell libraries (> 30 kb) to generate optimal read lengths, using the latest chemistry, will be presented. We also describe ways to maximize library yield (per preparation) from 5 µg sheared genomic DNA. The combination of these advances makes plant and animal whole genome sequencing a practical application of the Sequel System.

RNA Interference Protein Kinase Gene in the Fruiting Development of Module Mushroom—*Coprinopsis Cinerea*

Chan, Po Lam* (kathypolam@hotmail.com); Chang, Jinhui; Kwan, Hoi Shan

School of Life Science, The Chinese University of Hong Kong, Hong Kong SAR, China.

Introduction: Phosphorylation between different protein kinases trigger activation or inhibition of downstream targets affecting their activities and mediating their cellular and developmental responses to growth factors, environmental signals and internal processes as well as vital roles in signal transduction pathways. Protein kinases not only are crucial to the signal transduction, but also are master switches of development. Some protein kinases are proposed to be important in mushroom development processes. Activation or inhibition of interesting and putatively important kinase which involve in the signaling transduction pathway for fruiting, affect the development of fruiting. *Coprinopsis cinerea (C. cinerea)* is a model mushroom to study developmental processes in homobasidiomycetous fungi. The development of *C. cinerea* depends on the sensing of environment conditions, such as light, temperature, humidity and nutrients. The signal transduction pathway from the environmental condition sensors to the regulators and reproduction has been an important research topic.

Aims: This study investigated identified kinases which is involved in *C. cinerea* fruiting body development. Through the functional study of selected kinases and the putative regulatory kinase, we have a better understanding of the pathway between external stimuli and fruiting body initiation and development.

Methods: Small interfering RNA(siRNA) of different candidate genes, such as Cc.frg and Cc.neg, was applied to *C. cinerea* culture externally to transiently knockdown potential regulators for fruiting development of *C. cinerea*. With or without kinase inhibitor (LiCl) siRNA was applied to knockdown particular candidate genes under different unfavourable condition. The efficacy of siRNA was tested by real-time PCR assays. Results: The transient knockdown with Cc.frg siRNA cultures produced deformed fruiting body and some showed retarded fruiting body development. The culture with Cc.neg siRNA and without inhibitor under excessive glucose, on the other hands, continued fruiting body development gradually, when compared with the control.

Discussion: By gene knockdown, we confirmed that Cc.frg gene is involved in the fruiting body initiation and control further development of fruiting body. Besides, there was signal transduction pathway between Cc.frg and Cc.neg that regulates the fruiting of *C. cinerea*, as Cc.neg responded to nutritional stimuli.

eQTR: A Strategy to Identify the Association between RNA-Editing Events and Gene Expression—a Case Study in *Coprinopsis cinerea*

Chang, Jinhui* (tjchangjinhui@gmail.com)[1,2]; Kwan, Hoi Shan[1, 2]

[1]The School of Life Sciences, The Chinese University of Hong Kong, Hong Kong SAR, The People's Republic of China. [2]Food Research Centre, The Chinese University of Hong Kong, Hong Kong SAR, The People's Republic of China.

RNA-editing is a post-transcriptional process that introduces nucleotide changes in RNA different from its hereditary DNA template at specific sites. The transcriptome is modified by editing specific bases in a very quick manner, and then diversifies the proteome, through amino acid switching, alternative splicing, intron retention, modulating transcript stability, controlling translation and transport of the transcripts.

In this study, I have identified RNA editing events in the model mushroom *Coprinopsis cinerea* by in silico analyses, and predicted that some of these editing sites are involved in lignin degradation regulation. We have obtained RNA-seq gene expression profiles of 46 single spore isolates of *C. cinerea.* First, I compared the DNA and RNA sequence and called all RNA editing events. Then I developed an expression Quantitative Trait RNA editing sites (eQTR) strategy to identify regulators harboring RNA editing sites. eQTR represents a recoding base in the RNA different from its template DNA. Such variation between individuals could affect mRNA abundance, a quantitative gene expression trait. The basic model included the gene expression level as a response variable, with the RNA editing type as the predictor.

I have used the *C. cinerea* genome sequences and RNA sequences to call RNA editing sites, and characterized the RNA editing events, including function enrichment and genomic positions. Together with the gene expression profile from the same set of samples, some transcription factors were identified that link and may regulate CAZyme expression. To our knowledge, this would be the first functional study on RNA editing sites for roles in CAZymes production in mushroom forming fungi. Identification and functional characterization of regulators in CAZymes production will contribute to our understanding of the fungal lignocellulose degradation with potentials to benefit enzyme industry.

eQTL Mapping in *Populus* *(Short Talk)*

Chen, Jay* (chenj@ornl.gov)[1]; Yang, Yongil[1]; Muchero, Wellington[1]; Ranjan, Priya[1]; Weston, David J.[1]; Yang, Xiaohan[1]; Jawdy, Sara S.[1]; Gunter, Lee[1]; Lipzen, Anna[2]; Lindquist, Erika[2]; Barry, Kerrie[2]; Schmutz, Jeremy[2, 3]; Tuskan, Gerald A.[2, 3]

[1]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. [2]DOE Joint Genome Institute, Walnut Creek, CA 94598, USA. [3]HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA.

At the population level, phenotypic variation can be due to sequence polymorphism that produces altered or absent proteins (e.g., molecular variant) and/or due to differences in gene expression

that generate varying amounts of protein (e.g., expression trait). RNAseq applied to the progeny of a mapping population can identify the genetics underlying transcript variation, and further help determine phenotypic variation when combined with classical Quantitative Trait Locus (QTL) mapping. Similar to trait QTL mapping, when transcript levels in a given tissue are measured across a population of plants, the variation in mRNA abundance for each gene can be treated as a heritable trait. Expression QTL (eQTL) mapping can then locate and identify the genetic factors that control the observed variation in mRNA abundance. We performed analysis of RNAseq data on a total of 438 biological samples within the *Populus trichocarpa* × *Populus deltoides* pseudo-backcross pedigree. Among these samples, 124 samples have replicates. In total, the samples represent 312 unique genotypes of the QTL mapping pedigree and two parents. We found that among a total of 41,335 genes, the transcript of 37,588 genes can be detected in the developing xylems of at least one genotype. In average, a total of 26,923 genes are expressed in the developing xylem, representing approximately 65% of total genes in the *Populus* genome. Furthermore, we found that several hundreds of genes showing large transcript variation (>50 fold) within the mapping pedigree. We have identified over 10,000 cis-genetic and trans-acting genetic elements controlling transcript variation. Our studies also uncovered a number of unconventional transcriptional factors with no prior role in transcriptional regulation.

The Origin and Architecture of *Trichoderma* Hydrophobome

Chenthamara, Komal* (komal.chenthamara@tuwien.ac.at)[1]; Cai, Feng[1, 2]; Przylucka, Agnes[1]; Shen, Qirong[2]; Bayram Akcapinar, Günseli[1, 3]; Druzhinina, Irina[1]

[1]E166 - Institute of Chemical, Environmental and Biological Engineering, Research Area Biochemical Technology, Microbiology Group, TU Wien, Vienna, Austria. [2]Jiangsu Collaborative Innovation Center for Solid Organic Waste Resource Utilization, Nanjing Agricultural University, Nanjing, China. [3]Department of Statistics and Medical Informatics, School of Medicine, Acibadem University, Istanbul, Turkey.

A hyphal organization of the body in filamentous fungi provides a high surface area-to-volume ratio that is an essential adaptation for the effective extraction of nutrients while growing on or in (semi) solid substrates or a liquid. Hyphae are specially adapted for efficient attachment to a diversity of surfaces what is also required for an exertion of penetrative mechanical forces for the invasion of substrates and tissues. The development of many fungi includes interchanges between penetration of tissues/substrates for nutrition and growing out of them for dispersal. Consequently, filamentous fungi have unique molecular adaptations that evolved along with this distinctive lifestyle. Hydrophobins - the small amphiphilic surface-active extracellular proteins that are produced exclusively by filamentous fungi may provide a mechanistic understanding of such adaptations.

The diversity of hydrophobin-encoding genes varies in genomes of higher fungi while these proteins are not known from lower fungi. Some Basidiomycota fungi, in particular, those living in saline or xerophilic environments have enriched arsenals of hydrophobins that are required for their development and stress resistance, while the rest of them and the majority of Ascomycota have only a few hydrophobin-encoding genes. In this respect, the mycoparasitic genus *Trichoderma* (Hypocreales, Pezizomycotina), which also includes several cosmopolitan generalist species with high environmental opportunistic potential, is an exception as genomes of these fungi have expanded hydrophobomes (HFBomes). Previous studies have revealed that most of *Trichoderma* hydrophobins are the orphan genes that have no homologs in related organisms, while hydrophobins of other Pezizomycotina share their long evolutionary history.

In this study, we used the whole genomes and EST libraries for 11 *Trichoderma* and 8 other Hypocreales fungi to catalog HFBomes in *Trichoderma* and investigate the evolution of each protein. First, we tested the hypothesis on an operation of a purifying selection and 'birth-and-death' evolution of hydrophobins in this genus. Second, we found several cases of putative gene loss, gene duplication or horizontal gene transfer. The phylogenetic and phylogenomic analyses aided by the investigation of physical-chemical properties of individual hydrophobins such as the surface hydrophobicity, hydropathy plots, pI, and homology modeling, revealed at least 14 monophyletic clades containing active individual hydrophobin-encoding genes in the genus. It also allowed to differentiate plesiomorphic and propose apomorphic hydrophobins. Data mining performed for transcriptomes confirmed the functionality of majority of these genes across the genus. The analysis of the selection pressure revealed a highly complex pattern when individual genes in some species appeared to be under strong pressure of either purifying or positive selection while the respective homologous genes evolved neutrally in the other taxa. Comparisons with the specialized mycoparasitic hypocrealean fungus *Escovopsis weberi*, which has only a few hydrophobins, demonstrated that the expansion of HFBome in *Trichoderma* was closely linked to the evolution of its nutritional versatility. We propose that the entire HFBome contributes to the fitness of an individual species, while individual genes may emerge and get lost relatively quickly.

Communal Metabolism of Methane and the Rare Earth Element Switch

Chistoserdova, Ludmila* (milachis@uw.edu); Yu, Zheng; Krause, Sascha M. B.; Beck, David A. C.; Lidstrom, Mary E.

University of Washington, Seattle, WA, USA.

Background: Metabolism of methane is an important part of biogeochemical cycling of carbon. Methane is also a major contributor to climate change. A specialized group of microbes (the methanotrophs) that consume methane, gaining both energy and carbon from this chemically inert compound, represent a natural filter preventing an even faster accumulation of methane in the atmosphere. While methanotrophy has been studied for the past hundred years as a metabolic feature of individual pure cultures, a concept of communal function in methanotrophy has been gaining momentum. However, the mechanistic details are still missing of how and why the methanotrophs share their hard-earned carbon with other species, and whether and what they gain in return.

Methods: We utilized lake sediment methanotrophic community as a model. We manipulated complex natural communities using methane as the sole source of carbon, to determine species persisting in methane-consuming communities (the top-down approach). We also built synthetic communities of pure cultures of methanotrophs and non-methanotrophs and tested their behavior under a variety of environmental conditions (the bottom-up approach). We sequenced multiple (meta)genomes and (meta)transcriptomes to gain insights into the genomic potentials and gene expression patterns in relevant microbes.

Results: Through microcosm manipulation, using methane as the sole source of carbon, followed by metagenomic analysis, we identified key species active in methane consumption, the bacteria of the family Methylococcaceae. We further determined the primary and most abundant satellite types, the non-methanotrophic methylotrophic bacteria of the family Methylophilaceae. Two other persistent but less abundant types were identified as members of Burkholderiales and Flavobacteriales. Through manipulation of synthetic communities, followed by transcriptomic analysis, we identified at least one metabolic node at which community cross-talk takes place, the methanol oxidation step that involves

alternative methanol dehydrogenase enzymes, one requiring calcium as a cofactor, another requiring rare earth elements (REE), one of the first demonstrations of a biological function for this group of metals. Enzyme choice, in turn, appears to be determined by a number of environmental factors, such as oxygen and methane partial pressures, as well as sources of nitrogen.

Conclusions: We conclude that methanol must be the major carbon compound that the methanotrophs share with other community members, and that carbon flow is regulated by the REE switch, presenting an unexpected and unprecedented example for the important role of REEs in complex biological systems. Overall, our data shed new light on social lives of microbes involved in metabolism of methane in natural habitats and highlight some of the metabolic links among the community partners.

Full-Length cDNA Sequencing on the PacBio Sequel Platform

Clark, Tyson A.* (tclark@pacificbiosciences.com); Hon, Ting; Tseng, Elizabeth

Pacific Biosciences, Menlo Park, CA 94025.

The protein coding potential of most plant and animal genomes are dramatically increased via alternative splicing. Identification and annotation of expressed mRNA isoforms is critical to the understanding of these complex organisms. While microarrays and other NGS-based methods have become useful for studying transcriptomes, these technologies yield short, fragmented transcripts that remain a challenge for accurate, complete reconstruction of splice variants.

The Iso-Seq™ protocol developed at PacBio offers the only solution for direct sequencing of full-length, single-molecule cDNA sequences to survey transcriptome isoform diversity useful for gene discovery and annotation. Knowledge of the complete isoform repertoire is also key for accurate quantification of isoform abundance. As most transcripts range from 1 – 10 kb, fully intact RNA molecules can be sequenced using SMRT® Sequencing without requiring fragmentation or post-sequencing assembly. The PacBio Sequel platform has improved throughput thereby increasing the number of full-length transcripts per SMRT Cell. Furthermore, loading enhancements on the Sequel instrument have decreased the need for size fractionation steps. We have optimized the Iso-Seq library preparation process for use on the Sequel platform.

We applied the Iso-Seq method to the maize (*Zea mays*) inbred line B73. Full-length cDNAs from six diverse tissues were barcoded and sequenced across multiple size-fractionated SMRTbell libraries. A total of 111,151 unique transcripts were identified. More than half of these transcripts (57%) represented novel, sometimes tissue-specific, isoforms of known genes. In addition to the 2250 novel coding genes and 860 lncRNAs discovered, the Iso-Seq dataset corrected errors in existing gene models, highlighting the value of full-length transcripts for whole gene annotations.

Genome-Wide Identification of Bacterial Plant Colonization Genes

Cole, Benjamin J.* (bjcole@lbl.gov)[1]; Feltcher, Meghan E.[2]; Waters, Robert J.[1]; Wetmore, Kelly M.[3]; Mucyn, Tatiana S.[2]; Ryan, Elizabeth M.[1]; Wang, Gaoyan[1]; Ul-Hasan, Sabah[1,4];

McDonald, Meredith[2]; Yoshikuni, Yasuo[1]; Malmstrom, Rex R.[1]; Deutschbauer, Adam[3]; Dangl, Jeffery L.[2]; Visel, Axel[1]

[1]DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598. [2]Department of Biology, University of North Carolina, Chapel Hill, NC 27599; Howard Hughes Medical Institute, University of North Carolina, Chapel Hill, NC 27599; Curriculum in Genetics and Molecular Biology, University of North Carolina, Chapel Hill, NC 27599; Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, NC 27599; Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC 27599. [3]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA. [4]School of Natural Sciences, University of California Merced, 5200 N. Lake Road, Merced, CA, USA 95343.

Diverse soil-resident bacteria can contribute to plant growth and health, but the molecular mechanisms enabling them to effectively colonize their plant hosts remain poorly understood. We used randomly barcoded transposon mutagenesis sequencing (RB-TnSeq) in *Pseudomonas simiae,* a model root-colonizing bacterium, to establish a genome-wide map of bacterial genes required for colonization of the *Arabidopsis thaliana* root system. We identified 115 genes (2% of all *P. simiae* genes) whose function is required for maximal competitive colonization of the root system. Among the genes we identified were some with obvious colonization-related roles in motility and carbon metabolism, as well as forty-four other genes that had no or vague functional predictions. Independent validation assays of individual genes confirmed colonization functions for 20 of 22 (91%) cases tested. To further characterize genes identified by our screen, we compared the functional contributions of *P. simiae* genes to growth in 90 distinct *in vitro* conditions by RB-TnSeq, highlighting specific metabolic functions associated with root colonization genes. Our analysis of bacterial genes by sequence-driven saturation mutagenesis revealed a genome-wide map of the genetic determinants of plant root colonization and offers a starting point for targeted improvement of the colonization capabilities of plant-beneficial microbes.

The Effect of Drought and Host Selection on the Grass Root Microbiome

Coleman-Derr, Devin* (colemanderr@berkeley.edu)[1, 2]; Naylor, Daniel[1, 2]; DeGraaf, Stephanie[1]; Purdom, Elizabeth[1]

[1]University of California, Berkeley, CA, USA. [2]USDA-ARS, Albany, CA, USA.

Plant-associated microbial communities play crucial roles in determining the host phenotype. A subset of the plant-associated bacteria and fungi offer benefits to crop species through increasing nutrient and resource uptake efficiency, out-competing plant pathogens, and improving abiotic stress response. Despite much research, only a tiny fraction of plant microbiomes have been uncovered and evaluated, and many of the rules governing microbial community recruitment to the host microbiome remain unknown. The objectives of this project were to determine the general and species-specific plant microbiome responses to abiotic stress across the grasses, and to determine the correlation between host phylogenetic distance and phytobiome distance in a series of monocot species. Using 16S rRNA amplicon profiling of rhizosphere and root endospheres and phylogenetic trees based on the hosts' conserved chloroplast sequences, we have assessed the hypothesis that similar plant species share more similar microbiomes and uncovered conserved trends in root microbiome development in response to drought across the grasses. The identification of these trends may be of benefits to members of industry and academia alike that aim to improve crop performance through the use of biological control agents and other microbe-mediated means.

Phased Diploid Genome Assembly with Single Molecule Real-Time Sequencing

Concepcion, Gregory T.* (gconcepcion@pacificbiosciences.com)[1]; Chin, Chen-Shan[1]; Peluso, Paul[1]; Sedlazeck, Fritz J.[2]; Nattestad, Maria[4]; Clum, Alicia[5]; Dunn, Christopher[1]; O'Malley, Ronan[6]; Figueroa-Balderas, Rosa[7]; Morales-Cruz, Abraham[7]; Cramer, Grant R.[8]; Delledonne, Massimo[9]; Luo, Chongyuan[6]; Ecker, Joseph R.[6]; Cantu, Dario[7]; Rank, David R.[1]; Schatz, Michael C.[2,3,4]

[1]Pacific Biosciences, Menlo Park, CA, USA. [2]Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA. [3]Department of Biology, Johns Hopkins University, Baltimore MD, USA. [4]Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA. [5]DOE Joint Genome Institute, Walnut Creek, CA, USA. [6]Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, CA, USA. [7]Department of Viticulture and Enology, University of California, Davis, CA, USA. [8]Department of Biochemistry and Molecular Biology, University of Nevada, Reno, NV, USA. [9]Dipartimento di Biotechnologie, Universita' degli Studi di Verona, Verona, Italy.

While genome assembly projects have been successful in many haploid and inbred species, the assembly of non-inbred or rearranged heterozygous genomes remains a major challenge. To address this challenge, we introduce the open-source FALCON and FALCON-Unzip algorithms (https://github.com/PacificBiosciences/FALCON/) to assemble long-read sequencing data into highly accurate, contiguous, and correctly phased diploid genomes. We generate new reference sequences for heterozygous samples including an F1 hybrid of *Arabidopsis thaliana,* the widely cultivated *Vitis vinifera* cv. Cabernet Sauvignon, and the coral fungus *Clavicorona pyxidata,* samples that have challenged short-read assembly approaches. The FALCON-based assemblies are substantially more contiguous and complete than alternate short- or long-read approaches. The phased diploid assembly enabled the study of haplotype structure and heterozygosities between homologous chromosomes, including the identification of widespread heterozygous structural variation within coding sequences.

http://www.nature.com/nmeth/journal/v13/n12/full/nmeth.4035.html

A Mechanistic Understanding of Endophyte Mediated Plant Stress Responses

Connor, Elise R.* (eworchel@utexas.edu); Hawkes, Christine V.

University of Texas at Austin, Department of Integrative Biology, Austin, TX 78712.

Plant leaves harbor diverse communities of symbiotic fungi ('endophytes') that can affect plant physiology, growth, and stress responses. However, fungal effects are often context-dependent with both biotic and abiotic factors shifting effects from harmful to beneficial. Here, we present on two factors affecting the outcome of plant-endophyte symbioses: (1) intraspecific interactions among multiple endophytes simultaneously colonizing a plant and (2) environmental stress. We also examine mechanisms underlying fungal effects on plants, including niche overlap among interacting fungi and fungal induced metabolites within plants. To examine these mechanisms, we inoculated and grew Panicum grasses with 16 two-species endophyte mixtures composed of fungi with high or low niche overlap, monocultures of each fungus within a mixture, or a fungus-free inoculum in well-watered or drought conditions. Niche overlap among fungi was based on a wide range of traits in culture that

putatively relate to behavior in symbiosis, including fungal resource acquisition, stress tolerance, enzyme activity and metabolite production. Fungal induced metabolites within plants were identified from whole plant extracts from liquid chromatography coupled to time of flight mass spectrometry.

We found that plant-endophyte symbioses are highly context-dependent, with approximately 50% of the fungal pairs causing non-additive effects on plants. Interactions were both synergistic and antagonistic, in which fungal pairs increased or decreased plant performance relative to the corresponding monoculture inoculums. Additionally, dissimilarity in fungal stress tolerance and metabolite production was the best predictor of fungal interactive outcomes in the plant. We also identified fungal induced metabolites within plants that were unique to beneficial and negative symbiosis outcomes, suggesting that chemical mechanisms underlie fungal effects. Furthermore, many of the target metabolites were identified as important plant-microbe signaling molecules, further revealing a mechanism underlying plant-fungal symbiosis. These results are key to gaining a mechanistic understanding fungal interactions with the plant.

The Estuarine Metagenome of Microbial Communities in the Columbia River estuary

Crump, Byron C.* (bcrump@coas.oregonstate.edu)[1]; Fine, Lindy M.[3]; Herfort, Lydie[2]; Smith, Maria W.[2]; Payet, Jérôme P.[1]; Simon, Holly M.[2]

[1]Oregon State University, Corvallis, OR. [2]Oregon Health and Sciences University, Portland, OR. [3]University of Maryland Center for Environmental Science, Cambridge, MD.

Estuaries act as coastal filters for fluvial materials in which microbial, biogeochemical, and ecological processes combine to transform riverborne organic matter prior to export to the coastal ocean. This critical function of estuarine 'bioreactors' is linked to material residence times and is based on heterotrophic activity of unique estuarine microbial communities. However, little is known about the diversity of estuarine microbes and the genomic capabilities that allow them to thrive under estuarine environmental conditions. Here we present a seasonal study of bacterial production, microbial diversity (16S rRNA amplicon sequencing), and metagenomics in the Columbia River estuary, which, like many estuaries, supports highly active estuarine microbial communities on suspended particles that are retained in estuarine turbidity maxima (ETM). ETM particle trapping in this rapidly flushed estuary extends particle residence time from 1-2 days (the average water residence time) to several weeks, which facilitates organic matter transformations by microbes, and permits establishment of discrete and metabolically dominant particle-associated microbial communities. Size-fractionated samples were collected from several locations in the estuary in spring, summer, and fall 2012-2013. We identified diverse and seasonally variable estuarine microbial communities associated with ETM particles. These communities differed from river and ocean communities and included taxa that wash into the estuary from potential microbial seedbanks in four shallow lateral bays. Comparative metagenomics were used to identify "estuarine" genes, metabolic processes, and organismal capabilities that distinguish estuarine communities from river and ocean communities. These results provide important insights into the ecology, metabolism and ecosystem functions of estuarine microbial communities.

Viral Control of Halanaerobium Populations in Hydraulically Fractured Shales

Daly, Rebecca A.* (daly.130@osu.edu)[1]; Borton, Mikayla A.[1]; Johnston, Michael D.[2]; Wilkins, Michael J.[1, 2]; Wrighton, Kelly C.[1]

[1]Department of Microbiology, The Ohio State University, Columbus, OH. [2]School of Earth Sciences, The Ohio State University, Columbus, OH.

Hydraulic fracturing, the extraction of oil and gas from deep shale reservoirs via high-pressure injection of fluids, creates a novel ecosystem in the terrestrial deep subsurface. This process inadvertently injects surface microorganisms ~2,500 meters deep into the subsurface, exposing microorganisms to high pressures and brine-level salinities. *Halanaerobium* initially constitute <1% of the introduced microbes, but after ~50 days eventually dominate, up to 99% of the microbial community. The *Halanaerobium* population typically consists of multiple *Halanaerobium* strains that are dynamic through time; with diverse, lower abundance genotypes that persist along with the dominant *Halanaerobium* and occasionally become enriched. To identify the traits that support strain-level *Halanaerobium* heterogeneity, we isolated multiple *Halanaerobium* strains from produced fluids, and binned *Halanaerobium* from time-series metagenomic data collected from multiple, hydraulically fractured wells in Appalachian Basin shales. From our metagenomic datasets we also recovered 1,768 viral contigs across five wells, only 3% of which were shared between wells. Genomic analysis of the *Halanaerobium* isolates and metagenomic bins highlights key differences in the number of prophage and CRISPR-Cas mediated viral immunity despite high (~99.9%) average nucleotide identities. Laboratory experiments with isolates demonstrated that some *Halanaerobium* strains contain prophage that can be induced using stressors such as decrease in pH, metabolite accumulation and increases in heavy metal concentrations, lysing up to 58% of cells. Analysis of field data revealed that dramatic changes in strain abundances (66% decrease, 84% increase) occur concomitantly with viral blooms (66% of viral relative abundance). We hypothesize that viral predation opens niche space in this ecosystem, facilitating drastic changes in strain relative abundances. Extensive CRISPR-Cas spacer links between viruses and *Halanaerobium* strains indicate active viral predation in this ecosystem. For multiple *Halanaerobium* strains from metagenomic samples, we were able to find evidence for the temporal incorporation of new spacer sequences during the sampled time span within a well, suggesting that some strains can evade viral predation through acquired CRISPR-Cas immunity *in situ*. Together these data suggest that interactions between viruses and their *Halanaerobium* hosts are a primary driver of population dynamics in the ecosystem created by hydraulic fracturing in the deep terrestrial subsurface.

Activity of Novel Methanogenic and Methanotrophic Taxa in Freshwater Wetland Surface Soils

Daly, Rebecca A.* (daly.130@osu.edu)[1]; Angle, Jordan C.[1]; Smith, Garrett J.[1]; Solden, Lindsey M.[1]; Narrowe, Adrienne B.[2]; Morin, Tim H.[3]; Borton, Mikayla A.[1]; Hoyt, David W.[4]; Miller, Christopher S.[2]; Bohrer, Gil[3]; Wrighton, Kelly C.[1]

[1]Department of Microbiology, The Ohio State University, Columbus, OH. [2]Department of Integrative Biology, University of Colorado Denver, Denver, CO. [3]Department of Civil and Environmental Engineering and Geodetic Sciences, The Ohio State University, Columbus, OH. [4]EMSL, Pacific Northwest National Laboratory, Richland, WA.

Temperate freshwater wetlands are currently the greatest source of the potent greenhouse gas methane. Predicting methane emissions from wetlands is highly uncertain, partially confounded by limited knowledge about the physiology of organisms mediating methane cycling. To better understand these physiologies in freshwater wetland soils we sampled microbial communities from Old Woman Creek, a wetland adjacent to Lake Erie. We examined soils underlying three land-coverage types (vegetation, seasonal mud-flats, and continuously water flooded channels). Porewater gas measurements quantified methane dynamics monthly at 2.8 cm depth intervals and indicated that oxic surface soils frequently produced an order of magnitude more methane than did anoxic soils. Similarly, methanogen activity (mcrA transcript abundance via qPCR) was 9-fold greater in oxygenated surface soils as compared to deep, anoxic soils, and correlated positively to *in situ* porewater methane and acetate concentrations. To examine the metabolism of organisms fueling this unexpected methane cycle in oxic soils, we performed metagenomic and metatranscriptomic sequencing. We reconstructed multiple near complete (>80%) methanogen and methanotroph genomes, which constituted a new methanogen species (*Ca.* Methanosaeta oxydurans) and a new methanotroph genus within the Methylococcaeae (*Ca.* Methylosolum). Key functional genes (mcrA and pmoA) from these two organisms were constantly in the top 3% of transcripts from the soil surface microbial community, and represented the most active methanogens and methanotrophs in these soils across seasons and land-coverage types. Our findings demonstrate that new methanogen and methanotrophs, with the genomic capacity to withstand unexpected redox conditions, mediate active methane cycling in surface, oxic wetland soils. These findings challenge current assumptions about methane cycling metabolisms that are currently incorporated into global biogeochemical models.

Visualization and Exo-Proteomics of Plant Biomass Colonization by *Aspergillus*

de Vries, Ronald P.* (r.devries@westerdijkinstitute.nl)[1]; Khosravi, Claire[1]; Kowalczyk, Joanna[1]; Purvine, Samuel[2]; Dohnalkova, Alice[2]; Chrisler, William B.[2]; Orr, Galya[2]; Robinson, Errol[2]; Zink, Erika[2]; Wiebenga, Ad[1]; Battaglia, Evy[1]; Benoit, Isabelle[1]; Baker, Scott E.[2]

[1]Fungal Physiology, Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology, Utrecht University, Utrecht, The Netherlands. [2]EMSL, Department of Energy, USA.

Aspergillus niger is an ascomycete fungus able to secrete multiple extracellular enzymes that break down plant polymers. This ability has been extensively researched by scientists and industrial companies for many industrial applications. The production of these enzymes is tightly regulated by several regulators, such as the carbon catabolite repressor CreA and the (hemi-)cellulolytic activators XlnR and AraR. While the effects of regulator deletions on the production of the enzymes has been studied, how this affects the colonization and degradation of plant biomass by A. niger has so far not been studied in detail.

In this study we have addressed this topic by detailed visualization of the colonization and degradation of wheat bran by A. niger wt and mutants for the regulators mentioned above using high-tech microscopy of the DOE-EMSL institute (Helium Ion Microscope (HIM), ETEM, liquid cryoTEM, super resolution and standard confocal fluorescence microscopy, FISH). To explore the molecular basis for the differences observed between the strains we also performed exo-proteome analysis on these cultures. This way we could link the reduction in degradation efficiency of the mutants to the absence of subsets of enzymes. This study for the first time unearths the correlation between enzyme

production and substrate colonization and reveals which enzymes play an essential role in this process. Highlights from this study will be presented.

Dissecting Intraspecies Diversity in Fungal Wood Decay

de Vries, Ronald P.* (r.devries@westerdijkinstitute.nl)[1, 3]; Casado Lopez, Sara[1]; Purvine, Samuel[2]; Dohnalkova, Alice[2]; Orr, Galya[2]; Robinson, Errol[2]; Zink, Erika[2]; Baker, Scott E.[2]; Hildén, Kristiina S.[3]; Mäkelä, Miia R.[1, 3]

[1]Fungal Physiology, Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology, Utrecht University, the Netherlands. [2]Environmental Molecular Sciences Laboratory, Department of Energy, USA. [3]Department of Food and Environmental Sciences, University of Helsinki, Finland.

Dichomitus squalens is a white rot fungus that colonizes mainly softwood and is commonly found in the northern parts of Europe, North America and Asia. In nature, compatible monokaryons of *D. squalens* cross sexually to form a dikaryon. At this point it is not clear whether monokaryons can efficiently colonize and degrade wood, or whether this only occurs after the dikaryon is formed. Here we studied in detail the process of wood colonization, degradation and utilization by *D. squalens,* using a combination of transcriptomics, proteomics, metabolomics and high-resolution microscopy enabled by the combined facilities of the JGI and EMSL institutes.

In this study, we compared a set of related mono- and dikaryotic strains, originating from two isolates. All strains were grown on solid spruce sticks and sampling was performed after two and four weeks to evaluate the variation between the strains during the wood decay process. Monokaryons, compared to dikaryons, were notably slower in colonizing wood, which was reflected in the analysis of the various datasets obtained from the samples.

Studying the Genetics and Ecophysiology of Perenniality Using the Model Grass Genus *Brachypodium*

Des Marais, David L.* (desmarais@fas.harvard.edu)[1]; Catalan, Pilar[2]; Contreras-Moreira, Bruno[3]; Edger, Patrick P.[4]; Preston, Jill C.[5]; Woods, Daniel P.[6]; Amasino, Richard M.[6]

[1]Harvard University. Cambridge, MA, USA. [2]Universidad de Zaragosa. Huesca, Spain. [3]Estacion Experimental Aula Dei. Zaragosa, Spain. [4]Michigan State University. East Lansing, MI, USA. [5]University of Vermont. Burlington, VT, USA. [6]University of Wisconsin at Madison, Madison, WI, USA.

All plants must allocate limited resources to survival, growth, and reproduction. Decisions about allocation represent trade-offs between survivorship risk and subsequent fitness benefits with two extremes: annual species, which reproduce once and then die, and perennial species, which reproduce over multiple seasons, often with interim periods of quiescence. Whereas many cereal crops and most model plant species are annuals, most forage and cellulosic biofuel crops are perennials, for which maximizing vegetative production over multiple years is desirable. The grass genus *Brachypodium* is comprised of both annual and perennial species; one of these, the annual B. distachyon a well-developed model genomic system for studying the genetic, developmental, and physiological basis of biofuel feedstock production. In collaboration with JGI, we have begun a series of comparative genomic analyses with the broad aim of identifying transcripts and putative regulatory

control elements that are associated with the evolutionary transition between annual and perennial life history strategies. We will accomplish this via *de novo* sequencing of perennial *Brachypodium* species and a new outgroup species to facilitate phylogenomic analysis, deep resequencing of natural accessions of B. sylvaticum, and an RNASeq experiment assessing transcriptional control of meristem transitions in annual and perennial species of *Brachypodium*. In addition to allowing us to test hypotheses related to the environmental and genetic control of biomass partitioning, the new sequence data generated will provide an unprecedented set of community resources for comparative genomic analysis in the grasses.

Development of Luminescent Biosensors for Analysis of Bacterial Cyclic-di-GMP

Dippel, Andrew B.* (adippel@berkeley.edu)[1]; Evans, Robert[2]; Deutsch, Samuel[2]; Hammond, Ming[1]

[1]Chemistry Department, University of California, Berkeley, CA. [2]DOE Joint Genome Institute, Walnut Creek, CA.

The second messenger cyclic-di-GMP (cdiG) is a ubiquitous bacterial signaling molecule that regulates a large variety of processes, including virulence, motility, cell-cycle progression, and cell-cell communication. Despite the obvious importance of cdiG signaling in bacteria, our understanding of cdiG signaling networks is often incomplete, largely due to the fact that current assays for measuring cellular cdiG concentrations are laborious and limited in throughput. To develop a generalizable, high-throughput method for measuring cdiG levels in cells, we have created the first luminescent biosensor for cdiG, YNL-YcgR. By providing an intensity-based, luminescent signal (as opposed to fluorescence or FRET-ratio-based signals), YNL-YcgR can be used to measure cdiG levels in complex cell lysates that have high autofluorescence. We have successfully used this first generation YNL-YcgR biosensor to assay cdiG levels in *E. coli* lysates in a plate-reader format, but sensitivity of the assay is limited by its affinity (KD ~ 615 nM). We have developed a semi-high-throughput biosensor screening platform to improve the sensitivity of YNL-YcgR, and in collaboration with JGI we are screening a phylogenetic library containing ~100 different YcgR-like sequences in order to optimize the sensor's stability, affinity, signal-to-noise ratio. In the future we hope to use the optimized biosensor to measure cdiG levels of bacteria in response to various external signals and in co-cultures to help further our understanding of complex cdiG signaling networks.

Functional Diversity of Microbial Communities Associated with Deep-Sea Hydrothermal Deposits

Donaho, John A.*[1]; Reysenbach, Anna-Louise l. (bwar@pdx.edu)[1]; Kelley, John F.[1]; Flores. Gilberto E.[2]; Rhoades, Nicholas[2]; Seewald, Jeffrey S.[3]

[1]Portland State University, Biology Department, Portland, OR, USA. [2]California State University Northridge, Biology Department, Northridge, CA, USA. [3]Woods Hole Oceanographic Institution, Woods Hole, MA, USA.

Actively forming deep-sea hydrothermal vent sulfide deposits are rapidly colonized by a diversity of microbes. Many of these thermophilic Bacteria and Archaea obtain their energy from the geochemical gradients resulting from the hydrothermal fluid mixing with seawater. The microbial communities can

be in part predicted based on the geochemistry of the fluid and the type of deposit. The Eastern Lau Spreading Center (ELSC) vent fields provide excellent natural laboratories for exploring the factors that influence the diversity and relationships of microbial communities associated with actively forming deep-sea hydrothermal deposits. Here the ESLC provides large and systematic changes in fluid and rock geochemistry, spreading rate, magmatic/tectonic processes, and proximity to the volcanic arc over its relatively short length of about 397 km. Using 16S rRNA gene amplicon sequencing several interesting patterns in the microbial communities emerged. For example, the southern andesite-hosted vent field, with low pH, high CO_2, high Fe vent fluids was dominated by anaerobic Bacteria and Archaea, while the more northern basalt hosted vent field deposits were characterized with a diversity of microaerophilic to anaerobic thermophiles.

To explore the functional relationships within the microbial communities, we assembled the metagenomes of 10 different samples from these sites using MegaHit. Core functional gene patterns were explored using HMMER searches of 46 diagnostic genes. The contigs were binned using MetaBat and initial % completeness was checked using CheckM. The preliminary assemblies and binning confirmed our previous 16S rRNA gene amplicon data that the ELSC is a diversity hotspot for novel Crenarchaeota and in particular the Thermoprotei. For example, *Thermofilum* has never been cultured from deep-sea vents but our 89% complete draft *Thermofilum* genome from one vent field showed that it had relatively low similarity with published Thermofilum genomes. Furthermore, we obtained several almost complete (89%, based on CheckM) draft korarchaeotal genomes, and these are more divergent from the terrestrial Korarchaeota genomes than the terrestrial korarchaeotal genomes are from each other. Like their terrestrial relatives, the marine Korarchaeaota appear to have an anaerobic lifestyle by fermenting peptides. Furthermore, the metagenomes confirmed the importance of sulfur and nitrogen cycling at all vent fields.

The Genome of the Perennial Grain Crop *Thinopyrum intermedium* Provides Tools for Breeding and Unravels the Evolutionary History of the Genus

Dorn, Kevin M.* (dorn@ksu.edu)[1]; Kantarski, Traci[1, 2]; Crain, Jared[1]; Altendorf, Kayla[3]; Zhang, Xiaofei[3]; Anderson, James[3]; Wang, Shuwen[4]; Larson, Steve[5]; Plott, Christopher[6]; Jenkins, Jerry[6]; Grimwood, Jane[6]; Schmutz, Jeremy[6]; DeHaan, Lee[4]; Poland, Jesse[1]

[1]Kansas State University, Manhattan, Kansas, USA. [2]Columbia University, New York, New York, USA. [3]University of Minnesota, St. Paul Minnesota, USA. [4]The Land Institute, Salina, Kansas, USA. [5]USDA-ARS Forage and Range Research Lab, Logan, Utah, USA. [6]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA.

Thinopyrum intermedium (intermediate wheatgrass - IWG) is a perennial grass that has been targeted for domestication and improvement due to its prolific biomass production, relatively large seed size, and moderate grain yield. As a crop, IWG provides essential ecosystem services by stabilizing soil health with its deep (>3 meters) root structure and reducing nutrient runoff, all while producing biomass quantities similar to switchgrass and >1,000 kg/ha of grain that is similar to wheat in food products. Over the past decade, IWG has been dramatically improved with traditional breeding methods, focusing on traits relating to seed size/yield, free threshing seed, and early maturity. Accelerating these programs is possible with genomics-assisted breeding. There have been limited molecular breeding and genomic tools for IWG due to its large allohexaploid genome (2n = 6x = 42) and 1C genome size of 12.6 Gb. We have sequenced and assembled the genome of a single haploid plant using a combination of strategies, including *de novo* assembly with the NRGene DeNovoMagic pipeline, scaffolding with 10X Genomics reads, and anchoring/ordering using F1 population

sequencing. We have preliminary pseudochromosome assemblies completed that will be annotated this year. Using a series of genotype-by-sequencing experiments of putative diploid progenitor species, we have assigned 21 chromosomes to seven homoeologous groups. With these data, we have produced the first whole genome, sequence-based analysis of the diploid progenitor species for IWG and have proposed a new genomic designation for the species. The genome sequence is currently being used to assist genomic selection and several GWAS trials to increase the rate of genetic gain in IWG.

Functional Analysis of the Microbiome of *Populus:* Identification, Localization, and Characterization of Key Endophytes

Doty, Sharon L.* (sldoty@uw.edu)[1]; Firrincieli, Andrea[2]; Joubert, Pierre M.[1]; Sher, Andrew W.[3]; Tournay, Robert[1]; Khorasani, Mahsa[1]; Deutschbauer, Adam[5]; Ray, Jayashree[5]; Joseph, Matthew[1]; Scharffenberger, Sam[1]; Frank, Carolyn[4]; Huffaker, Alisa[3]; Mugnozza, Giuseppe[2]; Kim, Soo-Hyung[1]

[1]University of Washington, Seattle, WA, USA. [2]University of Tuscia, Tuscia, Italy. [3]University of California, San Diego, CA, USA. [4]University of California, Merced, CA, USA. [5]Lawrence Berkeley National Laboratory, Berkeley, CA, USA.

Just as research on the human microbiome has demonstrated the profound importance of our microbiota on human health, plants are also strongly influenced by the ecosystem within them. To improve the environmental and economic sustainability of biomass production, it is essential that the biological interactions between plants and associated beneficial microbiota be more fully understood. Poplar (*Populus*) trees are an early successional pioneer plant species able to colonize nutrient-limited, cobble-dominated riparian zones. Native poplar plants have a diverse microbiota including strains that can fix dinitrogen gas, and promote plant growth and health under abiotic stresses including nutrient limitation and drought. Our lab demonstrated using the 15N2 incorporation assay that N2 is fixed at high levels in wild poplar by endophytes, the microorganisms that live within plants. Using specific pollutant-degrading strains, we have demonstrated the effectiveness of endophyte-assisted phytoremediation of major environmental pollutants. Building upon the strong body of evidence that endophytes from poplar can improve the growth and stress tolerance of a broad range of plant species including eudicots, monocots, and conifers, we have three JGI projects that seek to more fully understand the plant-microbe interactions involved in endophytic symbiosis. These projects use the strengths of the JGI and EMSL to 1) compare the microbiome of wild poplar under contrasting environmental stresses to identify potential core species, 2) identify and characterize the key N-fixing strains from wild poplar using FISH and 15N labeling, and 3) use random barcoded TnSeq to elucidate the endophyte genes required for N-fixation, salt tolerance, TCE metabolism, arsenic tolerance, and plant colonization. Our ultimate goal is to optimize endophyte technologies towards increasing biomass yields on marginal or contaminated lands with reduced inputs.

http://depts.washington.edu/envaplab/

Novel, High-Severity Fire Influences on Soil Microbial Communities and Biogeochemical Processes: Opening the "Charcoal" Box

Dove, Nicholas C.* (ndove@ucmerced.edu)[1]; Hart, Stephen C.[2, 3]

[1]Environmental Systems Graduate Group, University of California, Merced. [2]Life & Environmental Sciences, University of California, Merced. [3]Sierra Nevada Research Institute.

Temperate coniferous forest soils store 102 Gt of carbon (C), constituting a major global C sink. However, forest fires in these ecosystems can be a major factor controlling soil C dynamics. While soil C is immediately released to the atmosphere through combustion, fire also changes the abiotic conditions of the soil (e.g., moisture, temperature, nutrients) that can impact microbial communities, decomposition, and other microbial processes (e.g., nitrogen cycling) influencing the longer-term stability of soil C. Globally widespread wildfire suppression practices over several decades in these ecosystems have led to massive accumulations of forest fuel; this management strategy has changed the fire regime from one of historically frequent, low-severity surface fires, to a regime characterized by infrequent, but stand-replacing canopy fires. This changing fire regime likely exacerbates the influence of fire on microbial communities and C cycling processes in these ecosystems. Despite the importance of the microbial community in decomposition, we know of no studies that have simultaneous investigated the long-term (i.e., > 40 y) metagenomic effects of novel, high-severity fire and quantified the associated change in C cycling processes in temperate forest ecosystems. To assess the long-term influence of novel, high-severity fire on microbial communities and C cycling processes, we sampled four high-severity burned areas of varying time since fire (3- 43 y; i.e., a fire chronosequence) and unburned controls along the South Fork of the American River within the Eldorado National Forest (CA, USA). We are pairing traditional biogeochemical assays (e.g., extracellular enzyme activity, basal respiration, net N mineralization, nitrification potential, and microbial biomass) with metagenomics and iTag characterization of the microbial community, to elucidate changes to the microbial community and C cycling processes with time since novel, high-severity fire. Preliminary results show an increase in phosphatase activity with time since fire, suggesting phosphorus (P) availability decreases and microbial P demand increases with ecosystem recovery. We hypothesize that microbial respiration of soil C will be lowest immediately after fire due to heat-induced mortality reducing microbial abundance. However, we also predict that microbial respiration will peak within a decade since fire due to increased nutrient availability, while eventually subsiding due to nutrient limitation. Additionally, we hypothesize that these changes will correspond with distinct microbial communities, which ultimately control microbial function.

Ecophysiology of Fungi from the Habitats of Borneo's 'Exploding Ants'

Druzhinina, Irina S.* (irina.druzhinina@tuwien.ac.at)[1]; Rahimi, Mohammad Javad[1]; Grujic, Marica[1]; Forte, Ana[1]; Pourmehdi, Shadi[1]; Pretzer, Carina[1]; Kopchinskiy, Alexey[1]; Chenthamara, Komal[1]; Fischer, Michaela[2]; Laciny, Alice[3]; Abu Salim, Kamariah[4]; Lim, Linda[5]; Chin Mei, Chan[5]; Parich, Alexandra[2]; Schuhmacher, Rainer[2]; Zettel, Herbert[3]

[1]Vienna University of Technology, Institute of Chemical Engineering, Research Area Biotechnology and Microbiology, Gumpendorferstraße 1a, 1060 Vienna, Austria. [2]University of Natural Resources and Life Sciences, Vienna (BOKU), Department of Agrobiotechnology, Center for Analytical Chemistry, Konrad-Lorenz-Straße 20, 3430 Tulln, Austria. [3]Natural History Museum Vienna, 2nd Zoological Department, Burgring 7, 1010 Vienna,

Austria. [4]Universiti Brunei Darussalam, Environmental and Life Sciences, Jalan Tungku Link, Bandar Seri Begawan BE 1410, Brunei Darussalam. [5]Universiti Brunei Darussalam, Chemistry Programme, Jalan Tungku Link, Bandar Seri Begawan, BE 1410, Brunei Darussalam.

The understudied carpenter ants in the *Colobopsis cylindrica* complex ('COCY', or exploding ants, Hymenoptera: Formicidae) dominate arboreal habitats in the rainforests of Borneo. They have evolved a remarkable behavior: In territorial combat with enemy ants and other arthropods they sacrifice themselves by rupturing (autothysis) and releasing sticky and irritant contents of their hypertrophied mandibular glands to kill rivals. Voluntary self-sacrifice is very rare in nature, undoubtedly due to attendant fitness losses. Contrary to autothysis in termites and honeybees, workers of COCY ants that forage solitarily, explode during one-on-one confrontations far from their nests. Thus, they are defending the territory against potential competitors probably for continuously renewing food resources such as phyllosphere microbes or other organisms in this habitat.

The phyllosphere is an ecosystem that hosts a vast and diverse microbial community. Such microbiomes are exceptionally rich in canopies of tropical rain forests where trees do not shed leaves annually. The adaptation of epiphytic and endophytic microorganisms to their habitat requires the development of such specialized functions as an efficient attachment to the leaf surface, resistance to oxidative stress and survival in the oligotrophic environment. On Borneo COCY ants possess hypertrophied mandibular gland reservoirs containing a diversity of chemical compounds, many of which have fungicidal activity. To solve the global and COCY-derived ecological challenges, phyllosphere microorganisms produce a diversity of secondary metabolites, secrete enzymes that hydrolyze lignocellulose and cutin and adapt to stress.

In this study, we focused on the ecophysiological properties of over 330 filamentous and unicellular fungi isolated from natural nests and canopy foraging grounds of the two distinct COCY species. DNA barcoding with the use of the primary (ITS) and taxon-specific secondary loci resulted in detection of over 100 phylotypes. DNA barcodes of only 70% of isolates had similar hits in public databases (>96% similarity) what allowed molecular identification on infrageneric levels. Molecular identification of 15% of isolates was uncertain, and the other 15% of strains were only attributed to high taxonomic ranks. Although many generalist fungi such as *Trichoderma* spp. and *Xenoacremomiun* sp., *Penicillium* spp., Mucoromycotina and members of Saccharomycetales were isolated, many fungi belonged to plant pathogens or had unique physiology. For example, the microbiome of BBQ COCY ants that are chemically similar to true 'exploding ants' but rarely exhibit autothysis, was dominated by a novel dimorphic melanised 'black yeast' fungus from Pleosporales that grew in its yeasty form *in situ* but formed mycelium *in vitro*. This slowly growing fungus has a remarkable antifungal potential and has a novel putatively nitrogen-fixing bacterial associate. The majority of isolated fungi were able to degrade natural polymers such as cellulose and cutin. However, the leaves and branches were healthy with less than 1% of their surface damaged with lesions due to fungal diseases. This preliminary study suggests the positive impact of COCY ants on the fitness of their host plant, while the details of ant-plant-microbe interactions remain to be elucidated. Interestingly, several phyllosphere fungi are capable to degrade polycaprolactone, a bidegradable plastic used in biomedical industry and construction.

www.cocy.tuwien.ac.at

The Plant Root-Associated Microbiome Viewed across Diverse Plant Species and Developmental Stages

Edwards, Joseph A.* (edwards@ucdavis.edu)[1]; Kilmer, John[2]; Santos, Christian[1]; Liechty, Zachary[1]; Nguyen, Bao[1]; Phillips, Gregory[2]; Sundaresan, Venkatesan[1]

[1]Department of Plant Biology, University of California, Davis. [2]Arkansas State University.

Soil microbes are important mediators of plant nutrition, growth, and resistance to biotic and abiotic stresses. Plants acquire root-associated microbiomes that are distinct from that of the surrounding soil. We have previously reported[1] that for cultivated rice, the root-associated microbiome has a relatively small dependence on the host plant genotype. We have now investigated the effect of wide interspecies host variation on assimilated root-associated microbiomes within a single environment. We find that host microbiome variation does not reflect host plant phylogenetic relationships and that rice plants enrich for a unique rhizospheric and endospheric microbiome compared to other host plants growing in the same environment. A comparison of the root-associated microbiomes of rice with microbiota of soil from cultivated fields and non-cultivated soils, suggests that rice cultivation changes the overall soil microbiome in rice fields not only through agricultural inputs and water submergence, but also through specific enrichment of microbial taxa by rice plants. We have also studied variations in the microbiome through the lifecycle of rice plants under field conditions. We find that the composition of the microbiome shifts throughout vegetative growth until the initiation of reproductive growth, whereupon the microbiome stabilizes. These results suggest that plants select for different microbiota as they undergo developmental changes from germination to reproduction.

[1]Edwards, Joseph, et al. "Structure, variation, and assembly of the root-associated microbiomes of rice." *Proceedings of the National Academy of Sciences* 112.8 (2015): E911-E920.DOI:10.1073/pnas.1414592112

Dynamics of Lignin-Degrader Communities in a Dry Seasonal Tropical Forest

Egerton Warburton, Louise M.* (louise.egerton08@gmail.com)[1, 2]; Ross, Jessica[2]

[1]Chicago Botanic Garden, Glencoe, IL, USA. [2]Northwestern University, Evanston, IL, USA.

Tropical forests account for a large proportion of global terrestrial C cycling and storage. Within these forests, saprophytic fungi are considered key transformative agents in litter decomposition and C cycling owing to their collective abilities to rapidly decompose abundant biopolymers (lignin, cellulose) using a diverse set of hydrolytic and oxidative mechanisms. The importance of saprophytic fungi in C-cycling processes raises the question of how these fungi, and subsequent C fluxes, could become increasingly vulnerable with the predicted reductions in mean annual precipitation. To provide insights into fungal dynamics in a drier climate, we conducted a three-year study of lignin-degrading fungi in a seasonally dry tropical forest (SDTF; Yucatán, MX) during wet (~1200mm precipitation yr-1) and dry-seasons (<200 mm yr-1). This strong seasonality is expected to affect soil C cycling because the alternation between oxygenated (dry) and anaerobic soil conditions (wet) can alter the composition and functioning of soil and litter fungi.

Traps containing a standardized lignin substrate were buried in the litter layer at nine locations. At the end of each dry and wet season, the residuum was removed, traps were re-packed with fresh substrate and reburied, and the residues were analyzed using high throughput sequencing with ITS

barcoded primers. We detected 98 fungal taxa of which 82 were putative lignin degraders. There was a high turnover of fungal taxa from year-to-year because certain taxa were restricted to a single sampling period (e.g. Dothidiales in 2014). Unique fungal assemblages were also detected in dry- versus wet-seasons. We found a significantly higher abundance of taxa affiliated with the Polyporales (e.g. Phlebia) and Magnaporthales (Mycoleptodiscus) in the dry season, whereas members of the Dothidiales (Aureobasidium), Capnodiales (Cladosporium), Trechisporales (Trechispora) were more abundant during the wet-season. Ectomycorrhizal (ECM) fungi were also recovered in both the wet (Laccaria, Inocybe) and dry season samples (Thelephoraceae, Scleroderma). However, many fungi were recovered in similar abundances in both wet and dry seasons (Eurotiales; Pleosporales). While these taxa all possess genes encoding for peroxidases and/or laccases, only the Polyporales and Scleroderma are known to endure extremely negative soil water potentials. Thus, we inferred that fungal degrader communities in the SDTF might respond to future reductions in precipitation with weaker inter-annual variations, and a shift towards communities with greater abundances of saprophytic and ECM taxa tolerant of drier conditions. How these changes feedback to litter decomposition and C cycling is being tested in a manipulative field experiment.

Identifying Shifts in Lake Sediment Microbial Communities in Response to Carbon Biogeochemistry in Stordalen Mire, Sweden

Emerson, Joanne B.*[1]; Marcus, Tamara S. (tsm1004@wildcats.unh.edu)[2]; Varner, Ruth K.[2]; Parks, Donovan H.[3]; Johnson, Joel E.[2]; Neumann, Rebecca B.[4]; Woodcroft, Ben J.[3]; Singleton, Caitlin M.[3]; Tollerson, Rodney[1]; Wik, Martin[5]; Owusu-Dommey, Akosua[6]; Binder, Morgan[6]; Freitas, Nancy L.[6]; Crill, Patrick M.[5]; Saleska, Scott R.[6]; Tyson, Gene W.[3]; Rich, Virginia I.[1]

[1]Ohio State University. [2]University of New Hampshire. [3]University of Queensland. [4]University of Washington. [5]Stockholm University. [6]University of Arizona.

Modeling the fate of carbon and understanding carbon (C) cycling dynamics in thawing permafrost peatlands has been a challenge due to the lack of empirical data regarding interactions of microbial communities and their impacts on metabolism of C. C released from these peatlands can flow via streams to lakes and be deposited to the sediments. Microbial communities play a large role in the ultimate fate of C but have yet to be analyzed in the post-glacial lake sediments associated with these landscapes. Sediment cores were collected from three post-glacial lakes in the Stordalen mire peatland complex located in the sporadic permafrost zone in northernmost Sweden. Sediment core samples were set within a multi-year ebullitive (bubbling) and diffusive CH_4 flux dataset and linked to aquatic vegetation mapping and sedimentary biogeochemical analyses. Microbial community composition was assessed using 16S rRNA gene amplicon sequencing, and microbial population genomes were reconstructed via metagenomics. Comparisons of microbial community composition, predicted metabolisms from the population genomes, and carbon biogeochemistry revealed key microbial lineages predicted to mediate carbon transformations and losses. Some population genomes revealed carbon-transforming metabolisms that deviate from previous predictions for their lineages, including methylotrophy in Candidate Phylum *Aminicenantes* and fermentation without the capacity for methanogenesis in the Methanomassiliicoccaceae. Findings have provided groundwork for ongoing sequence-based interrogation of the lake sediments in the Stordalen Mire ecosystem over a multi-year timescale.

Genomics of Individual Cells and Viruses Revamped: Enhanced Gdna Amplification and Direct, High Throughput Matching of Genomic and Physical Properties

Fergusson, Elizabeth A.*[1]; Stepanauskas, Ramunas X. (rstepanauskas@bigelow.org)[1]; Brown, Joseph[1]; Poulton, Nicole J.[1]; Tupper, Ben[1]; Povilaitis, Tadas[2]; Labonté, Jessica M.[1]; Becraft, Eric D.[1]; Brown, Julia M.[1]; Pachiadaki, Maria G.[1]; Jeremian, Richie[3]; Thompson, Brian P.[1]; Mascena, Corianna J.[1]; Bellows, Wendy K.[1]; Petronis, Arturas[3]; Lubys, Arvydas[2];

[1]Bigelow Laboratory for Ocean Sciences, East Boothbay, Maine, USA. [2]Thermo Fisher Scientific Baltics, Vilnius, Lithuania. [3]Centre for Addiction and Mental Health, Toronto, Canada.

Single cell genomics (SCG) is a transformative tool in studies of genomic differentiation in microbiomes and multicellular organisms. We present three major improvements to the SCG workflow:

1. Improved genome recovery using a thermostable mutant of the phi29 polymerase. Called WGA-X, this method reduces genome amplification biases against high G+C templates, as compared to the traditional multiple displacement amplification (MDA).
2. Linkage of genomic and physical properties for hundreds of individual cells and extracellular genetic elements in a single experiment. Index-sorting records physical characteristics such as size and fluorescence of each particle that is separated and subsequently subjected to genomic analysis.
3. Bias-free alternative to PCR for high- throughput single cell identification and initial characterization. Low coverage genomic sequencing (LoCoS) reveals many genomic properties in addition to cell identity.
4. We demonstrate how these new tools enhance genome recovery and the information gained from soil and marine prokaryotes, microalgae and human cells using SCG. Using these improved methods, we also present the first genomic analyses of individual marine extracellular particles (including viruses).

Under Attack: How a Marine Cyanobacterium Responds to Virus Infection and Grazer Predation

Floge, Sheri A.* (floge.1@osu.edu)[1]; Roux, Simon[2]; Bowen, Benjamin P.[3]; Northen, Trent R.[3]; Woyke, Tanja[2]; Sullivan, Matthew B.[1]

[1]The Ohio State University, Columbus, OH, USA. [2]DOE Joint Genome Institute, Walnut Creek, CA. [3]Lawrence Berkeley National Laboratory, Berkeley, CA.

Drifting photosynthetic microbes occupying the sunlit waters of the global ocean support complex food webs that ultimately provide essential nutrition to half the world's population. Marine microbes exist in a tightly interwoven web connected by a wide range of individual biotic and abiotic interactions that together, control global scale biogeochemical cycling. Of these microbes, the cyanobacterium *Synechococcus* is one of the most abundant and has been identified, along with it's viral and protistan predators, as a key player in oceanic carbon export. Using a combination of transcriptomics and metabolomics we examined the response of *Synechococcus* strain WH8102 to infection by the myovirus S-SSM5, predation by the single-celled dinoflagellate grazer *Oxyrrhis*

marina (CCMP3375) and to combined viral and grazer predation. Phage infection resulted in increased expression of genes regulating nucleotide biosynthesis as well as elevated intracellular guanine and increased release of adenine, guanine and cytosine into surrounding media. While the presence of the grazer alone had no measurable impact upon gene expression and intracellular metabolite composition, combined viral and grazer predation induced physiological changes in excess of those observed with viral infection alone. In particular, genes involved in carbon and nitrogen metabolism were upregulated under combined predation pressure. Our data suggest that the combined effects of viral and grazer predation can be greater than the sum of individual interactions and illustrate the potential importance of understanding the net impact of multiple microbial interactions when linking microscale cellular physiology to global scale biogeochemical processes.

Spatiotemporal Characterization of Microbial Communities Controlling Estuarine Nitrogen Cycling in the San Francisco Bay-Delta

Francis, Christopher A.* (caf@stanford.edu)[1]; Rasmussen, Anna[1]; Damashek, Julian[1, 2]; Lee, Jessica[1, 3]; Baker, Brett[4]

[1]Department of Earth System Science, Stanford University, Stanford, California, USA. [2]Department of Marine Sciences, University of Georgia, Athens, Georgia, USA. [3]Department of Biological Sciences, University of Idaho, Moscow, Idaho, USA. [4]Department of Marine Science, University of Texas Austin, Marine Science Institute, Port Aransas, Texas, USA.

Nitrogen (N) availability is an important factor controlling productivity in estuaries. Yet, we know relatively little about how the abundance, diversity, and distribution of estuarine N-cycling microbes are affected by environmental change. In this CSP project, we seek to investigate the structure and function of microbial communities, particularly those related to N and carbon cycling, in the San Francisco Bay-Delta—the largest estuary on the west coast of North America. Sequencing will span the diverse estuarine gradient, including both water and sediment from high-nutrient riverine regions, brackish transition zones, and marine regions collected as a 'time series' during USGS monthly monitoring cruises, which provide a vast array of environmental data to leverage against sequencing data. This project will generate catalogs of metagenomic, metatranscriptomic, and 16S rRNA sequences to compare with extensive physicochemical and biogeochemical meta-data at an unprecedented spatial and temporal scale. Additionally, prior work in our laboratory has produced extensive characterizations of N-cycling functional genes (e.g., *amoA, nirK, nirS*), 16S rRNA amplicons, and biogeochemical rates from many of these samples, facilitating direct comparison between metagenomic and gene-specific methods targeting the same organisms as well as their ecological impact in this nutrient-rich urban estuary.

Evolution of Mutualistic Cross-Feeding in a Synthetic Bacterial Coculture

Fritts, Ryan K.* (rkfritts@indiana.edu); Gliessman, Jennifer R.; Behringer, Megan G.; Lynch, Michael; McKinlay, James B.

Department of Biology, Indiana University, Bloomington, IN 47405, USA

Microbes typically reside in interactive multispecies communities. The metabolic interactions that occur between different microbes can profoundly affect community structure and biogeochemical cycling. Cross-feeding interactions in which one microbe releases nutrients into the environment that are consumed by other species are abundant in nature. If such cross-feeding interactions are bidirectional/reciprocal, mutualistic partnerships can form. To investigate the evolution of a nascent mutualism based on cross-feeding of essential nutrients, we established a genetically tractable, synthetic coculture comprised of the N2-fixing bacterium, *Rhodopseudomonas palustris*, and the fermentative bacterium, *Escherichia coli.* In this bacterial coculture, *R. palustris* fixes N2 gas and supplies essential NH_4^+ to *E. coli.* In return, *E. coli* ferments glucose and excretes organic acids as waste byproducts, which serve as the sole carbon source for *R. palustris.* Previous work from our lab has shown that engineered *R. palustris* strains with mutations in nifA, the master transcriptional regulator of nitrogenase, and amtB, the NH_4^+ uptake transporter, can support coculture growth by increasing the level of NH_4^+ excretion by *R. palustris*. In contrast, wild-type *R. palustris,* which lacks the aforementioned mutations, does not supply enough NH_4^+ to *E. coli* to support coculture growth. To determine if WT *R. palustris* can evolve to provide sufficient levels of NH_4^+ to *E. coli* to support mutualistic growth, we used an experimental evolution approach. We started six parallel coculture lineages of WT *R. palustris* and *E.coli* under both well-mixed and static conditions. Following inoculation, there was a two week period of little to no observable coculture growth, but eventually all coculture lineages grew and were serially propagated for one year. To determine if the genotype(s) underlying the spontaneously evolved NH_4^+ cross-feeding phenotype were similar to what we engineered or novel, we performed whole genome sequencing of coculture populations in collaboration with the Joint Genome Institute. We did not identify any mutations in either *nifA* or *amtB* in the evolved *R. palustris* population. However, we identified multiple parallel mutations in *R. palustris* and in *E. coli.* One of the genes in which we saw parallel mutations in *R. palustris* was *ntrY,* the sensor kinase of the poorly understood NtrYX two component system, putatively involved in regulating nitrogen metabolism. We hypothesize that these mutations increase NH_4^+ excretion, either by upregulating nitrogenase activity or by downregulating NH_4^+ recapture. We also identified mutations in *E. coli* in glnK, which functions as an inhibitor of the AmtB NH_4^+ transporter. We hypothesize that these mutations may increase NH_4^+ uptake by *E. coli.* Future work will focus on defining the role of these and other mutations in the evolved NH_4^+ cross-feeding phenotype using a combination of genetic, biochemical, mathematical modeling approaches.

ExaBiome: Exascale Solutions to Microbiome Analysis

Georganas, E.*[1]; Egan, Rob S. (rsegan@lbl.gov)[1, 3]; Buluc, A.[1]; Hofmeyr, S.[1]; Oliker, L.[1]; Rokhsar, Daniel[1, 3]; Yelick, Kathy[1]; Chain, P.[2]; Senin, P.[2]; Ivanova, N.[1, 3]; Kyrpides, Nikos[1, 3]; Neilson, T.[1, 3]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA. [2]Los Alamos National Laboratory, Los Alamos, NM. [3]DOE Joint Genome Institute, Walnut Creek, CA.

"Metagenomics" — the application of high-throughput genome sequencing technologies to DNA extracted from microbiomes — is a powerful method for studying microbiomes. But the first assembly step has high computational complexity, akin to putting together thousands of puzzles from a jumple of their pieces. Following assembly, additional data analysis is needed to find families of genes that work together and to compare across metagenomes. The ExaBiome team from Berkeley Lab, Los Alamos, and the Joint Genome Institute is developing exascale algorithms and software to address these challenges and will work with the vendor community to co-design systems that have the necessary network and memory features to address these and other large scale analytics problems. https://exascaleproject.org/2016/11/05/solutions-for-microbiome-analysis/

Quantifying Genomic Exchange in Chickpea-Nodulating *Mesorhizobium*
(Short Talk)

Greenspan, Alex* (greenspan@ucdavis.edu)[1];Alford, Betsy[1]; Chang, Peter[1]; Von Wettberg, Eric[2]; Kahraman, Abdullah[3]; Bükün, Bekir[4]; Mohammed, Zehara[5]; Fikre, Asnake[6]; Sarma, Birinchi[7]; Cook, Douglas R.[1]

[1]University of California, Davis, California. [2]Florida International University, Miami, Florida. [3]Harran University, Şanlıurfa, Turkey. [4]Dicle University, Diyarbakır, Turkey. [5]Addis Ababa University, Addis Ababa, Ethiopia. [6]Ethiopian Institute of Agricultural Research, Addis Ababa, Ethiopia. [7]Banaras Hindu University, Varanasi, India.

Legume crops are significant agriculturally and environmentally for their ability to form a symbiosis with specific soil-bacteria capable of nitrogen fixation. However, nitrogen fixation for a given legume in a given soil is limited by the availability of the plant's bacterial partners, and by variation in the effectiveness of those symbionts. In industrial agricultural systems where legume crops are provided specific inoculants, inoculation can fail if the added strains are unable to compete with more aggressive but less efficient soil-endemic strains. Biogeography—the study of what factors pattern the distribution of organisms in time and space—has historically ignored bacteria due to the limited genetic resolution of microbial taxonomy. However, biogeographic insight is vital to understand what factors affect nitrogen fixation in legume crops, and how nitrogen fixation can be enhanced through inoculation and other techniques. Similarly, understanding the relationship between a legume crop's symbionts in a geographic context can elucidate basic principles of microbial biogeography that have hitherto been unexplored.

We used a global-level hierarchical sampling scheme to comprehensively characterize the evolutionary relationships and distributional limitations of the nitrogen-fixing bacterial symbionts of the crop chickpea. This has been accomplished using culture-dependent and independent approaches to generate over 1,200 draft whole-genome assemblies at the level of bacterial populations, as well as 17 finished genomes for a subset of strains representing the full, observed geographic and phylogenetic diversity of chickpea's symbionts. These strategies have revealed a surprising diversity in chickpea's symbionts around the globe. The nature of sampling allows us to infer factors that pattern the distribution of chickpea's diverse symbionts. Experiments are currently underway in India and Ethiopia to test whether predictions of symbiotic effectiveness of variable strains based on whole-genome sequences hold in farmers' fields.

The biogeographic history of chickpea is well understood and we have sampled across the crop's full global distribution (including it's wild ancestors in their native range, as well as ancient and modern introductions of the domesticate); we can use this knowledge to infer patterns in the distribution and genetic relationships of its bacterial symbionts, additionally answering basic questions of bacterial biogeography. Comparative phylogenomic analysis reveals that despite chickpea's symbionts within and across regions coming from different taxa, all share almost identical genes for symbiosis. Full-chromosome genome assemblies reveal that this is due to the horizontal transfer of a 500 kbp chromosomal cluster known as the symbiosis island, between unrelated strains of the genus *Mesorhizobium.* Analyzing the symbiosis island at the population level reveals that the symbiosis island spreads repeatedly once introduced to a region, suggesting that strains well-adapted to a particular soil climate continue to dominate once the new host (chickpea) has been introduced, through repeated acquisition of the symbiosis island.

Unlocking the Photosynthetic and Genetic Diversity of Cryptophyte Algae Through Whole-Genome Sequencing of a Diverse Assemblage of Species

Greenwold, Matthew J.* (greenwold@biol.sc.edu); Cunningham, Brady R.; Richardson, Tammi L.; Dudycha, Jeffry L.

Department of Biological Sciences, University of South Carolina, Columbia, South Carolina.

Cryptophytes are eukaryotic microalgae that are important primary producers in aquatic environments, including ponds, lakes, estuaries, and oceans. They have unique photosynthetic pigments, the cryptophyte phycobilins, which are broadly classified into phycoerythrins (PE) and phycocyanins (PC). Here, we propose sequencing high quality *de novo* nuclear, nucleomorph, plastid, and mitochondrial genomes of 12 phenotypically and phylogenetically diverse cryptophytes. Our goal is to understand the genetic diversity of cryptophyte photosynthesis in the context of their ability to capture a broad spectrum of available light. These genomes will provide a valuable resource to researchers studying cryptophytes, microalgal photosynthesis and systematics, and marine metagenomics by providing a better catalog of the diversity within this clade of algae. The proposed species span all major clades within the phylum. Furthermore, they include freshwater and marine species, as well as PE- and PC-containing species spanning the diversity of cryptophyte coloration (Hoef-Emden 2008). Two species adapted to extreme environments (the Arctic and Southern Oceans) are also included.

Combining Multiple Functional Annotation Tools Increases Completeness of Metabolic Annotation

Griessemer, Marc*; D'haeseleer, Patrik M. (dhaeseleer2@llnl.gov); Kimbrel, Jeffrey; Navid, Ali;

Lawrence Livermore National Laboratory, Livermore, CA.

The dirty little secret behind "genome-wide" systems biology modeling efforts is that these models are based on very incomplete functional annotations. Annotated genomes typically contain 30-50% of genes with little or no functional annotation, severely limiting our knowledge of the "parts lists" that the organisms have at their disposal. In metabolic modeling, these incomplete annotations are often sufficient to derive a reasonably complete model of the core metabolism at least, typically consisting of well-studied (and thus well-annotated) metabolic pathways that are sufficient for growth in pure culture. However secondary metabolic pathways or pathways that are important for growth on unusual metabolites exchanged in complex microbial communities are often much less well understood, resulting in missing or lower confidence functional annotations in newly sequenced genomes. For example, one third of the EC database consists of "orphan enzymes" that have been described in the literature but for which no sequence data is available.

Individual metabolic annotation tools often return annotations for different subsets of genes, offering the potential to greatly increase the completeness of metabolic annotations by combining the outputs of multiple tools. Indeed, recent genome-scale modeling of *Clostridium beijerinckii* NCIMB 8052 demonstrated that the total number of genes and reactions included in the final curated model could be almost doubled by incorporating multiple annotation tools.

Here, we present preliminary results on a comprehensive reannotation of 27 bacterial reference genomes, focusing on enzymes with EC numbers annotated by KEGG, RAST, EFICAz, and the Brenda enzyme database.

A Genomic Perspective on Stoichiometric Regulation of Soil Carbon Cycling

Hartman,Wyatt H.* (whhartman@lbl.gov)[1]; Ye, Rongzhong[2]; Horwath, William R.[2]; Tringe, Susannah G.[1]

[1]DOE Joint Genome Institute, Walnut Creek, CA. [2]Department of Land, Air and Water Resources, University of California, Davis, CA.

Like plant growth, soil carbon (C) cycling is constrained by the availability of nitrogen (N) and phosphorus (P). We hypothesized that stoichiometric control over soil microbial C cycling may be shaped by functional guilds with distinct nutrient substrate preferences. Across a series of rice fields spanning 5-25% soil C (N:P from 1:12 to 1:70), carbon turnover was best correlated with P availability, and increased with experimental N addition only in lower carbon (mineral) soils with N:P < 16. Microbial community membership also varied with soil stoichiometry, but not with N addition. Shotgun metagenome data revealed changes in community functions with increasing C turnover, including a shift from aromatic C to carbohydrate utilization accompanied by lower N uptake and P scavenging. Similar patterns of C, N and P acquisition, along with higher ribosomal RNA operon copy numbers, distinguished microbial taxa positively correlated with C turnover. Considering such tradeoffs in genomic resource allocation patterns among taxa strengthened correlations between microbial community composition and C cycling, suggesting simplified guilds amenable to ecosystem modeling. Overall, our results suggest patterns of soil C turnover reflect community dependent metabolic shifts driven by resource allocation strategies, analogous to growth rate – stoichiometry coupling in animal and plant communities.

Efficient Integration of Phytozome Population Diversity and Expression Data with JBrowse

Hayes, Richard D.* (rdhayes@lbl.gov); Carlson, Joseph W.; Goodstein, David M.; Rokhsar, Daniel S.

DOE Joint Genome Institute, Walnut Creek, CA.

Phytozome (http://phytozome.jgi.doe.gov) is a plant comparative genomics web portal offering access to genome annotations and gene families for over 70 species from a diverse set of land plant and Chlorophyte/algal taxa. Our annotations are increasingly augmented by studies of gene expression across tissue types, developmental stages, and environmental conditions, and exploration of genetic diversity across natural and designed populations. Sample sizes continue to grow rapidly, posing unique challenges for visualization systems like JBrowse that perform graphical rendering within end-user web browser client software. Recent solutions to this problem are discussed, building on existing work to offer Phytozome users efficient display of this data in JBrowse while minimizing client-side data access and rendering resource consumption.

Identification and Characterization of the Core and Disposable Genome of *Setaria viridis* through Gene Presence/Absence Variation

Healey, Adam*[1]; Schmutz, Jeremy (jschmutz@hudsonalpha.com)[1, 3]; Jenkins, Jerry[1]; Mamidi, Sujan[1]; Sreedasyam, Avinash[1]; Barry, Kerrie[3]; Mockler, Todd[2]; Huang, Pu[2]; Brutnell, Tom[2]

[1]HudsonAlpha Institute for Biotechnology; Huntsville, Alabama, USA. [2]Donald Danforth Plant Science Center, St Louis, MO, USA. [3]DOE Joint Genome Institute, Walnut Creek, CA, USA.

Panicoid grasses (family Poaceae, subfamily Panicoideae), which includes Sorghum, Maize, Switchgrass, and *Miscanthus,* are a valuable agricultural resource, both as food crops and renewable sources for bioenergy production. However, given the size and complexity of their genomes, as well as long generation times, progress to understand the genetic basis of important agronomic traits within this group remains difficult. *Setaria viridis* (green millet), the undomesticated progenitor of *Setaria italica* (foxtail millet), is an ideal candidate model organism for panicoid grasses, based on its relatively small (550 Mb) diploid genome, short lifecycle and stature, prolific seed production and efficient transformation protocol.

Additionally, *S. viridis* displays wide phenotypic variation with respect to growth, flowering time, and inflorescence structure and is highly adaptable to environmental and climatic changes as well as utilizing efficient C4 photosynthesis. However, despite ample phenotypic diversity, the resources to dissect the underlying genetic mechanisms behind these desirable traits in *S. viridis* have not yet been developed. Traditional SNP based assays, which are reliant on mapping reads to a high-quality reference genome, are unable to capture structural variations within the genome, which can be responsible for large phenotypic differences among genotypes. Using *de novo* assembled genomes of 419 *S. viridis* accessions, we characterize the *Setaria* pan-genome through the identification of core and disposable gene sets through presence/absence variation analyses, as well as identify gene families significantly over/under- represented among various subpopulations. This information can be paired with SNP-based analyses to discover the genes and variation responsible for driving and maintaining phenotypic differences among wild accessions of *S. viridis,* as well as provide insight into the concept of the plant pan-genome, and evolution and domestication of panicoid grasses.

Profiling Complex Population Genomes with Highly Accurate Single Molecule Reads

Heiner, Cheryl R.* (cheiner@pacificbiosciences.com)[1]; Sharon, Itai[2]; Oh, Steve[1]; Hernandez, Alvaro G.[3]; Mizrahi, Itzhak[4]; Hall, Richard[1]

[1]Pacific Biosciences, Menlo Park, CA. [2]Tei-Hai College, Upper Galilee, and MIGAL Galilee Research Institute, Israel. [3]University of Illinois at Urbana-Champaign, IL. [4]Ben-Gurion University of the Negev, Israel.

Determining compositions and functional capabilities of complex populations is often challenging, especially for sequencing technologies with short reads that do not uniquely identify organisms or genes. Long-read sequencing improves the resolution of these mixed communities, but adoption for this application has been limited due to concerns about throughput, cost and accuracy.

The recently introduced PacBio Sequel System generates hundreds of thousands of long and highly accurate single-molecule reads per SMRT Cell.

We investigated how the Sequel System might increase understanding of metagenomic communities. In the past, focus was largely on taxonomic classification with 16S rRNA sequencing. Recent expansion to WGS enables functional profiling as well, with the ultimate goal of complete genome assemblies.

Here we compare the complex microbiomes in 5 cow rumen samples, for which Illumina WGS sequence data was also available. To maximize the PacBio single-molecule sequence accuracy, libraries of 2 to 3 kb were generated, allowing many polymerase passes per molecule. The resulting reads were filtered at predicted single-molecule accuracy levels up to 99.9%.

Community compositions of the 5 samples were compared with Illumina WGS assemblies from the same set of samples, indicating rare organisms were often missed with Illumina. Assembly from PacBio CCS reads yielded a contig >100 kb in length with 6-fold coverage. Mapping of Illumina reads to the 101 kb contig verified the PacBio assembly and contig sequence. Scaffolding with reads from a PacBio unsheared library produced a complete genome of 2.4 Mb.

These results illustrate ways in which long accurate reads benefit analysis of complex communities.

Novel Microbial Lineages Endemic to Geothermal Iron-Oxide Mats Fill Important Gaps in the Evolutionary History of Archaea

Inskeep, William P.*[1]; Inskeep, Dylan J. (dilly@live.com)[1]; Jay, Zackary J.[1]; Beam, Jacob P.[1]; Dlakic, Mensur[2]; Rusch, Douglas B.[3]; Kozubal, Mark A.[1]

[1]Thermal Biology Institute and Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT. [2]Department of Microbiology and Immunology, Montana State University, Bozeman, MT. [3]Center for Bioinformatics, Indiana University, Bloomington, IN.

Discovery of new archaeal lineages is critical to our understanding of the evolutionary history of Earth, and the important role that these organisms played in the evolution of Eukarya. There is strong support for the argument that life on Earth originated in thermal environments, and numerous thermophilic archaeal lineages occupy early branches on the Tree of Life. Extensive and geochemically diverse thermal environments in Yellowstone National Park (YNP) provide unprecedented opportunities for studying different groups of archaea in habitats that may represent analogs of those important on early Earth, and for discovering new members of this domain. Here we report the important discovery and thorough characterization of a deeply-rooted, phylum-level archaeal lineage (proposed and herein referred to as the 'Marsarchaeota', after the red planet), which contains at least two major subgroups that are endemic to moderately-acidic geothermal iron-oxide and jarositic microbial mats. Detailed metagenomics, single-cell sequencing and isolated consortia, as well as *in situ* transcriptional analyses and microscopy reveal their likely activity as facultative aerobic chemoorganotrophs and capabilities for dissimilatory Fe(III) reduction across a broad temperature range (circa 50-80°C) of geothermal Fe(III)-oxide deposits. These high-temperature, Fe-rich ecosystems support the growth and numerical abundance of the newly described Marsarchaeota as well as several other deeply-rooted archaeal lineages, and provide evidence for the importance of these habitat-types in the early evolution of Archaea.

Towards High-Quality *de novo* Plant Genome Assemblies with PACBIO

Jenkins, Jerry*[1]; Schmutz, Jeremy (jschmutz@hudsonalpha.org)[1, 2]; Grimwood, Jane[1, 2]; Barry, Kerrie[2]; Goodstein, David[2]; Shu, Shengqiang[2]; Daum, Chris[2]

[1]HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806. [2]DOE Joint Genome Institute, Walnut Creek, CA 94598.

Reference assemblies that are accurate, complete, and contiguous are essential for identifying important structural/functional elements and for analyzing genetic variation. Many of the plant reference genomes that are currently being sequenced pose unique challenges due to their large size (>10Gb), ploidy, prevalence of repetitive content, and high heterozygosity. Next generation sequencing technologies, such as Illumina, produced a large volume of accurate reads, but cannot adequately address the problems of ploidy, repeat content, and heterozygosity due to their short read length. However, our recent work in generating 8 reference assemblies using PACBIO reads has demonstrated the ability of long read technologies to produce genomes that are complete, contiguous, and accurate. Two specific examples of genomes we have produced will be presented to highlight some of the challenges that still exist with PACBIO assemblies, and how we are overcoming them. First, the problem of misjoins in FALCON assemblies and how to specifically pinpoint the misjoin will be addressed using the examples of *Phaseolus vulgaris* and *Salix purpurea* (var. Fishcreek). Second, the impact of high heterozygosity on consensus calling in FALCON, and the need to post-correct the consensus using Illumina will be highlighted in our work pertaining to *Populus deltoids* (var. WV94). Finally, we will compare the latest releases of *Oryza sativa* (var. Kitaake), *Sorghum rio, Brachypodium distachyon* (var. Bd21-3), and an algal genome *Botryococcus braunii* to our previous releases to demonstrate the significant improvement that long read technologies can bring to bear when faced with a challenging genome project.

Using Whole-Genome Sequencing to Identify the Components in GIPC-Related Plant Immunity Pathway

Jing, Beibei*[1]; Mortimer, Jenny C. (jcmortimer@lbl.gov)[1]; Ishikawa, Toshiki[2]; Lao, Jeemeng[1]; Pidatala, Ramana[1]; Anderbehan, Fekadu[1]; Yu, Xiaolan[3]; Baidoo, Edward[1]; Kawai-Yamada, Maki[2]; Li, Guotian[4]; Ronald, Pamela[4]; Dupree, Paul[3]

[1]Joint BioEnergy Institute, BioSciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [2]Graduate School of Science and Engineering, Saitama University, Japan. [3]Department of Biochemistry, University of Cambridge, Cambridge, UK. [4]Department of Plant Pathology, University of California, Davis, CA, USA.

The Golgi lumen is the site of many different glycosylation events, including cell wall polysaccharide biosynthesis and lipid glycosylation. Nucleotide sugar transporters (NSTs) are necessary for the import of the substrates (nucleotide sugars) required for glycosylation into the Golgi from the cytosol. Plants use four GDP-linked sugars to glycosylate macromolecules: GDP-L-Fucose, GDP-D-Mannose, GDP-L-Galactose and GDP-D-Glucose. Of all the predicted nucleotide sugar transporters in *Arabidopsis*, only four appear to contain the conserved motif needed for the transport of GDP-linked sugars, GOLGI LOCALIZED NUCLEOTIDE SUGAR TRANSPORTER (GONST) 1-4. GONST1 has been previously identified as a GDP-Man transporter1.

Sugars decorate glycosphingolipids. Glycosylinositolphosphoceramides (GIPCs) are the most abundant sphingolipid in the plant plasma membrane. gonst1 shows a dramatic reduction in GIPC mannosylation, which indicated GONST1 provides GDP-Mannose for glycosylation of GIPCs. gonst1 has a constitutive hypersensitive response phenotype including severely stunted stature, an early senescence phenotype, leaf lesions and dramatically increased Salicylic Acid (SA) level. NahG which encodes salicylate hydroxylase and converts SA to catechol partially suppresses the gonst1 phenotype. The characterization of gonst1 demonstrates that the loss of Man from GIPCs has a severe effect on plant development and immunity. To characterize the components in the GIPC but not SA related plant immunity pathway, we set up a forward genetic screen using the EMS mutagenized NahGgonst1-2 population. The introduction of NahG also made the plant culture much easier. So far, we have found one suppressor E1-5. Currently we are using next generation sequencing to identify the mutation which is responsible for the rescued phenotype.

[1]Mortimer et al. 2013, Plant Cell

Elucidation of Genes Required for Endophytic N-Fixation and Their Role in Plant Growth Promotion

Joubert, Pierre M.* (pierrj@uw.edu)[1]; Deutschbauer, Adam[3]; Doty, Sharon L.[2]

[1]Biology Department, University of Washington, Seattle, WA, USA. [2]School of Environmental and Forest Sciences, University of Washington, Seattle, WA, USA. [3]Lawrence Berkeley National Laboratory, Berkeley, CA, USA.

Nitrogen is one of the most important nutrients for plant growth, and fertilizers are heavily used worldwide to provide this nutrient. However, fertilizers are costly to produce and have significant negative impacts on the environment. Microorganisms within plants, termed endophytes, were isolated from wild poplar and willow plants and shown to promote growth in several other plant species as broadly diverse as rice, tomato, and Douglas fir. N-fixation was demonstrated in wild poplar and in hybrid poplar inoculated with a consortium of the endophytes from wild poplar. An N-fixation mutant endophyte was constructed and tested on crop plants. Additionally, in order to better understand the genes involved in N-fixation by these endophytes, mutant libraries will be created by random transposon mutagenesis in collaboration with the JGI and Dr. Deutschbauer of the Lawrence Berkeley National Lab. The libraries will be screened for inability to grow on N-free medium. The Tn insertion sites will be identified by sequencing. These results as well as a comparative genomics analysis with other diazotrophic bacteria may provide insight into how endophytes protect nitrogenase from inhibitory oxygen. The project will also provide additional mutants to be tested on plants for determining the relative importance of N-fixation in the observed growth promoting effects of endophytes. A better understanding of the processes involved in N-fixation could lead to the creation of an optimized consortia of naturally occurring, N-fixing endophytes that could reduce the need for fertilizers in the future.

Community Metagenomics of Bacterial Community in Fungus Gardens of Atta Leaf-Cutter Ants

Khadempour, Lily* (khadempour@wisc.edu); Currie, Cameron R.

Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, USA.

Herbivores, the most abundant animals on Earth, use symbiotic microbes to help derive energy and nutrients from plant material. Leaf-cutter ants are a paradigmatic example of this, as they use cut leaf material to cultivate a mutualistic fungus that they use as their primary food source in subterranean fungus gardens. Some leaf-cutter ants have transitioned to cutting grasses instead of the usual dicots. These ants have adapted both morphologically and behaviorally to grass cutting, and they form a monophyletic clade, which demonstrates an evolutionary transition to a new substrate. Their fungus gardens also contain a community of bacteria, the role of which is less understood. In order to address gaps in knowledge about the roles of the bacteria, we examined the bacterial community in the fungus gardens of four species of leaf-cutter ants: *Atta bisphaerica* and *Atta capiguara,* strict grass-cutters, *Atta sexdens,* a strict dicot-cutter, and *Atta laevigata,* which is a grass and dicot-cutter. Three bacterial community metagenomes were sequenced at the Joint Genome Institute with an Illumina HiSeq 2500 sequencer. In total, we sequenced 12 metagenomes, three from each ant species. The assembled metagenomes contained a total of 5.28 Gbp of sequence, and a total of 8.19 million genes. The metagenomes were dominated by the phylum Enterobacteriaceae. The bacterial communities in the gardens of grass-cutter ants, *A. bisphaerica* and *A. capiguara* were similar and had relatively low diversity, and were dominated by the genus *Pantoea*. The bacterial communities of *A. sexdens,* a dicot-cutter ant were more diverse, presumably because the plant substrates, and the respective plant defense compounds that the ants consume are more diverse. Interestingly, bacterial communities from *A. laevigata,* the grass and dicot-cutter, were the most variable, in one case being more similar to grass-cutter ant bacterial communities and in another, being more similar to dicot-cutter communities. While still in the preliminary stages of analysis, the work presented here demonstrates the connection between the functional and compositional diversity of a microbial community and how they relate to the substrate utilization of their host and can have implications for microbiome research and for functional community ecology.

Automated DNA Synthesis in a Microfluidic Chip

Kim, Peter W.* (pwkim@sandia.gov)[1, 2]; Arlow, Dan[1, 3, 4]; Palluk, Sebastian[1, 3, 4]; Bector, Rathin[1, 4]; Kang, Justine[1, 4]; Hillson, Nathan[1, 3]; Singh, Anup[1, 2]; Keasling, Jay D.[1, 3, 4]

[1]Joint BioEnergy Institute, Emeryville, CA. [2]Sandia National Laboratory, Livermore, CA. [3]Lawrence Berkeley National Laboratory, Berkeley, CA. [4]University of California, Berkeley, CA.

The availability of fast, reliable, and affordable synthesis of multikilobase DNA molecules with arbitrary sequences would dramatically accelerate workflows in life sciences research and bioengineering [1].Though the price of synthetic DNA continues to fall, incremental improvements on existing technologies are unlikely to yield a technique that will enable the reliable synthesis of arbitrary multi-kilobase pair (kbp) DNA molecules in under a week. Furthermore, existing methods struggle to synthesize "challenging" sequences containing repeats, structure-forming regions, or larger variation in GC content.

We are developing a microfluidic-based prototype of a new type of DNA synthesizer that, in principle, could synthesize 100% accurate multi-kilobase DNA with any specified sequence in a few days of continuous operation. Our approach achieves fast, reliable, and affordable multi- kbp DNA synthesis with no sequence restrictions by departing from the paradigm of bulk-scale oligonucleotide synthesis and assembly, and instead building literally one molecule of double-stranded DNA by stepwise enzymatic concatenation of ultra- pure fluorescent double-stranded DNA "Extension Units" (ExUs) that extend the growing DNA molecule 1 bp per cycle.

To meet this goal, we are developing an integrated microfluidic system with single- molecule fluorescence detection capability. This prototype chip has capability to dispense various reagents in nanoliter scale using computer-controlled pneumatic micro valve system. Automation of pneumatic valves and fluorescent imaging/analysis can achieve fully automated DNA synthesis. Also the substrate functionalization via "click" reaction showed very robust and selective labeling of DNA starting materials.

Isotope Tracing and Phylogenetic Composition of Simplified Bacterial Communities Conferring Growth and Biomass Enhancements to Biofuel-Producing Microalgae

Kimbrel, Jeffrey* (kimbrel1@llnl.gov)[1]; Samo, Ty[1]; Ligeti, Jorge[1]; Thelen, Michael[1]; Weber, Peter K.[1]; Siccardi, Anthony[2]; Lane, Todd[3]; Mayali, Xavier[1]

[1]Lawrence Livermore National Laboratory, Livermore, CA. [2]Texas A&M Agrilife, Corpus Christi, TX, [3]Sandia National Laboratories, Livermore, CA.

Mutualistic algal-bacterial interactions can arise when bacteria provide metabolically beneficial substances in exchange for energy-dense algal compounds. We have studied these positive exchanges in closely interacting algal-bacterial cultures, and simplified algal-attached bacterial communities developed via culture enrichments. We monitored the algae for elevated growth and biomass characteristics and quantified C and N exchanges with both bulk and single-cell approaches (using nanoscale secondary ion mass spectrometry (NanoSIMS)). Through successive rounds of culturing, we have established over 100 stable algal-bacterial co-cultures that exhibit increased algal productivity (as determined by chlorophyll fluorescence), and we have isolated several dozen phycosphere-attached bacteria.

We used two model biofuel-producing microalgal strains, *Phaeodactylum tricornutum* (Pt) and *Nannochloropsis salina* (Ns), to enrich for growth-promoting bacteria acquired either from the coastal Pacific Ocean or established algal raceway ponds in Texas. After enriching for phycosphere-attached bacteria, microbial communities (characterized by 16S rDNA gene sequencing) exhibited a higher abundance of some bacteria that were rare in the source communities, and lower abundance of bacteria with an exclusively "free-living" lifestyle (those abundant in the culture supernatant but below the limit of detection in the washed algal filtrate). The mechanisms leading to these community composition changes vary, and appear to depend on the source inoculum. Replicates from some sources had nearly identical emergent communities due to a consistent increase in bacteria that were rare in the source communities. However, in emergent communities from other sources, stochastic losses result in increased heterogeneity of community composition among the replicates. Inoculating a source community with either Pt or Ns had similar results, in both cases the host exerted strong selection to shape the microbial community. Compared to the original source communities,

enrichments led to significant increases in the Rhodobacterales, most notably *Loktanella, Ruegeria* and *Labrenzia* genera. Additionally, *Rhodobacterales, Sphingobacterales* and *Alteromonadales* genera were found to be significantly enriched in the phycosphere-attached over the free-living fractions.

We have established co-cultures of Pt and single bacterial species isolated from the community enrichments. Each isolated bacterium is highly abundant in the phycosphere-attached fraction of the enriched community, suggesting a role in algal population dynamics. To test the hypothesis that these isolates may shape these algal-bacterial symbioses, we quantified the exchanges of C and N between the bacteria and the algal host using NanoSIMS. Indeed, phycosphere attachment of bacteria lead to a greater incorporation of fixed algal $13CO_2$ products over unattached bacteria. While carbon transfer from algae to bacteria was confirmed, algal utilization of microbial-derived nitrogen via metabolite and/or vitamin release remains ambiguous. Algal health (quantified here as C fixation) is significantly affected by individual bacterial species, and the precise metabolites responsible for these observations are currently being examined. By combining tools such microbial community analysis with advanced isotopic imaging, we are generating a sensitive picture of how these interactions can be exploited to provide reliable and renewable fuels.

A High-Quality Genome Assembly of SMRT Sequences Reveals Long-Range Haplotype Structure in the Diploid Mosquito, *Aedes aegypti*

Kingan, Sarah B.* (skingan@pacb.com)[1]; Matthews, Ben[2]; Peluso, Paul[1]; Hall, Richard[1]; Vosshall, Leslie[2]; Korlach, Jonas[1]

[1]Pacific Biosciences, Menlo Park, CA. [2]HHMI-Rockefeller University, New York, NY.

Aedes aegypti is a tropical and subtropical mosquito vector for Zika, yellow fever, dengue fever, and chikungunya. We describe the first diploid assembly of an insect genome, using SMRT Sequencing and the open-source assembler FALCON-Unzip. This project is part of a large community effort leveraging multiple technologies in order to improve the genome assembly of this important disease vector. We use ~100 fold coverage of PacBio RS II data to produce an assembly with high contiguity (contig N50 1.3 Mb), that is more complete than previous assemblies (Length 1.45 Gb with 87% BUSCO genes complete), and is high quality (mean base >QV30 after polishing). We present a novel method to identify homologous primary contigs in the FALCON-Unzip assembly using conserved gene annotations in conjunction with raw read coverage depth and pairwise nucleotide alignments. Long-range haplotype structure, in some cases encompassing more than 4 Mb of extremely divergent homologous sequence with dramatic differences in coding sequence content, is resolved using this approach.

Eukaryotic Acquisition of a Bacterial Operon Identified through Comprehensive Comparative Genomics in the Yeast Subphylum

Kominek, Jacek Z.* (jkominek@wisc.edu)[1, 4]; Doering, Drew T.[1]; Opulente, Dana A.[1]; Shen, Xing-Xing[2]; Zhou, Xiaofan[2]; DeVirgilio, Jeremy[3]; Hulfachor, Amanda B.[1]; Kurtzman, Cletus P.[3]; Rokas, Antonis[2]; Hittinger, Chris T.[1, 4]

[1]Laboratory of Genetics, Genome Center of Wisconsin, Wisconsin Energy Institute, J.F. Crow Institute for the Study of Evolution, University of Wisconsin-Madison, Madison, WI 53706, USA. [2]Department of Biological Sciences, Vanderbilt University, Nashville TN 37235, USA. [3]Mycotoxin Prevention and Applied Microbiology Research Unit, Nation Center for Agricultural Utilization Research, Agricultural Research Service, US Department of Agriculture, Peoria, IL, USA. [4]DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI 53706, USA.

High-throughput sequencing technologies create great opportunities for novel insights into the genetic basis of the phenotypic diversity observed in living organisms. Low costs and a mature software base facilitate study of genomic diversity, not only in individual organisms or genera, but in entire higher-level taxonomic groups. Here, we undertook a massive sequencing and comparative genomics effort to uncover the genetic basis of metabolic diversity present in the entire fungal subphylum of Saccharomycotina, numbering more than 1000 yeast species (Y1000+, www.y1000plus.org). This taxonomic group exhibits great phenotypic diversity at the lifestyle, metabolism and ecology levels, and includes organisms of great scientific, medical, and industrial relevance. A study of this scale offers an unprecedented opportunity to investigate the genetic basis of various traits in a deep taxonomic context.

So far, we generated novel sequencing data and genome assemblies for over 700 yeast species, of which more than 350 are of sufficient quality and completeness to be directly comparable to the currently publicly available genomes. We identified numerous novel evolutionary events affecting, amongst others, genes involved in carbon, nitrogen, amino-acid, iron and sulfur metabolisms. These events represented the entire spectrum of genetic change: we observe gain, loss, amplification, and horizontal transfer of individual genes, as well as coordinated evolution of entire gene clusters.

One of our most suprising findings is that of a bacterial siderophore biosynthesis operon present in the genomes of a subset of yeast taxa. Although the ability to synthesize siderophores is widespread in bacteria and filamentous fungi, it was thought to be completely absent from the yeast subphylum. We identified clusters of genes responsible for biosynthesis of the bacterial catecholate-type siderophore enterochelin in several closely related yeast species. We employed sequence analysis and rigorous phylogenetic hypothesis testing to determine that it was likely acquired though a single ancestral horizontal gene transfer event from an Enterobacterial lineage into the ancestor of those species. This event was followed by genomic rearrangements and, in one lineage, further translocation of a yeast ferric reductase gene into the bacterial operon. We further experimentally demonstrated that siderophore biosynthesis and release into the surroundings actually takes place *in vivo*, with one species in particular (*Starmerella bombicola*) capable of producing copious amounts of bacterial siderophores.

Species harboring the siderophore biosynthesis operon often co-inhabit insect guts alongside various Enterobacteriaceae, and in this environment the ability of cells to acquire and utilize iron is a significant growth limiting factor. Our findings provide support for an evolutionary scenario where the transferred bacterial operon was maintained by selection in a eukaryotic organism as means of adaptation to a highly-competitive ecological niche.

Although gene transfer from bacteria into eukaryotes is not unheard of, maintenance of a functional, multi-gene bacterial operon in yeast have, to our knowledge, never been reported before. The fact that entire metabolic pathways can be transferred between bacteria and yeast expands our understanding of the extent and boundaries of gene flow across kingdoms of life.

Improving Cellulosic Ethanol Production via Quantitative Fitness Measurements of Rationally Designed Gene Expression Cassettes

Krause, David J.* (dkrause2@wisc.edu)[1, 2]; Alexander, William G.[1, 2]; Wrobel, Russell L.[1, 2]; Rovinskiy, Nikolay[1, 2]; Lopes, Mariana[2, 3]; Bukhman, Yury[1]; Peris, David[1, 2]; Kang, Jin[1, 2]; Tarver, Angela[4]; Harmon-Smith, Miranda[4]; Cheng, Jan-Fang[4]; Deutsch, Samuel[4]; Landick, Robert[1]; Piotrowski, Jeff S.[1]; Sato, Trey K.[1]; Gasch, Audrey P.[1, 2]; Hittinger, Chris T.[1, 2]

[1]Department of Energy Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI. [2]Laboratory of Genetics, University of Wisconsin-Madison, Madison, WI. [3]Universidade Federal de Minas Gerais, Belo Horizonte, Brasil. [4]DOE Joint Genome Institute, Walnut Creek, CA.

The production of cellulosic ethanol requires the fermentation of glucose and xylose in the presence of lignin-derived inhibitors. While *Saccharomyces cerevisiae* is highly capable of fermenting glucose into ethanol, it is not natively capable of utilizing xylose. Previously, we have engineered a strain of S. cerevisiae capable of fermenting xylose by expressing the xylose-activating genes xylose isomerase and xylulokinase, followed by adaptive mutations generated during directed evolution experiments. This project aims to implement and test a rational design approach to improving the production of biofuels through the expression of genes selected from the broad natural diversity of yeasts and other microbes. More than two hundred gene cassettes, containing a coding sequence of interest, *S. cerevisiae*-specific promoter and terminator, and a unique 50-nucleotide sequence flanked by universal priming sites were synthesized in collaboration with the JGI. These genes represent a wide range of functions, including sugar transporters, xylose metabolic genes, and a set of genes identified from a transcriptomics study of a naturally xylose-fermenting yeast. Strains containing these cassettes were pooled a subjected to competitive growth experiments under various conditions to measure their specific fitness effects. These experiments identified a xylulokinase that increases growth and ethanol production on xylose, whose fitness effects have been validated in independent culture. Current experiments aim to better resolve fitness effects by tracking strain frequencies during longer experiments with more generations. Also, by linking genes together while preserving their unique barcode information, we are beginning to quantitatively measure the competitive fitness of strains expressing multiple genes. To identify more genes that improve xylose fermentation, we are collaborating with the JGI to sequence the genomes of twenty-four yeasts that natively ferment xylose. By using PacBio's long read technology, we aim to improve the assembly of sub-telomeric regions that are often poorly assembled using short reads, yet contain novel metabolic genes. Further, we are performing RNA-Seq experiments on a subset of these species to identify conserved transcriptional programs associated with xylose fermentation. Future experiments will combine the synthesis and expression of candidate genes identified from these new sequencing efforts with quantitative fitness measurements of combinatorial cassettes to build optimal pathways for xylose fermentation in *S. cerevisiae*.

Genomics of the First 100 Aspergilli

Kuo, Alan J.* (akuo@lbl.gov)[1]; Riley, Robert[1]; Mondo, Stephen J.[1]; Haridas, Sajeet[1]; Salamov, Asaf[1]; Korzeniewski, Frank[1]; Simmons, Blake[2]; Baker, Scott E.[2]; Andersen, Mikael[3]; Grigoriev, Igor[1]

[1]DOE Joint Genome Institute (JGI), Walnut Creek, CA. [2]DOE Joint BioEnergy Institute (JBEI), Emeryville, CA. [3]Technical University of Denmark, Kgs. Lyngby, Denmark.

Aspergillus is a ubiquitous and phenotypically diverse genus of filamentous Ascomycota, many of which play key roles as fermenters in food production, platforms for biotechnology and industrial production of enzymes and chemicals, plant and opportunistic animal pathogens, and agents of agricultural toxigenesis and biomass conversion for bioenergy. As part of a JBEI initiative to characterize the entire genus, the JGI will sequence, assemble, and annotate the genomes of each of the ~300 species of the genus *Aspergillus*. To accomplish this massive task in a timely manner without sacrificing quality, we streamlined and optimized our processes for *Aspergillus* genomes. Over the past year we have released on MycoCosm the genomes of 100 *Aspergillus* sp. which represent a broad spectrum of phylogenetic diversity and gene content, including significant variability of transporters, carbohydrate-active enzymes, proteases, and secondary metabolism clusters. The high resolution of genomic differences between closely related species is being mapped to their distinctive phenotypes to improve gene annotation in the entire genus. The next 100 species are expected soon.

Evolutionary Biology of Mushrooms: Genomics and Transcriptomics

Kwan, Hoi Shan S.* (hoishankwan@cuhk.edu.hk); Au, Chun Hang; Wong, Man Chun; Li, Lei; Lee, Yung Yung; Huang, Qianli; Nong, Wenyan; Cheung, Man Kit; Chang, Jinhui; Cheng, Xuanjin; Xiao, Yang

Food Research Centre and School of Life Sciences, The Chinese University of Hong Kong, Hong Kong, China.

We aim to understand fruiting body development and evolution of mushrooms. We carried out genomics, transcriptomics, and evolutionary analysis of the Shiitake mushroom *Lentinula edodes* and the model mushroom *Coprinopsis cinerea*. The genomes of these two mushrooms have been completely sequenced. Over 13,000 protein-coding genes were predicted from each of the genomes. Comparative analyses on genome sequences of basidiomycetes and ascomycetes revealed genes expanded in genomes of mushroom-forming fungi. Five functional categories, General function prediction only [R], Signal transduction mechanisms [T], Posttranslational modification, protein turnover, chaperones [O], Transcription [K] and Carbohydrate transport and metabolism [G], dominate in the expanded families. We examined kinome, ubiquitome, transcription factories, and CAZy enzymes. AGC kinase subfamily, F-box and paracaspase domain-containing E3 like proteins are significantly expanded in mushroom-forming genomes. We sequenced the genomes of over 100 wild and cultivated *L. edodes* strains to explore their phylogenetic relationships. Their morphologies and cultivation characteristics were also investigated. With these data, we linked some phenotypes with the genotypes using a Genome-Wide Association Studies, GWAS, approach. We performed RNA-Seq of multiple stages and identified genes differentially expressed. Transcriptome age index (TAI) profile and transcriptome divergence index (TDI) profile showed molecular hourglass patterns of the two mushrooms over their developmental stages. Young fruiting body stage is the bottleneck expressing

the evolutionarily oldest and most conserved transcriptome. This phenomenon links fungi with plants and animals. Genomic sequences of mapping populations of both mushrooms generated high density genetic maps that supported expression quantitative loci (eQTL) analyses for the mushrooms growing in lignocellulolytic substrates. We revealed potential regulatory factors that may control CAZymes expressions, and Carbon Catabolite Repression (CCR) derepression. Our works have generated rich resources for genomics and transcriptomics of mushrooms and applied them to explore the evolution of mushrooms.

Sequence-Based Analysis of the Genus *Ruminococcus* Resolves Its Phylogeny and Reveals Strong Host Association

La Reau, Alex J.* (alareau@wisc.edu)[1]; Meier-Kolthoff, Jan P.[2]; Suen, Garret[3]

[1]University of Wisconsin-Madison, Madison, WI. [2]Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany.

It has become increasingly clear that the composition of mammalian gut microbial communities is substantially diet driven. These microbiota form intricate mutualisms with their hosts, which have profound implications on overall health. For example, many gut microbes are involved in the conversion of host-ingested dietary polysaccharides into host-usable nutrients. One group of important gut microbial symbionts are bacteria in the genus *Ruminococcus.* Originally isolated from the bovine rumen, ruminococci have been found in numerous mammalian hosts, including other ruminants, and non-ruminants such as horses, pigs and humans. All ruminococci require fermentable carbohydrates for growth, and their substrate preferences appear to be based on the diet of their particular host. Most ruminococci that have been studied are those capable of degrading cellulose, much less is known about non-cellulolytic non-ruminant-associated species, and even less is known about the environmental distribution of ruminococci as a whole. Here, we capitalized on the wealth of publicly available 16S rRNA gene sequences, genomes and large-scale microbiota studies to both resolve the phylogenetic placement of described species in the genus *Ruminococcus,* and further demonstrate that this genus has largely unexplored diversity and a staggering host distribution. We present evidence that ruminococci are predominantly associated with herbivores and omnivores, and our data supports the hypothesis that very few ruminococci are found consistently in non-host-associated environments. This study not only helps to resolve the phylogeny of this important genus, but also provides a framework for understanding its distribution in natural systems.

Multiplexing Strategies for Microbial Whole Genome Sequencing Using the Sequel System

Lambert, Christine C.* (clambert@pacb.com); Harting, John; Baybayan, Primo

Pacific Biosciences, Menlo Park, CA.

For microbial sequencing on the PacBio Sequel System, the current yield per SMRT Cell is in excess relative to project requirements. Multiplexing offers a viable solution; greatly increasing throughput, efficiency, and reducing costs per genome. This approach is achieved by incorporating a unique

barcode for each microbial sample into the SMRTbell adapters and using a streamlined library preparation process.

To demonstrate performance, 8 unique barcodes were assigned to *E. coli* and 12 unique barcodes assigned to *B. subtilis* and each multiplex was sequenced on a single SMRT Cell. To further demonstrate the applicability of this method, we multiplexed the genomes of 16 strains of *H. pylori.* Each DNA was sheared to 10 kb, end-repaired and ligated with a barcoded adapter in a single-tube reaction. The barcoded samples were pooled in equimolar quantities and a single SMRTbell library was prepared.

Successful *de novo* microbial assembly was achieved from all multiplexes tested (8-,12-, and 16-plex) using data generated from a single SMRTbell library, run on a single SMRT Cell with the PacBio Sequel System, and analyzed with standard SMRT Analysis assembly methods. For microbial whole genome assembly, we describe here a multiplexing workflow that facilitates the sequencing of multiple genomes (8-,12-, and 16-plex), up to 5 Mb genome size, and sequenced in one SMRT Cell.

Multi-Omics Analysis of an Ectomycorrhizal Interface

Larsen, Peter E.* (plarsen@anl.gov)[1, 2]; Shinde, Shalaka[1]; Cumming, Jonathan[3]; Zerbs, Sarah[1]; Korajczyk, Peter[1]; Noirot, Philippe[1]

[1]Argonne National Laboratory, Bioscience Division, Argonne, IL, USA. [2]University of Illinois at Chicago, Bioengineering Department, Chicago, IL, USA. [3]West Virginia University, Department of Biology, Morgantown, West Virginia, USA.

In terrestrial ecosystems, plants are rarely solitary entities. Rather, they exist as complex meta-organisms, comprised of plant hosts and communities of intimately interacting soil microorganisms, including mutualistic fungi and bacteria. Symbiosis between soil bacteria, mycorrhizal fungi, and tree species leads to coordinated resource exchange and enhanced productivity and resiliency in forest ecosystems. These symbiotic associations provide multiple benefits to the host tree, especially under conditions of limiting nutrient availability. In return, the rhizosphere microorganisms acquire photosynthetically-derived carbon from the plant in form of sugars and organic acids. There are significant knowledge gaps in our understanding of how rhizosphere communities establish and maintain symbiotic interactions that are beneficial to the plant host. We address these knowledge gaps using a laboratory model of mycorrhizal symbiosis between *P. tremuloides* (aspen) and the ectomycorrhizal fungi *Laccaria bicolor* (*Laccaria*) and *Paxillus involutus* (*Paxillus*), deep-multiomics analysis, and machine learning approaches. Laboratory cultures of aspen seedlings were grown in sand pots, alone or in mycorrhizal association with either *Laccaria* or *Paxillus*. Rhizosphere transcriptomic, community metabolomic, and root membrane proteomic data were collected, through collaborations with JGI and EMSL. Metabolic models of the mycorrhizal community were generated using transcriptomic data and JGI genome annotations. Computational models and multi'omics data were integrated into a system-scale representations of mycorrhizal interactions. By linking transcriptomics, proteomics, and metabolomics in a single model, it is possible to predict mechanisms of mycorrhizal interaction that span multiple 'omics data type, such as the biosynthesis of signaling molecules and the regulatory receptors that detect them, post-transcriptional modifications of mRNA and post-translational modification of proteins that are predicted to occur in response to mycorrhizal interaction, and linking metabolic models to biological observations for the deeper investigation of previously 'unknown' metabolites detected by metabolomics. Using machine learning methods, the relationships between the protein profile of the mycorrhizal interface and the aspen seedling phenotype can be used to infer specific molecular mechanisms of plant-mycorrhizal fungi interaction.

Crucially, this data integration drives novel hypothesis-driven laboratory experiments and prioritizes protein targets for molecular characterization. As supported by our previous research, we find that the sensors and transporters of the ectomycorrhizal interface play a key role in understanding the molecular mechanisms of interactions in the rhizosphere and unique patterns of gene expression, membrane-bound protein abundance, and metabolite concentrations identify both general symbiosis and species-specific responses in aspen seedling roots. The computational tools used for integration of multi-omics data in this Aspen-*Laccaria/Paxillus* laboratory system, including a novel approach for determining statistically significant protein abundance measures from proteomics data, establishes a robust analysis pipeline for future multi'omics experimental datasets.

Identification of microRNA-like RNAs in *Coprinopsis cinerea*

Lau, Yuet Ting* (1155032273@link.cuhk.edu.hk); Kwan, Hoi Shan

School of Life Sciences, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong.

Introduction. *Coprinosis cinerea* is a model organism for the study of homobasidiomycete fungi due to its short life cycle and easy cultivation in the laboratory. The function and regulation mechanism of microRNA (miRNA) of animals and plants have been extensively studied. Although emerging studies have suggested that microRNA-like RNAs (milRNAs) are present in fungi and their biogenesis pathway might not be identical as conventional miRNAs in plants and animals, milRNAs are still elusive for mushroom-forming fungi. The lack of studies on the role of milRNAs in fungi hinders the improvements of mushroom strains.

Aim. We aim to understand the development of mushroom-forming fungi at the molecular level, identifying the roles of milRNAs in the development of fruiting body.

Methods. The presence of milRNAs has been validated experimentally through Northern blotting and their expression level has been quantified using Stem-loop Reverse Transcription Real-time PCR (Q-PCR). Computational methods have been applied subsequently to characterize the roles of milRNAs during the transition from mycelium to primordium stage.

Results. We have identified, from *C. cinerea* genome, 16 putative milRNAs candidates with sizes around 21nt and a group of Dicer homologs that are specific to mushroom-forming fungi. The PZA domain is present in the Dicer-like (DCLs) proteins of mushroom, while it is absent in all other fungal DCLs. Besides, for these 16 putative milRNAs candidates, one of them showed higher expression in mycelium stage and 15 showed higher expression in primordium stage.

Discussion. The mushroom-specific PAZ domains were grouped closely with the animal PAZs and plant PAZs and the Dicer_dimer domain sequences alone could distinguish fungal DCLs from their homologs in animals and plants. The DCLs domains of mushroom suggest Dicer genes duplicated and diversified independently in early evolution of all three multicellular kingdoms-animals, plants and fungi. A better understanding of the biogenesis pathways and regulatory mechanisms milRNAs in *C. cinerea* could improve the development of homobasidiomycete fungi and contribute to evolutionary studies of miRNA among kingdom in the near future.

Nitrogen flows in an Autotrophic, Anaerobic Ammonium Oxidation (Anammox) Community

Lawrence, Jennifer E.* (jelawrence@berkeley.edu)[1]; Zhou, Lijie[2]; Antell, Edmund[1]; Keren, Ray[1]; Yu, Ke[2]; Alvarez-Cohen, Lisa[1]

[1]University of California, Berkeley, CA. [2]Shenzhen University, Shenzhen, China.

Anaerobic ammonium oxidation (anammox) is the basis for an innovative treatment process for the removal of reactive nitrogen from wastewater effluent. The bacteria responsible for anammox have the unique ability to convert ammonium and nitrite to dinitrogen gas with bicarbonate as their primary carbon source. However, these bacteria are easily inhibited by varying substrate and metabolite concentrations, and have excruciatingly slow doubling times (nine days) under typical bioreactor configurations. Insights into the anammox bacterial community and its functional profiles via metagenomic and 16S analyses are essential for optimizing anammox community performance and bioreactor design.

In this study, we assembled metagenomic and 16S data to investigate reactive nitrogen flows over time in an autotrophic, anammox bioreactor. Anammox community samples were collected from a one-liter, laboratory scale anaerobic membrane bioreactor in operation at the University of California, Berkeley. Three metagenome samples and 50 iTag samples were taken over the first 365 days of the bioreactor's operation. Initial results reveal an intricately connected flow of reactive nitrogen through anammox, denitrifying, and dissimilatory nitrate reduction to ammonium (DRNA) bacteria. A better understanding of the interactions between these community members will help to inform the design of full-scale, anammox bioreactors.

Tracking Interspecific Interactions in Symbiotic Fungus Gardens Raised by the Ant *Trachymyrmex septentrionalis*

Lee, Kevin*[1]; Klassen, Jonathan L. (jonathan.klassen@uconn.edu)[1]; Fox, Evan[1]; Puckett, Sara[2]; Balunas, Marcy J.[2]

[1]Department of Molecular and Cell Biology, University of Connecituct, Storrs, CT. [2]School of Pharmacy, University of Connecticut, Storrs, CT.

Fungus-growing ants grow symbiotic fungus gardens by feeding plant material to a specific cultivar fungus, which they later consume. Other bacteria also live in these fungus gardens, but their symbiotic interactions remain poorly understood. Here, we examine symbiotic interactions that are mediated by secondary metabolites within fungus gardens raised by the ant *Trachymyrmex septentrionalis*. *T. septentrionalis* is the northernmost of all fungus-growing ants, occurring throughout the Eastern USA as far north as Long Island. We have combined metagenomics and metabolomics to identify symbiotic interactions that are mediated by secondary metabolites in the *T. septentrionalis* symbiosis. These results indicate a wide diversity of secondary metabolite biosynthetic gene clusters and metabolites in *T. septentrionalis* fungus gardens, which vary as a function of geography and symbiont diversity. Our results highlight the importance of secondary metabolites as traits that directly mediate interactions in the *T. septentrionalis* symbiosis, and the potential for combining metabolomic and metagenomics data to identify interspecific interactions in microbial communities.

Genetic Determinants of Bacterial Adaptation to Plants

Levy, Asaf* (alevy@lbl.gov)[1]; Clingenpeel, Scott[1]; Gonzalez, Isai S.[2, 3, 4]; Paredes, Sur H.[2, 3, 4]; Stillman, Kyra[1]; Monteiro, Freddy[2, 3]; Alvarez, Bryan R.[1]; Lundberg, Derek S.[2, 3, 6]; Lu, Tse-Yuan[10]; Sarah, Lebeis[9]; Jin, Zhao[11]; McDonald, Meredith[2, 3, 7]; Feltcher, Meghan E.[2, 3], del Rio, Tijana G.[1], Grant, Sarah R.[2, 3, 6], Doty, Sharon L.[8], Ley, Ruth E.[12], Pelletier, Dale A.[10], Tringe, Susannah G.[1], Woyke, Tanja[1], Dangl, Jeffery L.[2, 3,5,7]

[1]DOE Joint Genome Institute, Walnut Creek, CA, USA. [2]Department of Biology, University of North Carolina, Chapel Hill, NC, USA. [3]Howard Hughes Medical Institute. [4]Curriculum in Bioinformatics and Computational Biology, University of North Carolina, Chapel Hill, NC, USA. [5]Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC, USA. [6]Curriculum in Genetics and Molecular Biology, University of North Carolina, Chapel Hill, NC, USA. [7]Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, NC, USA. [8]School of Environmental and Forest Sciences, University of Washington, WA, USA. [9]Department of Microbiology, University of Tennessee, Knoxville, TN, USA. [10]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA. [11]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, USA. [12]Max Planck Institute for Developmental Biology, Tübingen, Germany.

Plants intimately associate with an array of diverse bacteria. Plant-associated (PA) bacteria have evolved a gene set enabling adaptation to the plant environment. However, the identity and functions of these genes are poorly characterized. Here, we sequenced 484 genomes of bacterial isolates from the roots of Brassicaceae, poplar, and maize. We then performed a large-scale comparative genomics analysis encompassing 3837 bacterial genomes to identify PA and root-associated genes and operons. PA bacterial genomes are larger, and encode more carbohydrate metabolism functions and fewer mobile elements than related non-plant associated genomes. Novel PA proteins include members of a predicted pathogen-specific type VI effector family and a phage-like secretion system. We also identified 113 PA protein domains that are apparent mimics of plant domains, many of which are also shared with PA fungi and oomycetes. This work significantly expands the genome-based understanding of plant-microbe interactions and could open new avenues of efficient and sustainable agriculture through microbiome engineering.

The Sequence of 1504 Mutants in the Model Rice Variety Kitaake Facilitates Rapid Functional Genomic Studies

Li, Guotian* (guotianli@lb.gov)[1, 2, 3]; Jain, Rashmi[1, 2, 3]; Chern, Mawsheng[1, 2, 3]; Pham, Nikki[1, 2]; Martin, Joel[4]; Wei, Tong[1, 2, 3]; Schackwitz, Wendy[4]; Lipzen, Anna[4]; Duong, Phat[1]; Jones, Kyle[1, 2]; Jiang, Liangrong[1, 5]; Ruan, Deling[1, 2, 3]; Bauer, Diane[4]; Barry, Kerrie[4]; Schmutz, Jeremy[4,6]; Ronald, Pamela[1,2,3]

[1]Department of Plant Pathology and the Genome Center, University of California, Davis, CA 95616, USA. [2]Feedstocks Division, Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [3]Environmental Genomics and Systems Biology, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [4]DOE Joint Genome Institute, Walnut Creek, CA 94598, USA. [5]School of Life Sciences, Xiamen University, Xiamen 361102, China. [6]HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA.

The availability of a whole-genome sequenced mutant population and the cataloging of mutations of each line at a single-nucleotide resolution facilitates functional genomic analysis. To this end,

we generated and sequenced a fast-neutron-induced mutant population in the model rice cultivar Kitaake (*Oryza sativa* L. ssp. japonica), which completes its life cycle in 9 weeks. We sequenced 1,504 mutant lines at 45-fold coverage and identified 91,513 mutations affecting 32,307 genes, 58% of all rice genes. We detected an average of 61 mutations per line. Mutation types include single base substitutions, deletions, insertions, inversions, translocations, and tandem duplications. We observed a high proportion of loss-of-function mutations. Using this mutant population, we identified an inversion affecting a single gene as the causative mutation for the short-grain phenotype in one mutant line with a small segregating population. This result reveals the usefulness of the resource for efficient identification of genes conferring specific phenotypes. To facilitate public access to this genetic resource, we established an open access database called KitBase that provides access to sequence data and seed stocks, enabling rapid functional genomic studies of rice.

Unraveling the Core Mechanisms of *Suillus-Pinus* Symbiosis through Time-Course RNAseq and Omics *(Short Talk)*

Liao, Hui-Ling (Sunny)* (sunny.liao@ufl.edu)[1, 2]; Talbot, Jennifer M.[3]; Bruns, Tom[4]; Taylor, John[4]; Peay, Kabir[5]; Branco, Sara[6]; Hoyt, David W.[7]; Purvine, Samuel[7]; Kuo, Alan J.[8]; Barry, Kerrie[8]; Grigoriev, Igor[8]; Vilgalys, Rytas[2]

[1]University of Florida, GNV, FL. [2]Duke University, Durham, NC. [3]Boston University, Boston, MA. [4]University of California, Berkeley, CA. [5]Stanford University, Stanford, CA. [6]University of Paris, Paris. France. [7]EMSL Pacific Northwest Laboratory, WA. [8]DOE Joint Genome Institute, Walnut Creek, CA.

Despite the importance of ectomycorrhizal fungi (EMF) and other soil fungi to forest health, we still know little about mechanisms by which EMF associate and communicate with their plant hosts. We are using the *Pinus-Suillus* EMF symbiosis as a model for study of host-specific interactions and their effects on C/N/P fluxes in forest soils. Here, we applied integrated multiple 'omics approaches to characterize sets of interacting genes/ between compatible *Suillus-Pinus* pairings (*S. spraguei x P. strobus; S. decipiens x P. taeda*) versus incompatible pairings (*S. spraguei x P. taeda; S. decipiens x P. strobus*). Combined 'omics approaches include *de novo* fungal genome sequencing and assembly (JGI), metatranscriptomics (JGI) and proteome/metabolome profiling (by EMSL) to identify genes/transcripts/proteins/metabolites. Using time-course studies, genes and gene products involved in early and late stages of fungal-plant communication and nutrient reallocation were identified. For example, a novel Pinus gene group coding for 1-helix-1-signalp-microbe-induced like proteins were upregulated 100 to 6000-fold in EMF-roots versus control roots, suggesting a role as key signaling molecules involved in maintaining EMF symbiosis. The enrichment of plant RNA and gene products for flavonoid pathway and defense (e.g. LRR, PR, peroxidase, JA pathway. chitinase) in incompatible roots also suggests their involvement in *Pinus-Suillus* incompatibility. Integration of transcriptomic-proteomic-metabolomic data confirm fructose metabolites to be the major forms of plant carbon relocated to the fungal symbiont followed by conversion to mannitol, trehalose and arabinitol. Several other highly expressed gene groups were identified (e.g. UDP-glucosyltransfereses, involved in catalyzing plant oligosaccharide, inorganic phosphate, water or protein residues during symbiotic exchange). Integrated transcriptomic/ proteomic studies also reveal extracellular C/N degradation activities involving *Suillus* mycorrhizae, and suggest a more active role for *Suillus* in (soil) nutrient capture. Combined, these integrated 'omics approaches provide functional insights into mycorrhizal symbiosis involving EMF.

An Enzyme Cocktail with a Cellulose Booster

Lin, Yu-Ju (Lulu)* (superolulu@gmail.com)[1, 4]; M., Anandharaj[1, 2, 3]; Lu, Mei-Yeh[1]; Huang, Chieh-Chen[4]; Chang, Jui-Jen[5]; Li, Wen-Hsiung[1]

[1]Biodiversity Research center, Academia Sinica, Taipei 115, Taiwan. [2]Molecular and Biological Agricultural Sciences Program, Taiwan International Graduate Program, National Chung Hsing University and Academia Sinica, Taipei, Taiwan. [3]Graduate Institute of Biotechnology, National Chung Hsing University, Taichung, Taiwan. [4]Department of Life Sciences, National Chung Hsing University, No. 250, Kuo Kuang Rd, Taichung 402, Taiwan. [5]Department of Medical Research, China Medical University Hospital, China Medical University, No.91 Hsueh-Shih Road, Taichung 402, Taiwan.

Lignocellulosic biomass is a potential feedstock for a cellulolytic biorefinery to produce valuable products such as biofuel and biochemicals. However, degradation of plant biomass into simple sugars is a major obstacle. To address this issue, researchers have developed biomimic hosts displaying multiple cellulase enzymes, which can degrade the crystalline cellulose. Apart from the hydrolytic enzymes, a new class of copper contained-oxidative enzyme 'lytic polysaccharide monooxygenases' (LPMO) and its co-enzyme cellobiose dehydrogenase (CDH) can boost the activity of cellulases to effectively degrade the crystalline cellulose. Hence, we studied the synergistic effect of cellulase booster via co-transforming three endo-glucanases (EgI, EglA and EgIII), two exo-glucanases (CBHI and CBHII), and a β-glucosidase (NpaBGS) along with the cellulase boosters TaLPMO and MtCDH by genomic integration. The cellulase booster containing strain increased the sugar release to 1.5-fold relative to the control strain. Our result showed that the cellulase booster can significantly improve the cellulolytic activity of the designer strain, and has potential for biofuel application.

Identification of Uncultured and Candidate Lineages as Viable Microorganisms in the Deep Terrestrial Biosphere

Lopez-Fernadez, Margarita* (margarita.lopezfernandez@lnu.se)[1]; Simone, Domenico[1]; Broman, Elias[1]; Wu, Xiaofen[1]; Bertilsson, Stefan[2]; Dopson, Mark[1]

[1]Centre for Ecology and Evolution in Microbial Model Systems (EEMiS), Linnaeus University, Kalmar, Sweden. [2]Department of Ecology and Genetics, Limnology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden.

The deep biosphere is the largest 'bioreactor' on earth, containing an estimated 2 to 19% of the total terrestrial biomass. Nevertheless, due to the difficulty of obtaining samples it is one of the least understood ecosystems on the planet. Microorganisms inhabiting this biome strongly influence the global nutrient and energy cycles. An important question for deep biosphere microbiology is not only its diversity but also whether or not specific populations are viable or non-viable. The Äspö Hard Rock Laboratory provides access to investigate the microbial life in the deep terrestrial subsurface. Previous studies at the site revealed the presence of microbial anaerobes, putative H2 oxidizers and NO_3^-, Fe_3^+, SO_4^{2-}, and Mn_2^+ reducers along with acetogens and methanogens. The current work expands on these earlier findings to identify the viable (i.e. having an intact cellular membrane) and non-viable (i.e. having a compromised membrane) populations in three aquifers with different chemistries and depth below the surface. High throughput 16S rRNA gene sequencing of total and viable cells revealed significant differences between the total and the viable subset of the community and that the viable diversity decreased with temporal separation from the surface. The viable population was

mainly related to uncultured candidate phyla such as Parcubacteria (OD1), Gracilibacteria (GN02), Atribacteria (OP9), Katanobacteria (WWE3), Microgenomates (OP11), Candidate TA06, and unclassified Euryarchaeota. Our results emphasize the importance of unclassified and candidate phyla and the need to further investigate microbial activities in the deep biosphere.

Multiplication Is Vexation: A Genomic Perspective on Cell Division and DNA Replication in the Large Sulfur Bacteria

MacGregor, Barbara J.* (bmacgreg@unc.edu)[1]; Flood, Beverly[2]; Bailey, Jake[2]; Kanke, Matthew[3]

[1]Department of Marine Sciences, University of North Carolina - Chapel Hill, Chapel Hill, NC. [2]Department of Earth Sciences, University of Minnesota, Minneapolis, MN. [3]Department of Genetics, University of North Carolina - Chapel Hill, Chapel Hill, NC.

The large sulfur bacteria (LSB) are a morphologically diverse group of sulfide-oxidizing Beggiatoaceae (Gammaproteobacteria), including free-living and sheathed filaments, vacuolated and non-vacuolated species, and some of the largest known bacteria. The family Beggiatoaceae includes the genera Beggiatoa, Thioploca, Thiomargarita, Marithrix, and several Candidatus groups. They are found at sulfidic/oxic interfaces in both marine and freshwater settings, and may store nitrate, elemental sulfur, polyphosphate, and/or carbohydrates, allowing them to tolerate environmental fluctuations or migrate between sulfidic and oxidizing conditions. Only two species are in cultivation, both non-vacuolated freshwater Beggiatoa spp., but their large size has allowed several others to be collected for genome sequencing. Comparison with related Gammaproteobacteria reveals several changes that may have been key to the evolution of the LSB, in particular the loss of canonical genes for septum formation and for DNA replication initiation by DnaA. The evolutionary origin and mechanism of division of the central vacuoles found in some LSB remains a puzzle; we suggest acquisition of dynamin-family proteins may have been one key step. LSB genetic traits related to DNA replication have parallels to those in the Cyanobacteria, another morphologically diverse group, with which the LSB appear to have a history of gene exchange. From these observations, we propose a model for the evolution of the LSB. We hope that our new Community Sequencing Program project will fill in this picture by expanding the LSB genome sequence collection to represent additional morphologies and habitats, and perhaps by identifying epibiont species that may have exchanged genes with the LSB.

Wisconsin Crop Innovation Center: A New Public Resource for Plant Genome Research

Malueg, Cera Danielle* (cdmalueg@wisc.edu)[1, 2, 3]; Petersen, Mike[1, 3]; Martinell, Brian[1, 3]; Miyamoto, Amy[1, 3]; Harnish, Rob[1, 3]; Mahoy, Jill[1, 3]; Lor, Vai[1, 3]; Spalding, Edgar[1, 3]; Miller, Nathan[1, 3]; Kaeppler, Heidi[1, 2, 3]; Kaeppler, Shawn[1, 2, 3]

[1]University of Wisconsin - Madison, Madison, WI, USA. [2]Great Lakes Bioenergy Research Center, Madison, WI, USA. [3]Wisconsin Crop Innovation Center, Middleton, WI, USA.

The Wisconsin Crop Innovation Center (WCIC) is a new public research and service facility generously donated by Monsanto to the University of Wisconsin-Madison. The mission of WCIC is to advance basic and applied translational and functional genomic research in crop science. WCIC was initiated

to dramatically increase U.S. research capacity through technology development, collaboration, and offer fee-for-services including transformation, genome editing, and phenotyping. We intend that WCIC will work together with other public crop biotechnology facilities to increase research capacity and accelerate public agricultural research outputs. WCIC is seeking collaborations with partners that have the need for large-scale projects. Initial fee-for-service rates for specific services are available on our website (cropinnovation.cals.wisc.edu), with additional services to be added over time. We are willing to discuss start-up collaboration rates for significant projects with partners that can provide constructs. Partners would then receive final product in the form of seed for further research, with options of plantlets or callus available upon inquiry. We are developing customizable constructs including over-expression, RNAi and genome editing for numerous crop plants transformed at the WCIC. Initial species of focus include soybean, maize, sorghum, and medicago. The WCIC team has the expertise, and the facility has the capacity, for most crop species and crop biotechnology approaches, so please inquire if you have projects of interest.

https://cropinnovation.cals.wisc.edu

Genome-Resolved Metagenomic and Geochemical Analysis of East River Riparian Zone Soils Supports the 'Systems within Systems' Approach for Watershed Analysis

Matheus Carnevali, Paula B.* (pmatheus@berkeley.edu)[1]; Williams, Kenneth H.[2]; Dong, Wenming[2]; Thomas, Brian C.[1]; Hubbard, Susan[2]; Banfield, Jillian F.[1, 2, 3]

[1]Department of Earth and Planetary Sciences, University of California, Berkeley, CA, USA. [2]Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [3]Department of Environmental Science, Policy and Management, University of California, Berkeley, CA, USA .

Microorganisms largely control the turnover of carbon and mediate other important biogeochemical processes in soils. Information regarding microbial impacts at larger spatial scales is required for understanding of ecosystem processes. Our approach to scaling genome-resolved metagenomics to study the riparian zone was to target discrete reaches of the East River as a potentially representative repeating motif. We are testing this hypothesis using extensive sampling and DNA sequencing based analyses. 94 soil samples were collected over the 10-25 cm depth interval from floodplains associated with three meandering reaches: upstream (MG), midstream (ML), and downstream (MZ). Genomic DNA was extracted and then sequenced at the Joint Genome Institute on a HiSeq Illumina instrument (~ 5 Gbp per sample). DNA sequence information from individual samples was processed and assembled using standard methods. Based on profiling of community composition using scaffolds encoding rpS3, we identified 813 distinct strains, many of which were found at multiple sites within a meander and also occurred in two or all three meanders. Although the data were only assembled approximately one week prior to this writing, we have recovered 164 draft or higher quality genomes that represent 88 distinct microorganisms. Most of the relatively abundant organisms in the three systems are represented by draft genomes, with the main exception being some members of a group of highly abundant and closely related strains of Betaproteobacteria. For many organisms, the same genome was repeatedly recovered from multiple samples. For example, for one novel Betaproteobacteria we reconstructed 11 near complete genomes and a genotype of Gammaproteobacteria is represented by 11 genomes. Other groups of abundant, novel and genomically defined Bacteria include Acidobacteria, Ignavibacteria, Nitrospirae, Gemmatimonadetes, Deltaproteobacteria, Verrucomicrobia, and Planctomycetes. Thaumarcheota (Nitrosopumilales) were

the most abundant Archaea observed in these samples. In addition, we recovered genomes for a variety of very little understood organisms known only from prior genomic studies, mostly at the Rifle SFA research site. These include Rokubacteria, WS3X, CHLX, Zixibacteria, OD1 (various Parubacteria), and DPANN (Pacearchaeota). Widely distributed functions in addition to carbon turnover include sulfur oxidation, iron oxidation, ammonia oxidation, nitrate reduction and nitrogen fixation. Despite some heterogeneity in species distribution, there are strong overlaps in organism membership and similar functions are well represented across the three floodplain localities; however, different organisms may be responsible for some functions at the different sites. Overall, we conclude that genome-resolved metagenomics, when analyzed in the context of physical and chemical information, can be scaled to provide information about the distribution of ecosystem processes. Insights regarding community composition from a single floodplain site may be generalizable, at least under some conditions.

Zmm22 Gene in Maize Has Pleiotropic Effects on Traits Important for the Production of Food, Feed, and Fuel

Mazaheri, Mona M.* (mmazaheri@wisc.edu)[1, 2]; Burdo, Brett[1]; Heckwolf, Marlies[1, 2]; Vaillancourt, Brieanne[3, 4]; Gage, Joseph[1]; Buell, Robin C.[3, 4]; de Leon, Natalia[1, 2]; Kaeppler, Shawn[1, 2]

[1]Department of Agronomy, University of Wisconsin, Madison, Wisconsin 53706. [2]Department of Energy, Great Lakes Bioenergy Research Center, University of Wisconsin, Madison, Wisconsin 53706. [3]Department of Plant Biology, Michigan State University, East Lansing, Michigan 48824. [4]Department of Energy, Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, Michigan 48824

Replacing fossil fuels with sustainable energy sources is a major current challenge. This challenge can be met by utilizing agricultural residues such as corn stover. With an annual production of about 300 million tons, corn stover is a major source of biofuel feedstock in the US. However, to replace 30% of transportation fuel with biofuels, stover biomass needs to be increased 40% by 2030. To achieve this goal, characterizing the genetic mechanism underlying stover biomass is essential. In this study, we investigated the natural genetic variation among maize inbred lines to detect candidate genes associated with biomass traits in genome wide association studies (GWAS). In doing so, we assembled a panel of 835 inbred lines and phenotyped these lines for major biomass traits such as stalk diameter, plant height, leaf number, and flowering time. The population was genotyped with 430,947 RNA-Seq based single nucleotide polymorphism (SNP) markers. Our results showed that *zmm22* is the most significant candidate gene associated with all biomass traits measured in this study. To explain the pleiotropic effect of *zmm22*, we proposed that elevated expression of *zmm22* later in plant development triggers the suppression of vegetative growth and the activation of flowering transition. We supported this hypothesis by modifying the expression level of *zmm22*. Transgenic lines with high expression of *zmm22* had reduced vegetative growth; they flowered four days earlier and showed a decrease in stalk diameter, leaf number, and plant height by 14%, 10%, and 9% respectively, compared to the wild types. In turn, reduced expression of *zmm22* enhanced the biomass traits. Our results demonstrated that *zmm22* is a central gene in regulating maize development. Detecting this gene could open promising opportunities to increase stover biomass for sustainable energy production without compromising grain yield for food and feed.

Methane Cycling by Novel Archaea in Yellowstone National Park Hot Springs

McKay, Luke J.* (mcgucas@gmail.com)[1, 2]; Hatzenpichler, Roland[3]; Fields, Matthew W.[2, 4]; Inskeep, William P.[1]

[1]Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA. [2]Center for Biofilm Engineering, Montana State University, Bozeman, MT, USA. [3]Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT, USA. [4]Department of Microbiology and Immunology, Montana State University, Bozeman, MT, USA.

The discovery of near-complete methanogenesis pathways in the genomes of members of recently described phyla Bathyarchaeota and Verstraetearchaeota has drawn into question whether the ability to generate or oxidize methane and other hydrocarbon gases is limited to a single archaeal phylum, the Euryarchaeota. This phylum-level taxonomic expansion supports the hypothesis that methanogenesis could be one of the first metabolic modes of life because an early metabolism would likely leave genetic remnants across broad, deeply-rooted lineages. Washburn Hot Springs (WS) in Yellowstone National Park (YNP) is an anoxic, highly sulfidic geothermal system where source gases are comprised almost entirely of carbon dioxide (~92%), hydrogen (~2%), and methane (~6%)—one of the highest known methane fluxes in YNP. In subsurface sediments at WS, porewater methane concentrations are between 50 – 100 ppm. Stable isotopic signatures of methane gas from WS have suggested that the majority of methane flux is geogenic, but these high levels of methane could also reflect a contribution from methanogens and/or support anaerobic methanotrophy even if the major flux is not biotic. We employed a combined approach of targeted amplicon and random metagenomic sequencing at WS to investigate the likelihood that novel methane cycling archaea are present and/or active in this extreme environment. Our data indicate a predominance of Bathyarchaeota in WS and a wide diversity of mcrA genes related to the Euryarchaeota, Verstraetearchaeota, and Bathyarchaeota. Sequences detected from RNA demonstrate that mcrA genes are actively transcribed by all three phyla. Further, reconstructed genomes yielded information regarding the ubiquity and phylogeny of single carbon processing genes across the archaeal domain. We interpret these results as an indication that previously unidentified archaea are likely major contributors to global fluxes of methane, and that the WS system is replete with the ingredients for a chemolithotrophic origin of life: abundant carbon dioxide and hydrogen, intense thermal activity, and a cesspool of deeply-rooted archaea involved in single carbon metabolism.

Operation-Driven Heterogeneity and Overlooked Feed-Associated Populations in Anaerobic Digester Microbiome

Mei, Ran* (ranmei2@illinois.edu)[1]; Nobu, Masaru[1]; Narihiro, Takashi[2]; Kuroda, Kyohei[1, 3]; Sierra, Julian Muñoz[4]; Wu, Zhuoying[5]; Ye, Lin[6]; Lee, Patrick[7]; Lee, Po-Heng[5]; van Lier, Jules[4]; McInerney, Michael[8]; Kamagata, Yoichi[2]; Wen-Tso Liu[1]

[1]Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [2]Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan. [3]Department of Environmental Systems Engineering, Nagaoka University of Technology, Kami-tomioka, Niigata, Japan. [4]Section Sanitary Engineering, Department of Water Management, Delft University of Technology, Delft, The Netherlands. [5]Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. [6]School of the Environment, Nanjing University,

Nanjing, Jiangsu, China. [7]School of Energy and Environment, City University of Hong Kong, Kowloon, Hong Kong. [8]Department of Botany and Microbiology, University of Oklahoma, Norman, OK, USA.

Anaerobic digestion (AD) is an essential waste(water) treatment technology that not only biodegrades waste but also produces methane as a sustainable energy source. Microbes involved in this process include fermenters, syntrophs and methanogens as three conventionally recognized major guilds. While these endogenous anaerobic microorganisms with diverse functions interact closely to decompose complex organic matters and produce biogas, exogenous aerobic microorganisms have often been observed in AD with high abundance with unknown functions. In this study, 148 anaerobic digester sludge samples are taken from 51 municipal wastewater treatment plants in North America, Japan, Hong Kong, and the Netherlands. The genomic DNA extracted from individual sludge samples are used for analyses of 16S rRNA gene through Illumina Miseq sequencing. A total of 7.04 million effective sequences are obtained, which generate 263 thousand operational taxonomic units (OUTs at 97% sequence similarity). Clustering analyses based on weighted UniFrac distance matrix and UPGMA algorithm demonstrate that AD microbiome could be grouped into eight clusters, corresponding to operation conditions instead of geological location. By comparing the AD microbiome and the upstream feed sludge microbiome, we observe that the presence of exogenous microbial populations in AD is a common phenomenon in all clusters of AD communities. These populations are likely resulted from incomplete digestion of feed sludge or residues, as they are more affiliated with aerobic or facultative anaerobic microbes that are mostly dominant populations found in the feed sludge. Their abundances decrease after entering AD but could persist to different extent. Some populations presumably possess capacity of anaerobic respiration and prolong their persistency, which requires future in-depth studies. These populations account for around 10% of populations in digesters operated under regular conditions (mesophilic, 20-40 days of sludge retention time), and increased to 35% in digesters operated at low temperature (e.g., under 30°C). Pre-treatment technologies, such as thermal hydrolysis, could effectively reduce the abundance of these populations down to 1.5%. Overall, these populations could represent an indicator of efficiency of cell lysis, which is the first but challenging step of the entire digestion process, to guide the improvement of AD.

Combining Sequencing, Synthesis, and Evolution to Discover and Optimize Pathways for Lignin Bioconversion

Michener, Joshua K.* (michenerjk@ornl.gov); Clarkson, Sonya M.; Kridelbaugh, Donna; Elkins, James G.; Guss, Adam M.

Oak Ridge National Laboratory, Oak Ridge, TN.

Cellulosic biofuel production yields a substantial lignin byproduct stream that currently has few applications. Biological conversion of lignin compounds into chemicals and fuels has the potential to improve the economics of cellulosic biofuels. Microbial strains have been isolated with the ability to specifically degrade challenging lignin dimers, and the associated pathways could substantially improve the conversion of lignin into valuable biochemicals. However, no dimer degradation pathway has been characterized to the level necessary for reconstruction in a heterologous host, largely due to the combined challenges of fully identifying the relevant enzymes and then optimizing the resulting pathways for heterologous function. We use a barcoded transposon library to rapidly identify entire catabolic pathways in genetically-intractable microbes such as the lignin-degrading strain Novosphingobium aromaticivorans. Simultaneously, we are reconstructing pathways for lignin

catabolism in Escherichia coli to enable pathway characterization in a tractable and well-defined host. Evolutionary optimization of these heterologous pathways identifies factors limiting pathway activity and will allow the design of characterized, optimized synthetic operons for lignin dimer degradation.

Pervasive Adenine N6-methylation of Active Genes in Fungi *(Short Talk)*

Mondo, Stephen J.* (sjmondo@lbl.gov)[1]; Dannebaum, Richard O.[1]; Kuo, Rita C.[1]; Louie, Katherine[1]; Bewick, Adam J.[2]; LaButti, Kurt[1]; Haridas, Sajeet[1]; Kuo, Alan J.[1]; Ahrendt, Steven R.[1, 3]; Lau, Rebecca[1]; Bowen, Ben[1]; Lipzen, Anna[1]; Sullivan, William[1]; Andreopoulos, William B.[1]; Clum, Alicia[1]; Lindquist, Erika[1]; Daum, Chris[1]; Northen, Trent R.[1]; Ramamoorthy, Govindarajan K.[1]; Schmitz, Robert J.[2]; Graganskyi, Andrii[4]; Culley, David[5]; Magnuson, Jon[5]; James, Timothy Y.[6]; O'Malley, Michelle A.[7]; Stajich, Jason E.[8]; Spatafora, Joseph W.[9]; Visel, Axel,[1, 10]; Grigoriev, Igor V.[1, 3]

[1]DOE Joint Genome Institute, Walnut Creek, CA, USA. [2]Department of Genetics, University of Georgia, Athens, GA, USA. [3]Department of Plant and Microbial Biology, University of California, Berkeley, CA, USA. [4]L. F. Lambert Spawn Co, Coatesville, PA, USA. [5]Pacific Northwest National Laboratory, Richland, WA, USA. [6]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA. [7]Department of Chemical Engineering, University of California, Santa Barbara, Santa Barbara, CA, USA. [8]Department of Plant Pathology and Microbiology, University of California, Riverside, California, USA. [9]Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon, USA. [10]School of Natural Sciences, University of California, Merced, CA, USA.

N6-methyldeoxyadenine (6mA) is a non-canonical DNA base modification present at low levels in plant and animal genomes, but its prevalence and association with genome function in other eukaryotic lineages remains poorly understood. Here we report that abundant 6mA is associated with transcriptionally active genes in early-diverging lineages of the fungal kingdom. Using single-molecule long-read sequencing of 16 fungal genomes, we observed that up to 2.8% of all adenines were methylated in early-diverging fungi, far exceeding the levels observed in other eukaryotes and more derived fungi. 6mA occurred symmetrically at ApT dinucleotides, was concentrated in dense 'Methylated Adenine Clusters' (MACs) surrounding the transcriptional start sites (TSSs) of expressed genes, and its distribution was inversely correlated with 5mC. Our results reveal a striking contrast in the genomic distribution of 6mA and 5mC and reinforce a distinct role of 6mA as a gene expression-associated epigenomic mark in eukaryotes.

Interpopulation Mating to Associate the Phenotype of Growth at Low Temperature With Specific Genes in *Neurospora crassa* Populations

Montoya, Liliam A.* (liliam_m007@berkeley.edu)[1]; Catcheside, David[2]; Ellison, Chris[3]; Freitag, Michael[4]; Gladieux, Pierre[5]; Daskalov, Asen[1]; Brem, Rachel[6]; Schackwitz, Wendy[7]; Hann-Soden, Christopher[1]; Glass, Lousie N.[1]; Taylor, John[1]

[1]University of California, Berkeley, Plant and Microbial Biology, Berkeley, CA, 94720-3102, USA. [2]Flinders University, School of Biological Sciences, GPO Box 2100, Adelaide 5001, South Australia. [3]University of California, Berkeley, Integrative Biology, Berkeley, CA, 94720-3102, USA. [4]Oregon State University, Department of Biochemistry and Biophysics, Corvallis, OR 97331-7305. [5]Campus International de Baillarguet, UMR BGPI, INRA,

34398 Montpellier, France. [6]Buck Institute, 8001 Redwood Blvd. Novato, CA 94945. [7]DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA, 94598.

Understanding causative gene differences that confer population variation of complex traits is crucial in the study of evolutionary biology. The search for genes that have been selected to promote adaptation using a "reverse ecology" approach detects regions of high differentiation referred to as "islands" of differentiation that arise by hybridization and introgression due to the presence of adaptive genes. Previous researchers used this approach to make and test hypotheses about genes that promote adaptation to low temperature in populations of *Neurospora crassa* from subtropical Louisiana. There, the average yearly minimum temperature is 9°C cooler than the Caribbean basin. Two genomic "islands" in Chr 3 and 7 were detected; genes related to cold temperature response where located in these islands. However, the reverse ecology approach cannot be used to identify genes with more subtle signatures of selection. To find additional genes involved in adaptation to growth at low temperature, we mated strains from Louisiana and the Caribbean, selected 250 progeny, evaluated their growth at 10°C and 25°C, and sequenced the parental and progeny genomes. We performed QTL analysis using SNPs scored in the progeny genomes and the phenotype of the ratio of growth rates at 10°C and 25°C. The analysis showed regions of significant interest, particularly in Chr 5, where we can now study the population variation in the parental genomes for candidate genes. We plan on testing the detected genes using the *Neurospora crassa* gene deletion collection.

http://www.pnas.org/content/108/7/2831.abstract

The Role of Priming Effects on the Conversion of Blue Carbon to CO in the Coastal Zone

Morrison, Elise S.* (emorrison@ufl.edu)[1]; Bianchi, Thomas[2]; Ward, Nicholas[3]; Liu, Yina[4]; Arellano, Ana[2]; Osborne, Todd[1]; Ogram, Andrew[1]

[1]University of Florida, Dept. of Soil and Water Sciences, Gainesville, FL, USA. [2]University of Florida, Dept. of Geological Sciences, Gainesville, FL, USA. [3]Pacific Northwest National Laboratory Marine Sciences Laboratory, Sequim, WA, USA. [4]Pacific Northwest National Laboratory Environmental Molecular Sciences Laboratory, Richland, WA, USA.

Blue carbon, or the organic carbon (OC) stored within coastal ecosystems, has been increasingly recognized as an important, but vulnerable, global carbon sink. Blue carbon habitats, particularly those along low-relief coastlines such as Florida, are susceptible to erosion and carbon loss as sea levels rise. Peat-derived organic matter (OM) may be degraded within downstream estuarine systems, especially via priming effects in the presence of labile algal-derived OM. We conducted a series of incubation experiments to determine the fate of peat-derived organic carbon in the presence of labile, algal-derived OM. Four treatments were established: seawater with peat and algal leachate (SWPA), seawater and peat leachate (SWP), seawater and algal leachate (SWA), and seawater alone (SW). Treatments containing peat leachate (SWPA and SWP) harbored greater total DOC concentrations compared to SWA and SW treatments. Total dissolved nitrogen (TDN) concentrations tended to decrease over the course of the incubation, and significantly greater TDN concentrations were seen in the SWPA treatment compared to the SW treatment. Over the course of the incubation, CO_2 concentrations increased in all treatments, with the highest CO_2 levels in treatments with algal-derived DOM (SWA and SWPA). Both treatments with algal-derived DOM (SWA and SWPA) treatments showed an increase in 13C-labeled CO_2 over the course of the incubation, and stable isotope mass balance

indicated that the conversion of peat-derived OC to CO_2 occurred approximately 30% faster with the presence of algal-derived DOC. Aromaticity indices from absorption spectra were significantly lower in the SWP treatment when compared to the SWPA treatment. Dissolved organic matter molecular formulae detected by Fourier-transformed ion cyclotron resonance spectrometry indicated an increase in the degradation of peat-derived compounds when algal material was present. These findings suggest that there is an increase in microbial degradation of peat when in the presence of algal-derived DOM. Further analysis of microbial community composition will be conducted with metagenome analysis at the Joint Genome Institute to identify taxa that may drive the conversion of blue carbon to CO_2 within primed systems.

Artificial Microbiome-Selection to Engineer Microbiomes that Confer Salt-Tolerance to Plants

Mueller, Ulrich G.* (umueller@austin.utexas.edu)[1]; Kardish, Melissa R.[1]; Smith, Chad C.[1]; Juenger, Thomas E.[1]; Des Marais, David L.[2]; Tringe, Susannah G.[3]

[1]Department of Integrative Biology, University of Texas at Austin, 1 University Station #C0990, Austin TX 78712, USA. [2]Department of Organismic & Evolutionary Biology, Arnold Arboretum, Harvard University, 1300 Centre Street, Boston MA 02131, USA. [3]DOE Joint Genome Institute, Walnut Creek, CA 94598, USA.

We use a differential microbiome-propagation method to artificially select for rhizosphere microbiomes that confer salt-tolerance to the model grass *Brachypodium* distachyon. We combine several protocol steps to maximize evolutionary changes due to differential microbiome-propagation, while minimizing some (but not all) ecological processes affecting microbiome composition. Specifically, our methods aim to improve microbiome perpetuation and response to artificial microbiome-selection by (a) controlling microbiome assembly when inoculating seeds; (b) low-carbon soil to enhance host-control during initial microbiome assembly and subsequent microbiome persistence; (c) microbiome-fractionation to propagate and select only on bacterial and viral (but not fungal) microbiome components; and (d) ramping of salt-stress between selection-cycles to minimize the chance of either under-stressing or over-stressing plants. Depending on salt-stress and control treatments, our protocol generates microbiomes that enhance plant fitness after only 1-3 rounds of differential microbiome propagation. When testing microbiomes after nine rounds of differential microbiome propagation, the effect of bacterial microbiomes selected to confer tolerance to sodium-salt stress appears specific (these microbiomes do not confer such tolerance under aluminum-salt stress), but the effect of microbiomes conferring tolerance to aluminum-salt stress appears non-specific (selected microbiomes ameliorate both sodium- and aluminum-salt stresses). We have posted a pre-publication at www.biorxiv.org/content/biorxiv/early/2016/10/17/081521.full.pdf .Ongoing metagenomic analyses of the artificially selected microbiomes will help elucidate metabolic properties of microbiomes that confer specific versus non-specific salt-tolerance to plants.

Genomes OnLine Database (GOLD): A Curated Catalogue of Genome and Metagenome Sequencing and Analysis Projects

Mukherjee, Supratim* (supratimmukherjee@lbl.gov); Stamatis, Dimitrios; Bertsch, Jon; Ovchinnikova, Galina; Kyrpides, Nikos; Reddy, T.B.K.;

DOE Joint Genome Institute, Walnut Creek, CA, USA.

The Genomes Online Database (GOLD) is a freely available database of sequencing projects, which provides an up-to-date status of finished and ongoing sequencing projects along with their associated metadata. GOLD provides a simplified web interface for users to access genome reports and launch advanced search tools and is available at the following URL: https://gold.jgi.doe.gov. There are three different sources for projects in GOLD: internal projects from the Department of Energy Joint Genome Institute (DOE-JGI), external projects entered by GOLD users and projects sourced from public databases such as NCBI. The current version of GOLD is based on a four level classification system in the form of a Study, Biosample (Organism), Sequencing Project and Analysis Project. As of Feb 2017, GOLD contains information for nearly 28,750 studies, 130,450 sequencing projects, 19,300 biosamples, 270,000 organisms and 109,500 analysis projects. The GOLD web interface facilitates submission of a diverse range of sequencing projects (such as isolate genome, single-cell genome, metagenome, metatranscriptome) and complex analysis projects (such as genome from metagenome, combined assembly from multiple sequencing projects). It provides a seamless interface with the Integrated Microbial Genomes (IMG) family of analysis tools and also supports and promotes the Genomic Standards Consortium (GSC) Minimum Information standards.

Targeted Exploration of Archaeal Genomic, Phylogenetic, and Spatial Diversity in Freshwater Wetland Soils

Narrowe, Adrienne B.* (adrienne.narrowe@ucdenver.edu)[1]; Angle, Jordan C.[2]; Daly, Rebecca A.[2]; Stefanik, Kay C.[3]; Wrighton, Kelly C.[2]; Miller, Christopher S.[1]

[1]Department of Integrative Biology, University of Colorado Denver, Denver, Colorado. [2]Department of Microbiology, The Ohio State University, Columbus, Ohio. [3]School of Environment and Natural Resources, The Ohio State University, Columbus, Ohio.

Temperate freshwater wetlands represent the single largest source of atmospheric methane, and yet we have a poor process-level understanding of the biogeochemical controls affecting microbial distributions or methane emissions in these ecosystems. Despite being key contributors to biogeochemical processes, including methane cycling, archaea are underrepresented in DNA-sequencing-based molecular surveys and existing genomic sampling. Using a novel domain-specific amplicon V3-V6 rRNA 16S sequencing method, we first collected a high-resolution view of archaeal diversity and geochemistry in hydric soils sampled across a model methane-emitting freshwater wetland (Old Woman Creek National Estuarine Reserve, Ohio, USA). In general, archaeal communities were reproducibly structured by soil depth and geochemistry across our replicated study design. Although methanogens with diverse metabolisms were abundant across the wetland, some showed surprising OTU-level partitioning by depth. The vast majority of the archaeal community detected is not known to produce methane. *Candidatus* Methanoperedens spp. archaea thought to perform

anaerobic oxidation of methane linked to iron reduction were abundant, as were many OTUs within the underexplored Woesearchaeota and Bathyarchaeota.

To further explore the genomic determinants underlying observed habitat preferences, we selected samples enriched in archaeal populations of *Candidatus* Methanoperedens spp. and Bathyarchaeota species for metagenomic and metatranscriptomic sequencing. We performed targeted reassembly and coassembly of binned population genomes of interest, and refined bins leveraging differential coverage that included metagenomic sequencing of additional samples across the wetland. Resulting population genome bins included the targeted *Candidatus* Methanoperedens spp. and Bathyarchaeota, but also many other archaeal taxa, including Thermoplasmatales, Methanomicrobiales and other enigmatic archaea. Many genome bins come from archaeal taxa with little previous genomic sampling, and some taxa are observed for the first time in freshwater settings. The distribution of these populations across the wetland agrees with the distributions we observed in the marker-gene survey, providing the opportunity to link genomic content to the spatial and geochemical habitat preferences we inferred from the high-resolution amplicon survey. Metatranscriptomic sequencing revealed that in the mudflat site dominated by *Candidatus* Methanoperedens spp., these archaea are not only abundant but are active, with multiple transcribed genes in the (reverse) methanogenesis pathway providing *in situ* evidence of methane-cycling activity for this potentially globally important methane consumer.

The Effect of Drought and Host Genotype on the Grass Root Microbiome

Naylor, Daniel T.* [1, 2]; Coleman-Derr, Devin (colemanderr@gmail.com)[1, 2]; DeGraaf, Stephanie[3]; Purdom, Elizabeth[3]

[1]Department of Plant and Microbial Biology, 111 Koshland Hall, University of California, Berkeley, CA 94720, USA. [2]Plant Gene Expression Center, USDA-ARS, Albany, CA 94710, USA. [3]Department of Statistics, University of California, Berkeley, CA 94720, USA.

Plant health is intimately intertwined with the communities of microbes that live on and within their roots. However, the processes that determine root microbiomes are poorly understood. Community composition might be expected to vary by a number of factors, including proximity to the root, water availability, and plant host genotype. We hypothesized that microbiome dissimilarity would positively correlate with host phylogenetic distance, and furthermore that imposing drought would provoke species-specific drought stress responses in plants that would strengthen this correlation. Furthermore, we hypothesized that drought enrichment of bacterial class Actinobacteria, seen previously in soils and plant rhizosphere in a limited number of species, would also be observed in the root endosphere in many different species. To test these hypotheses, we examined root microbial communities within bulk soil, rhizosphere, and root endosphere samples, for 18 distinct Poaceae grass lineages for both drought and well-watered conditions. Host phylogenetic distances based on chloroplast sequences correlated with microbiome dissimilarity in most root and rhizosphere samples, though the effect of this correlation was less in drought samples than control. We also found a significant enrichment of certain bacteria lineages by watering regime, most notably a universal enrichment of Actinobacteria in drought that was enhanced in roots compared to soil. Our results suggest evolutionary history might be a predictor of root microbiome composition, and that drought provokes common responses in the root microbiome that might serve as a basis for microbial-based soil amendment strategies for drought tolerance in crops.

Elucidating the Ecology of *Fibrobacter* spp. in the Herbivore Gut Using Comparative Genomics

Neumann, Anthony P.* (apneumann2@wisc.edu); Suen, Garret

University of Wisconsin-Madison, Madison, WI.

Herbivores rely on symbioses with fiber-degrading microbes living in their gut because these animals alone cannot efficiently digest the plant cell wall polysaccharides that make up their diet. One example is bacteria in the genus *Fibrobacter*, which includes important cellulose-degraders from the rumen. However, other members of this genus are poorly understood and are known to be conserved members of the gut microbiota of a diversity of herbivores. Culture-independent studies indicate that *Fibrobacter* populations in hindgut-fermenting herbivores (e.g. horses) are phylogenetically distinct from those populations in the rumen, but our ability to investigate their physiology has been hampered by a lack of representative axenic cultures. To address this knowledge gap, we developed a novel method for recovering axenic *Fibrobacter* cultures from herbivore gastrointestinal samples, and applied it toward the isolation of 45 novel *Fibrobacter* strains from 11 different hosts. These 45 strains represent 9 different phylotypes, and include the first confirmed representatives of phylogenetically distinct *Fibrobacter* populations common in the horse hindgut. Based on our phylogenetic analysis, we hypothesized that these *Fibrobacter* phylotypes occupy distinct ecological niches *in vivo*. In conjunction with the JGI, we tested this by performing whole genome sequencing on a subset of our cultured isolates using an Illumina HiSeq 2500 sequencer. In total, 23 high-quality draft genomes were produced, and an analysis of their predicted proteins identified potential functional differences in both carbon and nitrogen metabolism that corresponded to their phylogenetic placement. In particular, genes encoding proteins enabling the ability to transport and use urea as a nitrogen source were predicted for all 6 *Fibrobacter* strains representing phylotypes associated with the horse hindgut, but were only found in 1 of the 13 strains representing phylotypes associated with the rumen. Furthermore, we identified a gene encoding for a pyruvate kinase in the 6 horse-associated strains, but no gene encoding for a pyruvate-formate lyase. Importantly, pyruvate kinase is not present in the genome of the type strain *F. succinogenes* S85 from the rumen, nor was it found in any of the other 13 rumen-associated genomes examined here. This observation could explain the lack formate detected for the horse-associated strains, while rumen-associated strains typically produced small amounts of this fermentation product. The availability of these novel *Fibrobacter* strains and their genomic sequences promises to greatly expand on our knowledge of these enigmatic, yet important, fiber-degrading gut bacteria.

Agricultural Nitrogen Management Affects Functional Genes and Organisms in Nitrogen Cycle Processes

Norton, Jeanette M.* (jeanette.norton@usu.edu)[1]; Ouyang, Yang[2]; Habtesselassie, Mussie Y.[3]; Shapiro, Nicole[4]

[1]Utah State University, Logan UT, USA. [2]Michigan State University,East Lansing, MI, USA. [3]University of Georgia, Tifton GA, USA. [4]DOE Joint Genome Institute, Walnut Creek CA, USA.

Improved understanding of nitrogen (N) cycling in agroecosystems is essential for increasing N use efficiency and sustainable food production. Availability of N from organic sources and fertilizers is the

result of the enzymatic processes that comprise N mineralization, immobilization and nitrification. These transformations between organic N and inorganic N form a central part of the internal soil nitrogen cycle. Nitrification is mediated by ammonia oxidizing bacteria or archaea, and nitrite oxidizing bacteria. Understanding links between process rates, enzyme activities and the communities of microbes that cycle nitrogen may contribute to sustainable management. A multi-year experiment was conducted in Utah and Georgia to examine N-source effects on nitrification and mineralization in these agricultural systems. N-sources include low and high levels of ammonium sulfate fertilizer (100 and 200 kg N/ha) and steer or poultry manure composts. We used a combination of enzyme activity, real-time quantitative PCR for target genes, amplicon pyrosequencing and metagenomic sequencing to examine functional changes in responsible organisms. Key enzymes and their relevant marker genes included ammonia monooxygenase (amoA), nitrite oxidoreductase (nxrB), protease (npr and sub), chitinase (chiA), and urease (ureC). The overall soil microbial community composition was assessed targeting ribosomal genes. Differential inhibition and quantitative PCR revealed that while ammonia-oxidizing archaea (AOA) gene counts were higher, ammonia-oxidizing bacterial populations were more dynamic and responsible for an equal or greater fraction of the ammonium oxidized. Nitrite oxidation rates and nitrite oxidizer populations were strongly stimulated by N fertilizers. *Nitrospira* was the only known nitrite oxidizer genus identified from the Utah soils. The question of whether the nitrite oxidizers of the *Nitrospira* had the gene inventory to also perform ammonia oxidation using ammonia monooxygenase (i.e. comammox process) was investigated with metagenomics. Preliminary results were intriguing but inconclusive. The combination of enzyme kinetic and metagenomic approaches has brought us closer to the goal of linking the capable organisms to the process rate and extent in the soil environment.

Support from DOE-JGI Community Sequencing/Science Program 2010, 2014, Utah Agric. Expt. Sta. #UTA00371, 01260, USDA Agriculture and Food Research Initiative Competitive Grants Program Grant no. 2011-67019-30178 and 2016-35100-25091.

Dimethylsulfoniopropionate Degradation in the Coastal Ocean: Gene- and Taxon-Centric Approaches

Nowinski, Brent C.* (nowinski@uga.edu)[1]; Preston, Christina[2]; Kiene, Ronald P.[3]; Scholin, Christopher[2]; Birch, James[2]; Whitman, William B.[4]; Moran, Mary Ann[1]

[1]Department of Marine Sciences, University of Georgia, Athens, GA. [2]Monterey Bay Aquarium Research Institute, Moss Landing, CA. [3]Dauphin Island Sea Lab, University of South Alabama, Dauphin Island, AL. [4]Department of Microbiology, University of Georgia, Athens, GA.

Dimethylsulfoniopropionate (DMSP) is an abundant organic sulfur and carbon compound produced by marine phytoplankton. Marine bacteria transform DMSP via two pathways: the demethylation pathway retains DMSP-derived sulfur in the cell where it can be incorporated into biomass or oxidized, while the cleavage pathway releases volatile DMS with potential implications for cloud formation. The prominent hypothesis explaining differential regulation of these pathways poses that demethylation is favored when DMSP dominates the organic sulfur pool, while cleavage is favored when other organic sulfur compounds can substitute as the cellular sulfur source. Marine phytoplankton groups differ in their production and release of DMSP and other organic sulfur compounds, so shifts in phytoplankton community composition alter sulfur source availability. We tracked bacterial DMSP gene abundance in Monterey Bay surface waters during a 21-day study in which the Environmental Sample Processor (ESP) autonomously filtered and archived the seawater microbial community.

Sequencing of twelve metagenomes during the ESP deployment showed taxa known to carry DMSP genes made up 23-39% of the bacterial community, with SAR11-like cells representing the most abundant DMSP degraders, followed by Roseobacter-, SAR116-, and marine gamma proteobacterium HTCC2080-like cells. Placement of metagenomic reads to reference, assembled, and single-cell genomes revealed nine major clades of DMSP demethylation (dmdA) genes, with SAR11-like cells the source of 53%. Roseobacter-like cells were the source of most of the DMSP cleavage genes, averaging 63% of dddP, dddK, dddQ, and dddD genes.

Over the ESP deployment, chlorophyll a measurements remained relatively low for this coastal system (~1-7 ug L-1). The abundance of picoeukaryotes and the diatom *Pseudo-nitzschia* was elevated at the end of the deployment, co-occurring with low DMSP concentrations and consumption rates. However, dynamics in the phytoplankton community were not associated with changes between the four major groups of the bacterial DMSP degrading community. DMSP degraders compose a large fraction of the bacterial community in this coastal ecosystem, and changes in the relative abundance of these cells may not be closely connected to the phytoplankton community in the absence of blooms of DMSP-producing phytoplankton.

Re-Evaluating the Salty Divide: A Meta-Analysis of 16S rRNA Gene Sequences from Marine and Freshwater Systems

Paver, Sara F.* (sara.paver@gmail.com)[1]; Muratore, Daniel J.[1]; Newton, Ryan J.[2]; Coleman, Maureen[1]

[1]University of Chicago, Chicago, Il. [2]University of Wisconsin-Milwaukee, Miwaukee, WI.

Marine and freshwater communities are phylogenetically distinct and transitions between habitat types are thought to be infrequent. We revisited the concept of transition frequency and identified specific microbial lineages exhibiting comparatively high or low similarity between marine and freshwater ecosystems using a sequence-based meta-analysis of 16S rRNA amplicon datasets from 15 marine and 22 freshwater sites. Marine and freshwater microbial communities differed in the relative abundance of major phyla and proteobacterial classes and contained habitat-specific lineages while simultaneously including shared taxa observed in both environments. Alphaproteobacteria contained sequences with the highest similarity between marine and freshwater samples and the most shared Minimum Entropy Decomposition nodes, including non-LD12 SAR11 nodes observed near detection limits in the Laurentian Great Lakes. Detection limitation and observed undersampling of shared nodes suggest that transition frequency between habitat types is potentially underestimated and beyond the scope of available datasets. Identified lineages with a high degree of shared taxa or habitat-specific diversification are potential targets for future investigations into microbial adaptations and evolution. Near- or below-detection diversity, including lineages that appear to have crossed the salty divide relatively recently, may provide insurance of microbial biodiversity and have opportunities for niche expansion as environmental conditions change.

Understanding the Genomic Basis of Syntrophic Relationships between Rumen Anaerobes

Peng, Xuefeng (Nick)* (xpeng@engineering.ucsb.edu)[1, 2]; Gilmore, Sean P.[1]; Henske, John K.[1]; Swift, Candice L.[1]; Theodorou, Michael K.[3]; Valentine, David L.[2]; O'Malley, Michelle A.[1]

[1]Department of Chemical Engineering, University of California, Santa Barbara, CA, USA. [2]Department of Earth Science, University of California, Santa Barbara, CA, USA. [3]Animal Production, Welfare and Veterinary Sciences, Harper Adams University, Newport, Shropshire, UK.

Anaerobic microbial consortia in rumens of herbivore are highly effective in converting lignocellulose-rich biomass into monomeric sugars, fatty acids, and biogas (mainly carbon dioxide and methane). They consist of bacteria, protozoans, fungi, and methanogens, whose interdependency remains elusive largely due to difficulties in cultivation. We developed a cultivation method that allows us to investigate competitive and mutualistic interactions among these different groups of microorganisms. In particular, we seek to determine the specificity of the syntrophy between fungi and methanogens, and whether the composition of biomass substrates exerts selective pressure on anaerobic consortia.

We present preliminary results from the cultivation experiments conducted in 2016. Freshly produced fecal pellets from goat at the Santa Barbara Zoo were collected as source material to inoculate rumen fluid-based media with three different types of plant material and xylan as carbon substrates. One group of culture media were amended with chloramphenicol, and a consortia of anaerobic fungi dominated the enrichment culture, producing high levels of hydrogen. A second group of culture media were amended with penicillin and streptomycin, and a consortia of anaerobic fungi and methanogen dominated the enrichment culture, producing high levels of methane and there was no hydrogen accumulation. The third group of culture media were not amended with any antibiotics, and were dominated by a consortia of cellulolytic bacteria and methanogens, producing high levels of methane. These enrichment cultures were maintained for a month by passaging every three days. In addition to gaseous products, liquid media were sampled daily for analysis of primary aqueous metabolites on HPLC. At the end of each three-day passage, all remaining biomass including plant material were harvested and stored for nucleic acids extractions.

Genomic DNA and RNA were extracted using the QIAGEN AllPrep DNA/RNA/miRNA Universal kit with an initial bead beating step to achieve effective cell lysis. Genomic DNA samples were submitted for ribosomal iTag sequencing using 16S, 18S, and ITS2 primer sets, and a subset of them will be selected for metagenome minimal draft and metatranscriptome assembly. Preliminary results showed that different carbon substrates selected for different taxa of bacteria in the media untreated with antibiotics after 27 days of culturing. Bacteroidetes dominated cultures supplemented with alfalfa stems (pectin-rich). Clostridia dominated cultures supplemented with bagasse (lignin-rich). Lactobacillales and Clostridia dominated cultures supplemented with reed canary grass. Selenomonadales and Lactobacillales dominated cultures supplemented with xylan. In media amended with penicillin and streptomycin, Methanobrevibacter was the primary methanogen present in the co-culture with anaerobic fungi, and Erysipelotrichales were also a major part of the enrichment community. With the complete dataset, we will test the hypothesis that co-culturing with methanogens reprograms fungal metabolism at the transcription level through upregulation of specific biomass-degrading regulons, secretion of fungal cellulosomes, and/or sugar-specific membrane transporters.

Long term, understanding the genomic basis of anaerobic consortia from rumens will provide a framework to build synthetic co-cultures that convert biomass into value-added products.

Protecting Photosynthesis During Desiccation: Do the Genomes of Desert Derived and Aquatic Scenedesmus Species Hold the Key to Understanding Extreme Desiccation Tolerance among Green Algae?

Peredo, Elena L.* (elperedo@mbl.edu); Cardon, Zoe G.

The Ecosystems Center, Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA.

Desert-evolved green algae are particularly well suited system for exploration of the genomic basis of the adaptations that emerged in plants during the colonization of land. The specialized physiologies of desert algae, not found in their aquatic counterparts, evolved in multiple lineages of green algae (in up to seven classes). Independently evolved terrestrial green plant lineages display convergent habitat adaptations, including high tolerance to desiccation, that are achieved, at least in part, through ancestral metabolic pathways shared across the plant clade.

Water loss is especially harmful for photosynthetic organisms and often results in extensive photodamage, as light continues to be absorbed by the photosynthetic apparatus even when carbon fixation is limited by suboptimal conditions. During desiccation, a very powerful photoprotective mechanism is induced in desert-evolved algae, and rapidly reversed upon rehydration. Our work identifies commonalities in photoprotection across desert-evolved taxa in multiple algal classes. These commonalities may be the signature of an ancestral origin deeply rooted in the plant clade. Further support is provided by the shared photosynthetic traits between drought tolerant mosses and desert green algae. Our analysis of the expression patterns during desiccation and upon rehydration in desert and aquatic species of *Scenedesmus* sp. (green algae) found parallels to shifts in gene expression during seed maturation, and during desiccation of vegetative tissues in resurrection seed plants. Water loss triggers upregulation of genes dealing with water management, stress and photoprotection such as aquaporins, lipid and pigment synthesis, oxidoreductases, Late Embryogenesis Abundant proteins, Early Light Induced Proteins, and DNA repair. Our results further suggest that, despite differences in specialized life cycles or morphologies, desiccation tolerance in the green plant clade is achieved by co-opting and rewiring existing gene pathways of ancestral origin.

All this work to date has been conducted with RNASeq without genomes. To characterize the genomic basis of desiccation tolerance in green plants, we will work with JGI to sequence a suite of closely-related habitat-adapted *Scenedesmus* species (Sphaeropleales, Chlorophyta). Desert-derived *S. deserticola* and aquatic *S. obliquus* (Utex 72) differ by only 5 nt in the 18S sequence, yet S. deserticola is single-celled, thick-walled, and extremely desiccation tolerant while *S. obliquus* (Utex 72) is coenobial (colonial), thin-walled, and completely desiccation intolerant. Another desert inhabitant, *S. rotundus*, is distinguished by only 23 nt from aquatic *S. costatus* (CCAP 276-31) yet *S. rotundus* is desiccation tolerant when dried slowly, and aquatic *S. costatus* shows slight tolerance to single episodes of slow desiccation. The overall genetic background of all four species should be very similar, but comparative analysis should uncover habitat-driven genomic changes, leading to the range of desiccation tolerance we observe no matter the common garden conditions in which the algae are grown.

DOE JGI Plant Flagship Gene Atlas: Physcomitrella patens

Perroud, Pierre-François*[1]; Haas, Fabian B. (fabian.haas@biologie.uni-marburg.de)[1]; Schmutz, Jeremy[2, 3]; Grimwood, Jane[2]; Stacey, Gary[4]; Sreedasyam, Avinash[2]; Rensing, Stefan A.[1]

[1]Plant Cell Biology, Faculty of Biology, University of Marburg, Germany. [2]HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA. [3]DOE Joint Genome Institute, Walnut Creek, CA, USA. [4]Divisions of Plant Science and Biochemistry, National Center for Soybean Biotechnology, University of Missouri, Columbia, MO, USA.

In the context of the Department of Energy's Joint Genome Institute (DOE JGI) Plant Flagship Gene Atlas project, we present here the dataset generated for the model moss *Physcomitrella patens*. Our in-house RNA-seq data management pipeline was used to map and associate sequence reads with differently expressed genes (DEGs). Subsequently, stage and condition specific gene analysis as well as transcriptome-wide Gene Ontology (GO) term bias analyses were performed. Clear separation of the RNA-seq experiments and clustering of the biological replicates were found.

The three dominant morphological stages, protonemata, gametophore and sporophyte can be separated clearly at transcriptome level using a GO term bias approach. Additionally, comparison across all replicates of treated and untreated stages allowed us to define stage specific transcript sets, present only in one type of tissue and absent in all other regardless of the treatment applied. In parallel stage specific sampling and treatment were applied to a given developmental stage. For example protonema tissue at 24 h of ABA was correlated with control. Also, effect of ammonium was evaluated by comparing culture growing with NH3 or nitrogen and NH_4. Diverse and specific expression patterns emerge from the hormonal and metabolic treatments.

The homogeneity of protonemata makes this tissue, in principle, the simplest to experimentally manage and should be the easiest to compare within and between laboratories. Comparison of experiments confirmed the tight correlation between replicates within a single laboratory. Yet, replication of the same conditions between laboratories underscored the difficulty to perform such comparison. Even with medium, growth conditions and tissue identically sets, different laboratories generated significantly different transcriptome profiles, showing that seemingly minimal variations during the experiment leads to variation difficult to explain and control.

For cross-experiment normalization the External RNA Controls Consortium (ERCC) spike-ins were added to all experiments. Spike-ins are polyadenylated artificial transcripts that mimic natural eukaryotic mRNA. Normalization is difficult for any expression profiling methodology and not well developed for RNA-seq data yet. A methodology to make use of ERCC spike-ins for cross-experiment normalization is under development. So far, DEG calling with the use of spike-in as a normalization is significantly different from the DEG calling with in silico normalization methods. Additionally, unexpected fluctuations of spike-in sequences between experiment RNAs were found.

All presented experiment results are based on the first set of 99 samples. Another 72 samples have been processed and will significantly broaden the range of conditions and perturbations.

Use of a Single Cell Type Model, the Root Hair Cell, to Advance Our Understanding of the Soybean and Sorghum Transcriptomic and Epigenomic Responses to Various Environmental Stresses

Pingault, Lise M.* (lise.pingault@ou.edu); Qiao, Zhenzhen; Ren, Yuqi; Libault, Marc

University of Oklahoma, Norman, OK.

The structural organization of eukaryotic genomes depends on various epigenomic marks including, among others, the methylation pattern of the genomic DNA (gDNA) and the pool of small non coding RNAs. These epigenomic marks also regulate gene transcription, gDNA replication, recombination and repair. However, mostly due to the cellular complexity of the tissues collected, modifications of the plant epigenome in response to environmental stresses and their impact on gene expression remain unclear. To better characterize the role of the epigenome in controlling gene expression, we are focusing our effort on one single plant cell type: the root hair (RH) cell. The main function of these single tubular root cells is to improve water and nutrient uptake by the roots. As a consequence, RHs are highly responsive to environmental stresses, and they are considered as a model to investigate in detail the molecular response of a single cell type to various biotic and abiotic stresses including those related with climate change.

Using soybean as a model and taking advantage of the recent updates of the annotation of its genome, we initiated the project by comparing the modifications of the soybean RH methylome and transcriptome in response to 11 environmental stresses including extreme temperatures, pH variations, nutrient deprivations, elevated atmospheric CO_2 concentration and salinity.

Integrating the transcriptomic datasets generated from this project to the previously published soybean RH transcriptome (Libault et al., 2010), we updated to 726 the number of soybean genes specifically expressed in RHs. Considering the entire soybean genome and the soybean RH-specific genes, we are analyzing our transcriptomic and epigenomic datasets to (1) draw expression profiles of each soybean gene in response to environmental stress, (2) identify new non-coding transcripts, (3) study in details the transcriptional evolution between soybean paralogs, (4) and identify differentially methylated regions of the soybean genome and their impact on gene expression. The detailed results of our analysis will be presented as well as the relationships existing between gene space organization and genome methylation under the context of plant response to different environmental stresses.

The Nuclear Genome of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta), a Commercially Harvested Representative of the Ancient Bangiophytes

Prochnik, Simon E.*[1]; Brawley, Susan H. (brawley@maine.edu); Schmutz, Jeremy[2]; Jenkins, Jerry[2]; Barry, Kerrie[1]; Lindquist, Erika[1]; Rokhsar, Daniel[1]; Blouin, Nicolas A[3]; Bhattacharya, Debashish[4]; Grossman, Arthur R.[5]; Stiller, John E.[6]

[1]LBNL Joint Genome Institute, Walnut Creek, CA. [2]HudsonAlpha Institute for Biotechnology, Huntsville, AL. [3]University of Wyoming, Laramie, WY. [4]Rutgers University, New Brunswick, NJ. [5]Carnegie Institution for Science, Stanford, CA. [6]East Carolina University, Greenville, NC. [7]University of Maine, Orono, ME.

Biogeochemical cycling, aquatic food webs, and human health are strongly influenced by plastid and nuclear genes from red algae because of secondary endosymbioses between heterotrophs

and captured red algae that led to the evolution of diatoms, dinoflagellates, apicomplexans, and haptophytes. Moreover, the oldest taxonomically-resolved fossil (1.2 Ga) of a multicellular eukaryote belongs to the class Bangiophyceae; however, few red algal nuclear genomes have been sequenced. Here, we discuss features of the haploid nuclear genome (87.7 Mbp) of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta). One of the notable features of the genome is its high G + C-content; protein-coding regions have a G + C-content averaging 72.9 %. Prior to use of PacBio, the genome sequence was incomplete, and sequencing reactions disproportionately recovered antibiotic-resistant bacterial contaminants with moderate G + C-content. An improved and highly contiguous (contig L50 = 189.9 kb) draft assembly was generated with whole genome shotgun (WGS) sequencing using the PacBio platform, with homozygous insertions and deletions corrected using Illumina WGS. Nearly 98% of the sequenced transcripts (ESTs) can be mapped to the genome assembly and a complete complement of genes encoding RNA polymerase subunits and other conserved informational proteins involved in transcription, translation, and DNA synthesis are present, suggesting that the genome is nearly complete. A typical gene has ~2 exons, implying abundant splicing for a red alga; however, only 235 alternative splice-forms were identified from expression data. This genome encodes the largest number of gene loci (13,125) among the characterized red algal nuclear genomes and a notable portion (~44%) of repeat sequences (largely LTR and DNA transposable elements). Despite the preparation of cultured, clonal, unialgal blades from medium containing antibiotics to eliminate associated bacteria, and purification of genomic DNA on cesium chloride gradients, approximately half of the scaffolds in the PacBio assembly did not belong to *Porphyra*. Existing metagenomic binning tools were unable to accurately partition the *Porphyra* genome scaffolds, therefore a PCA-based method was developed. This allowed us to confidently separate scaffolds belonging to the *Porphyra* genome from the metagenomic scaffolds and to generate a high quality assembly.

Invasive Dreissenid Mussels Induce Shifts in Bacterioplankton Diversity through Selective Feeding on High Nucleic Acid Bacteria

Props, Ruben* (Ruben.Props@UGent.be)[1, 2]; Schmidt, Marian[2]; Heyse, Jasmine[1]; Vanderploeg, Henry[3]; Boon, Nico[1]; Denef, Vincent[2]

[1]Center for Microbial Ecology and Technology (CMET), Ghent University, Coupure Links 653, B-9000 Gent, Belgium. [2]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, US. [3]NOAA Great Lakes Environmental Research Laboratory, Ann Arbor, MI, US.

Species invasion is an important disturbance to ecosystems worldwide, yet knowledge about the impacts of invasive species on bacterial communities remains sparse. Using a novel approach, we simultaneously detected phenotypic and derived taxonomic change in a natural bacterioplankton community when subjected to feeding pressure by invasive dreissenid mussels (IDMs), a widespread aquatic invasive species. We detected a significant decrease in diversity within one hour of feeding, and a total diversity loss of 11.6 ± 4.1 % after 3h. This loss of bacterial diversity was caused by the selective removal of high nucleic acid (HNA) populations (29 ± 5% after 3h). We were able to track the community diversity at high temporal resolution by calculating phenotypic diversity estimates from flow cytometry data of minute amounts of sample. Through parallel flow cytometry and 16S rRNA gene amplicon sequencing analysis of environments spanning a broad diversity range, we showed that the two approaches resulted in highly correlated diversity measures and captured the same seasonal and lake-specific patterns in community composition. This allowed us to predict the magnitude of community shifts solely based on flow cytometric measurements, facilitating a temporal

resolution of this biological process that would not have been possible otherwise. Based on our results, we predict that selective feeding by IDMs directly impacts the microbial component of the carbon cycle, as it may drive bacterioplankton communities toward less diverse and less productive states.

Metagenomics of Methane-Oxidizing Mesocosms from the Gulf of Mexico and U.S. Atlantic Margin

Redmond, Molly C.* (mcredmond@uncc.edu)[1]; Chan, Eric W.[2]; Shiller, Alan M.[2, 3]; Kessler, John D.[2]

[1]University of North Carolina Charlotte, Charlotte, NC. [2]University of Rochester, Rochester, NY. [3]University of Southern Mississippi, Stennis Space Center, MS.

Microbial methane oxidation at marine hydrocarbon seeps significantly reduces the amount of methane escaping to the atmosphere. However, questions remain about the environmental factors that control methane oxidation rates and organisms responsible. To address these questions, we designed methane oxidizing mesocosms with seawater collected from methane seeps on the U.S. Atlantic Margin (Hudson Canyon) and the Gulf of Mexico. We measured aerobic methane oxidation with high resolution measurements of methane, carbon dioxide, and oxygen concentrations, stable isotopic changes in methane and carbon dioxide, trace metals, nutrients, and 16S rRNA gene and metagenomic sequencing in ten replicate mesocosms from each site. Hudson Canyon seep mesocosm communities were dominated by methanotrophs from the family Methylococcaceae. Methylococcaceae were also present in the Gulf of Mexico mesocosms, but were much less abundant and methane was consumed less rapidly than in the Hudson Canyon mesocosms. The Gulf of Mexico seeps also emit ethane, propane, and other hydrocarbons, and hydrocarbon degraders such as Colwellia and Cycloclasticus were abundant in the Gulf of Mexico mesocosms. Metagenomes from both locations included genes encoding methane monooxygenase and methanol dehydrogenase (both mxaF and xoxF types). The Gulf of Mexico metagenomes also contained genes involved in the oxidation of longer chain alkanes and aromatic compounds, including several divergent methane monooxygenase genes that may be involved in ethane or propane oxidation.

Plant-Microbe Interfaces: Characterizing the Diversity and Function of the Ectomycorrhizome of *Populus trichocarpa*

Rojas, Alejandro J.* (jar115@duke.edu)[1]; Looney, Brian[2, 3]; Liao, Hui-Ling (Sunny)[1]; Chen, Jay[2]; Labbé, Jessy[2]; Vilgalys, Rytas[1]

[1]Duke University, Durham, N.C [2]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN. [3]University of Tennessee, Knoxville, TN.

Project Goals: The goal of the PMI SFA is to understand the genome-dependent molecular and cellular events involved in establishing and maintaining beneficial interactions between plants and microbes. *Populus* and its associated microbial community serves as the experimental system for understanding how these molecular events manifest themselves within the spatially, structurally, and temporally complex scales of natural systems. To achieve this goal, we focus on 1) characterizing host

and environmental drivers for diversity and function in the *Populus* microbiome, 2) utilizing microbial model system studies to elucidate *Populus*-microbial interactions at the molecular level and dissecting the signals and pathways responsible for initiating and maintaining microbial relationships and 3) develop metabolic and genomic modeling of these interactions to aid in interpreting the molecular mechanisms shaping the *Populus*-microbial interface.

The *Populus* root microbiome harbors a diverse community of ectomycorrhizal fungi (EMF) that significantly increases nutrient uptake and acquisition by the plant host while also providing protection against antagonistic parasites. Over 30 genera of EMF are known to associate with *Populus*, including many groups of mushroom-producing families including Boletaceae, Russulaceae, Cortinariaceae, Tricholomataceae, and Amanitaceae. A major aim of the PMI project is the collection, isolation, and characterization of the major EMF fungal associates of *P. trichocarpa* across its range in the Pacific Northwest. In Autumn 2015 and 2016, we surveyed macrofungal diversity under native *P. trichocarpa* forests from five core watersheds on the Squamish (BC), Snohomish (WA), Puyallup (WA), Columbia (OR and WA), and Willamette (OR) rivers. This resulted in over 150 collections of EMF fruit body collections and sampling of bulk soil used in bioassay studies. All sporocarp collections were plated on modified Melin- Norkrans medium, photographed, spore printed, and dried for identification and accession into fungal herbaria. The first fungal collections have been identified while the second samples from the second campaign are being identified by a consortium of taxonomic experts using the ITS barcode marker as well as morphological features.

The Plant Microbe Interfaces Scientific Focus Area is sponsored by the Genomic Science Program, U.S. Department of Energy, Office of Science, Biological and Environmental Research. http://pmi.ornl.gov

Extremophile Microbial Bioprospecting for Biomass Production and Biorefinery

Ruiz-Font, Angelica C.* (convenios22@hotmail.com)[1, 2, 3]; Melo-Perez, Flor M.[2]; Betto-Sagahon, Janeli[2]; Campos-Vazquez, Maria E.[2]

[1]Applied Biotechnology Research Center, National Polytechnic Institute. Ex Hacienda San Juan Molino, Tlaxcala, Mexico. 90700. [2]Productive Biotechnology Graduate Program at National Polytechnic Institute. CIBA-IPN. Mexico. 90700. [3]Doctorado en Ciencias Naturales para el Desarrollo, Universidad Nacional de Costa Rica.

Global threats of oil shortages in the near future and climate change due to green-house gas emissions are posing serious challenges and hence and it is imperative to explore means for sustainable ways of averting the consequences. The dual application of forest-agave production for soil restoration and biomass production for sustainable biorefinery production is a feasible option. This project aims to evaluate and select endophytic microorganisms from extreme environments (arid land) that promote plant growth and promote the restoration of deteriorated soils. And later in environments with high acidity, we have selected microorganisms with capacity to produce hydrocarbon precursors.

In the first stage, bioprospecting was developed in soil seed banks of arid zone plants. Plant associated microbes may help mediate such dry and salt stress. We analyzed rhizospheric, soil and leaf litter microbial communities associated with two saline-adapted chenopod plants, Suaeda mexicana, from central Mexico and Atriplex canescens, from the Chihuahuan Desert region of the United States. In order to characterize the cultivable microbial community, samples were processed and analyzed by traditional surface spread plating methods on sixteen different culture media, each contained either

4% or 10% NaCl (w/v). 43 morphotypes were selected, run them with PCR, amplification of 16S rDNA was carried out using the primer pairs F984GC-R1378 for bacteria and ITS1F-ITS4 for the sole fungal isolate. PCR products from Atriplex isolates were homologous to sequences of the bacterial genera *Penibacillus, Streptomyces, Promicromonospora, Rhodococcus, Bacillus* and *Pseudomonas,* and the fungal genus *Aspergillus*. Sequences homologous to the genera *Arthobacter, Streptomyces, Nocardia, Cellulosimicrobium, Pseudomonas* and *Bacillus* were amplified from Suaeda isolates. Each morphotype were used for the inoculation of axenically grown Amaranth spp seedlings and plant growth promotion was evaluated.

Enhanced Productivity of Biofuel-Producing Marine Microalgae Using Ecologically Engineered Bacteria—Examination Using Isotope Tracing and Marker Gene Sequencing

Samo, Ty* (samo1@llnl.gov)[1]; Kimbrel, Jeffrey[1]; Nilson, Daniel J.[1]; Siccardi, Anthony[3]; Lane, Todd[2]; Pett-Ridge, Jennifer[1]; Weber, Peter K.[1]; Mayali, Xavier[1]

[1]Lawrence Livermore National Laboratory, Livermore, CA. [2]Sandia National Laboratories, Livermore, CA. [3]Texas A&M Agrilife, College Station, TX.

We sought to ecologically engineer mutualistic microalgal-bacterial interactions by establishing bacterial enrichments using live algal cultures as organic matter sources. After serially diluting these bacterial communities in secondary enrichments, we focused on those associated with the algal phycospheres of two model biofuel-producing microalgal strains, *Phaeodactylum tricornutum* (Pt) and *Nannochloropsis salina* (Ns). Algal growth and biomass monitored using chlorophyll and absorbance measurements revealed elevated growth and/or yield in 4.5% of Ns and 7.5% of Pt secondary enrichments. Using isotope tracing with H13CO3 and 15N-leucine additions followed by NanoSIMS imaging of single cells, 4 of the initial enrichments indicated that microalgae with attached bacteria had decreased inorganic 13C fixation while those attached bacteria had higher 15N-leucine incorporation (growth) than free-living bacteria. The 4 secondary enrichments exhibited the opposite patterns; increased C fixation by microalgae with attached bacteria (up to 66% compared to axenic microalgae) and faster growing free-living bacteria. The high bacterial diversity of primary enrichments, analyzed by 16S rRNA gene iTag sequencing, likely caused minimal productivity responses due to competing effects on algal metabolism. The simplified nature and presence of mutualistic bacteria in the secondary enrichments enabled a conferral of favorable traits by bacteria to algae, and vice versa. Indeed, the Rhodobacteraceae family, known for probiotic interactions with algae, dominated the secondary enrichments. Productivity enhancements also correlated with genera within families of the Alteromonadaceae, Cyclobacteriaceae, and unclassified members of Bacteroidetes, Gamma-, and Alphaproteobacteria phyla. One filamentous genera, *Haliscomenobacter* sp., was noted to provide nitrogenous substances to Pt in exchange for organic compounds, and may represent an important bacterial component of healthy and productive algal ponds. Overall, the growth effects observed here suggest that cell attachment and cooperative symbioses of bacteria interacting with microalgae and each other are critical for stabilizing mutualisms that may enable economically feasible and sustainable algal biofuel production.

bio-sfa.llnl.gov

Driving Change in the Bacterial Microbiome Using Growth Promoting Inoculums in *Nicotiana benthamiana* Plants

Sanchez Barrios, Andrea M.* (asanchez-barrios@uky.edu)[1]; Lundberg, Derek[3]; Shearin, Tom[1]; De Lorenzo, Laura[2]; Nair, Meera[1]; DeBolt, Seth[1]

[1]Department of Horticulture and Plant Physiology/Biochemestry/Molecular Biology Program, University of Kentucky Lexington, Kentucky 40546, USA. [2]Department of Plant and Soil Sciences, University of Kentucky, Lexington, KY, 40546-0312, USA. [3]Department of Molecular Biology, Max Planck Institute for Developmental Biology, Tübingen 72076, Germany.

Microbial communities in soil are subjected to constant changes driven by the environment. Same way, plant recruitment of microbes relays on any alteration introduced to the system in which they are growing. For this project, our overall hypothesis was that the core microbiome of a plant will be flexible as a response to the introduction of foreign bacteria used as growth promoters. Plant-microbe interactions have been a subject of interest as a means to improve the health and development of the plant. Application of microbes as inoculants for growth promotion has been analyzed, but it has mostly been directed towards root mass studies and nutrient availability. It is because of this that we identified microbial inoculums capable of influencing plant cell expansion. *Nicotiana benthamiana* seeds were treated with three different bacterial strains, which we called S41 –*Micrococcus* sp.- (isolated from switchgrass), S413 –*Bacillus* sp.- and S343 –*Bacillus cereus*- (isolated from tomato plants), as well as mock treatments. CARD-FISH revealed the presence within intact plant tissue of the bacteria through time in treated and control plants. In addition, differential plant height (H), length (L), leaf width (LW), number of leaves (NL), and number of flowers (NF) were documented to analyze plants response to the inoculum (cold treated, exudates, live cells, death cells and as a whole) when applied on the seed coat. Treatments S413 and S343 inoculants had beneficial effects on seedling development and vegetative growth of plants; in contrast, S41 had a retarded growth effect. Auxin reporter lines suggested a correlation between cellular hormone shift and influence of morphological changes due to hormonal alteration which could be linked to cell expansion. The evaluation of endophyitic bacterial communities 16S rRNA amplicons in inoculated-plants vs. controls showed striking changes in the composition of the microbial communities in roots of the different treatments. In summary, bacterial inoculants had an effect on plant development and growth, indicating their potential as a tool in crop production.

Drought Stress Alters the Structure of Microbial Communities Associated with Rice Roots

Santos-Medellin, Christian M.* (cmsantosm@ucdavis.edu); Liechty, Zachary; Edwards, Joseph A.; Nguyen, Bao; Sundaresan, Venkatesan

Dept. of Plant Biology, University of California, Davis, CA.

Plants continuously interact with diverse communities comprised of beneficial, commensal, and pathogenic microorganisms via their roots. Environmental shifts can potentially restructure this intricate network of biotic interactions by affecting the physical, chemical, or biological properties across the plant-soil continuum. Drought, in particular, can reduce soil moisture, increase oxygen availability, and limit substrate diffusion. Furthermore, it can reshape the plant-microbial crosstalk

by triggering stress responses in both the plant host and associated microorganisms. To test if this complex environmental disturbance translates into changes in the microbial composition of root communities, we set up a multi-factorial greenhouse experiment designed to explore the conservation and variability of the community response to drought across key determinants of root microbiome structure. Briefly, we grew four rice cultivars in three different soils, and exposed them to a dry-down regime for three weeks. Samples were collected from unplanted soils and two distinct root compartments: rhizosphere (the soil microenvironment immediately affected by the root), and endosphere (the root interior). Drought significantly altered the overall microbial structure in all three compartments with the endosphere communities showing the greatest divergence from well-watered communities. Moreover, significant interactions between irrigation regime and the other experimental factors were detected, indicating that the impact of drought varied across soils, and cultivars. An analysis at the highest taxonomic levels also revealed that water-deficit decreased the relative abundance of multiple bacterial taxa with Deltaproteobacteria and Acidobacteria being the most affected. In contrast, Chloroflexi and Actinobacteria were constantly overabundant in drought-stressed communities. Notably, the relative increase in Actinobacteria resulted from the simultaneous enrichment of several microbial species that spanned multiple classes, orders, and families within the phylum. This pattern was consistent regardless of the soil and cultivar used, suggesting an ecological coherence among this phylogenetically diverse group of microorganisms. Together, these results show that drought reshapes the microbial composition of root-associated microbiomes towards an Actinobacteria-enriched state.

JGI Plant Genome Re-Annotation Using NGS Sequences

Shu, Shengqiang* (sqshu@lbl.gov)[1]; Herring, Robert[2]; Jenkins, Jerry[2]; Goodstein, David[1]; Schmutz, Jeremy[2]; Rokhsar, Daniel[1]

[1]DOE Joint Genome Institute, Walnut Creek, CA. [2]HudsonAlpha Institute for Biotechnology, Huntsville, AL.

We report a new pipeline, gene model improvement (GMI), yielding high and stable gene sets. GMI uses transcriptome to update the gene structure if and only if transcriptome data dictate. With short read transcriptome assemblies and ESTs/cDNA if any, we achieved about 95% mapping ratio in several plant genomes and 99% mapping ratio using assembled corrected PacBio Iso-Seq circular consensus sequencing (CCS) reads in *Arabidopsis*. Assembles from corrected CCS reads are near full length, high quality and contribute to high mapping ratio in *Arabidopsis*.

Corrected CCS reads yield higher number of alternative transcripts comparing to PacBio Iso-Seq transcripts.

Discovery, Transfer, and Characterization of Novel Phytase Genes for Utilization of Recalcitrant Phosphate

Shulse, Christine N.* (cnshulse@lbl.gov)[1]; Chovatia, Mansi[1]; Agosto, Carolyn[1]; Wang, Gaoyan[1]; Lei, Yan[1]; Hamiton, Matthew[1]; Deutsch, Samuel[1]; Yoshikuni, Yasuo[1]; Deutschbauer, Adam[1, 2]; Blow, Matthew J.[1]

[1]DOE Joint Genome Institute, Walnut Creek, CA. [2]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA.

High yield agricultural plant growth is currently dependent on costly and environmentally damaging phosphate fertilizers. One approach to alleviating this dependency is to develop bacterial strains that convert existing phosphorus sources in the soil to soluble forms available for plant uptake. Among potential sources, phytic acid is an abundant organic phosphorus-containing compound accounting for up to 40% of soil phosphate in some agricultural soils. We therefore aimed to engineer plant-associated bacteria with the ability to hydrolyze phytic acid, and release plant-available phosphate. We first searched all available microbial genomes and environmental metagenomes in the Integrated Microbial Genomes database and selected 92 sequences that represent the diversity of phytase genes. Using JGI DNA synthesis program capabilities, we refactored these sequences for optimal expression in Proteobacteria, synthesized the genes and engineered them into the genomes of four known plant-associated bacteria. We next determined the ability of these engineered strains to solubilize phytic acid in liquid culture assays. While host strains have no or low native phytase activity, a total of 29 engineered strains were capable of high levels of phytic acid hydrolysis, with at least one representative for each of the host strains used. Finally, we tested 10 strains in plant assays and identified 3 strains from two hosts that confer a significant growth advantage on the model plant *Arabidopsis thaliana* Col-0 when phytate is the sole phosphate source. These data provides proof of principle that DNA synthesis approaches can be used to generate plant associated strains with novel capabilities benefitting plant growth, and are a first step in the development of alternative approaches to sustainable phosphorus use in agriculture.

Metagenomic study of Microbial Community in Seasonal Stratified Lakes

Sjogaard, Kamilla S.* (kamillasjogaard@biology.sdu.dk)[1]; Forth, Michael[1]; Schramm, Andreas[2]; Woyke, Tanja[3]; Treusch, Alexander H.[1]

[1]Nordic Centre for Earth Evolution (NordCEE), Department of Biology, University of Southern Denmark (SDU), Odense, Denmark. [2]Section for Microbiology and Center for Geomicrobiology, Department of Bioscience, Aarhus University, Denmark. [3]DOE Joint Genome Institute, Walnut Creek, CA, USA.

Danish lakes often have a high level of anthropogenic influence and a feature regularly observed in them is seasonal stratification with hypoxic or anoxic bottom waters during summer. This oxygen gradient creates variation in biogeochemical cycling and diversity in microbial functions.

The aim of our study was to find key microbes in the carbon cycle along an oxygen gradient in lakes with different organic carbon sources. The two Danish stratified lakes, Lake Almind (mesotrophic/oligotrophic) and Lake Hund (dystrophic) were sampled for DNA in sediment and at multiple water depths, where also nutrient and oxygen levels were measured. The microbial communities were studied with both 16S rRNA gene amplicon sequencing and shot gun sequencing for metagenomic

analysis. The most abundant OTUs were of the acI lineage (Actinobacteria) and known lake cosmopolitans of β-proteobacteria (Albidiferax, Limnohabitans, Polynucleobacter). Specifically for Lake Almind were LD12 freshwater group (α-proteobacteria) most abundant, while Flavobacterium (Bacteroidetes) and Mycobacterium (Actinobacteria) were mostly abundant in Lake Hund.

Meiosis in *Porphyra umbilicalis*

Sprecher, Brittany N.* (brittany.sprecher@uconn.edu)[1]: Lin, Senjie[1]; Prochnik, Simon E.[2]; Brawley, Susan H.[3]

[1]Dept of Marine Sciences, University of Connecticut, 1080 Shennecossett Road, Groton, CT 06340, USA. [2]DOE Joint Genome Institute, 2800 Mitchell Dr, Walnut Creek, CA 94598. [3]School of Marine Sciences, University of Maine, Orono, ME 04469.

Meiosis, an essential step for sexual reproduction, is crucial for maintaining genetic diversity and is widespread in eukaryotic lineages. Sexual organisms are often determined through the presence or absences of necessary meiotic genes in a genome. The red algae, *Porphyra umbilicalis*, that inhabits the northeastern United States is believed to be asexual. To find molecular evidence to support this notion, we analyzed the genome of this strain of *Porphyra umbilicalis* and looked for meiosis genes. Six (SPO11, HOP1, HOP2, MND1, MSH4, MSH5) of the nine (DMC1, REC8, MER3) most common eukaryotic meiosis-specific genes were found. In addition, *P. umbilicalis* appears to lack six other genes found in several eukaryotes that simultaneously encompass roles in DNA repair, mitosis, and meiosis: MLH2, MLH3, PDS5, RAD21, RAD51, and SSC3. However, the significance of these absences requires further study, as some of these genes are not present in clearly sexual organisms and it is possible that other proteins support these functions in this *P. umbilicalis* strain. For future work, we propose to further investigate the missing genes and compare *P. umbilicalis's* genetic composition to a clearly sexual organism.

JGI Plant Gene Atlas: An Updateable Transcriptome Resource for JGI 'Flagship' Plants

Sreedasyam, Avinash*[1]; Schmutz, Jeremy (jschmutz@hudsonalpha.org)[2]; Plott, Christopher[1]; Hossain, Shakhawat[2]; Qualls, Morgan[1]; Jenkins, Jerry[1]; Grimwood, Jane[1]; Carlson, Joseph W.[3]; Goodstein, David[3]; Kennedy, Megan[3]; Wang, Mei[3]; Johnson, Jenifer[3]; Sandor, Laura[3]; Amirebrahimi, Mojgan[3]; Zane, Matthew[3], Chris, Daum[3], Torres-Jerez, Ivone[4], Udvardi, Michael[4]; Jawdy, Sara S.[5]; Tang, Yuhong[4]; Tuskan, Gerald A.[5]; Juenger, Thomas E.[6]; Brutnell, Tom[7]; Merchant ,Sabeeha S.[8]; Rensing, Stefan A.[9]; Mullet, John[10]; Mockler, Todd[7], Stacey, Gary[3]; Schmutz, Jeremy[1, 2]

[1]HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA. [2]Divisions of Plant Science and Biochemistry, National Center for Soybean Biotechnology, University of Missouri, Columbia, MO, USA. [3]DOE Joint Genome Institute, Walnut Creek, CA, USA. [4]Plant Biology Division, Samuel Roberts Noble Foundation, Ardmore, OK, USA. [5]Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA. [6]University of Texas at Austin, Department of Integrative Biology, Austin, TX, USA. [7]Donald Danforth Plant Science Center, St. Louis, MO, USA. [8]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA, USA. [9]University of

Marburg, Karl-von-Frisch-Str. 8, 35043 Marburg, Germany. [10]Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX, USA.

Availability of up-to-date structural and functional annotations is crucial in gaining insights into the plant biological processes at the system and molecular levels. Analyzing the gene expression landscape of plants in different tissues and conditions can lead us towards deciphering the genetic architecture and understanding the biological role of genetic elements. To contribute to that goal, JGI Plant Gene Atlas project was undertaken. Gene Atlas Project is an updateable plant transcriptome resource consisting of 1,087 transcriptomic datasets of standard tissues and conditions generated across thirteen plants: *Chlamydomonas* (algal model), *Physcomitrella* (moss model), *Brachypodium* (a C3 grass model), switchgrass (a woody perennial crop plant), Hall's panicgrass (grass model), *Setaria italica* (grain and forage crop), *Setaria viridis* (model C4 grass), Sorghum (a C4 grass bioenergy crop and model) [6 monocots], *Arabidopsis* (model for plant genetics and biology), soybean (legume model and crop plant), *Medicago* (legume model), *Eucalyptus* and poplar (2 biomass tree crops).

Employing a genome version controlled computational pipeline, we catalogue 1) expression profiles of annotated genes in these 13 plant genomes, 2) clusters of genes exhibiting tissue/condition specific expression patterns, 3) co-expression networks and 4) transcriptional modulation observed in response to changes in nitrogen regimes. We performed interspecific comparisons that shed light on orthologous genes displaying similar expression patterns and potentially involved in similar functional roles in common tissues and conditions across the surveyed phylogeny. Based on these data sets we identified primary targets for functional characterization. We systematically assigned experimental and orthology-based functional descriptions to 17.7% to 49.6% genes in JGI plant flagship genomes and created an updateable transcriptomic data resource available through the JGI Plant Portal at phytozome.jgi.doe.gov.

Identifying and Characterizing the Secondary Metabolites of Anaerobic Fungi

Swift, Candice L.* (cswift@umail.ucsb.edu)[1]; O'Malley, Michelle A.[1]; Louie, Katherine[2]; Northen, Trent R.[2]

[1]University of California, Santa Barbara, CA. [2]DOE Joint Genome Institute, Walnut Creek, CA.

Fungal secondary metabolites have provided humans with abundant natural products with applications in medicine, biofuels, agriculture, and others. Multimodular enzymes produce secondary metabolites in a fashion similar to molecular assembly lines, which are typically clustered within genomes. Much scientific effort has been devoted to the study of biosynthetic gene clusters with the end goal of controlling the production of diverse structural products. Understanding natural biosynthetic machinery is the first step towards designing new enzymes for engineering natural products. Here, we uncovered a plethora of gene clusters encoding biosynthetic enzymes for secondary metabolites from diverse chemical classes by mining the genomes and transcriptomes of three anaerobic fungi from the primitive phylum *Neocallimastogomycota*. Key secondary metabolite clusters include those that build polyketides and nonribosomal peptides. Anaerobic fungi thrive in competitive microbial environments such as the digestive tract of many large herbivores due to an unprecedented suite of biomass degrading enzymes that degrade lignocellulosic material without pretreatment. We hypothesize that these unique non-model organisms also possess a rich array of biosynthetic machinery for producing secondary metabolites, whose function remains to

be determined in their host microbial communities. Unlike many aerobic fungi which transcribe a low percentage of clusters under standard laboratory conditions, two of the anaerobic fungal isolates studied transcribe greater than 30% of clusters during lab-scale growth on various cellulosic substrates. These clusters have low sequence similarity to biosynthetic gene clusters in other organisms with known products, suggesting that they may produce previously uncharacterized compounds. The secondary metabolism of anaerobic fungi represents a completely untapped reservoir of biosynthetic potential, which could be drawn on for therapeutics, new chemical building blocks, and enzymes for bioengineering natural products.

JGI CSP #1657: Genome Sequencing of *Arabidopsis thaliana* Powdery Mildew Pathogen *Golovinomyces orontii* Isolate MGH1

Taneja, Jyoti* (jyotitaneja@berkeley.edu); Andrew, Liu; Wildermuth, Mary

Department of Plant and Microbial Biology, University of California, Berkeley, CA.

Powdery mildews are widespread fungal pathogens that infect plants including economically important cereal crops, fruits, vegetables, ornamentals and tree species. As part of their obligate biotrophic life cycle, they develop feeding structures in the plant epidermis and alter host plant metabolism to acquire nutrients for their growth and reproduction while limiting plant defense. Little is known about the virulence functions that target host metabolism and immune system.

In the past few years, efforts have been made to sequence genomes of 5 different powdery mildews that led to the identification of some specific structural features. The key identified features include larger estimated genome sizes (from 120 to 160 Mb) than other Ascomycetes, many fewer genes, the absence of specific metabolic pathways, many more candidate effector proteins and an extensive proliferation of DNA repeat elements. Comparative analysis of the genomes of powdery mildews with those of other obligate plant biotrophs, suggests that obligate biotrophy is associated with genome size inflation and substantial irreversible gene losses.

To gain insights into the molecular mechanism of powdery mildew infection process and the evolution of obligate biotrophy, a JGI-CSP Project #1657 "Comparative Genomics of Powdery Mildews and Associated Plants" co-led by Mary Wildermuth and Shauna Somerville at UC Berkeley, aims to sequence genomes of 11 powdery mildew strains representing all five tribes of powdery mildews (Erysiphaceae). This project involves the contribution of genomic DNAs submitted by labs working on powdery mildews across different continents of the world.

Wildermuth Lab works with powdery mildew *Golovinomyces orontii* that infects *Arabidopsis* and is a well-established model for studying obligate biotrophy in dicots. Methodologies for spore collection, high molecular weight DNA isolation, quality control and contamination assessment were developed/optimized. The *G. orontii* genome was sequenced by the JGI using PacBio20Kb+ library and was assembled using FALCON assembly tool. The initial post assembly BLAST analysis returned significant hits to *Albugo laibachii*, an oomycete pathogen of *A. thaliana*, which suggested contamination with *A. laibachii*. However, further analysis determined this was not the case. Instead, the *A. laibachii* genome sequence contained powdery mildew sequence, which was reported to be a contaminant.

Currently, the automated annotation of *G. orontii* genome is underway by Igor Grigoriev and group at JGI. To facilitate the annotation, RNASeq was performed from RNA isolated from germinating *G. orontii* spores on glass plates and pooled RNAs isolated from different growth stages of powdery mildew on *A. thaliana*. The automated annotation will be supplemented with manual curation.

G. orontii genome appears to be the most complete powdery mildew genome to date. Interestingly, the estimated genome size of *G. orontii* in current assembly is almost half of the previously estimated size of any powdery mildew genome. The sequencing of other powdery mildews as part of this project is underway.

Characterizing Novel Archaeal Lineages in Salton Sea Sediments

Tarn, Jonathan*; Tarn, Ann H. (cupoflifenoodles@gmail.com); Valentine, David L.

University of California, Santa Barbara, CA.

Biological communities in extreme environments are often dominated by microorganisms of the domain Archaea. Abundances of Archaea are found in the hottest, saltiest, and most thermodynamically-limited ecosystems on earth. These taxing surroundings are thought to impose a state of chronic energy stress on resident organisms due to high costs of cellular maintenance relative to resource availability. Their adaptation to extreme, biologically-limiting conditions may therefore be an ancestral, domain-wide trait. In this study, we seek to characterize the archaeal community of the Salton Sea, where members of this domain are dominant over Bacteria. Previous work by Swan et al. in 2010 showed that gradients in salinity, sulfate, carbon and nitrogen across sediment horizons of the Salton Sea are linked to changes in archaeal dominance and community structure. In light of recent taxonomic revisions of the domain, we have reclassified the 107 published unique small subunit rRNA Archaeal sequences from the 2010 study, with most sequences falling into either the DPANN or TACK superphyla. DPANN sequences were very abundant in shallow, organic-rich sediments. In deeper, energy-limited strata, divergent clades of TACK superphylum Archaea were more prominent. Ongoing metagenomic work on these sediment communities has resulted in the assembly of several draft genomes of novel Archaeal groups. These data will help define genomic adaptations of Salton Sea Archaea to varying levels of energy stress as well as inform future cultivation efforts.

Response of a Low-Salinity Ammonia-Oxidizing Thaumarchaeote to Shifts in Environmental Conditions

Tolar, Bradley B.* (btolar1@stanford.edu); Francis, Christopher

Department of Earth System Science, Stanford University, Stanford CA.

The *Thaumarchaeota* are a globally abundant phylum of Archaea found in nearly every environment, ranging from the ocean to hot springs to soils. This group is involved in the nitrogen cycle through the oxidation of ammonia to nitrite, completing the first step of nitrification. Although broadly distributed, very few cultivated representatives of the Thaumarchaeota exist for targeted study. Most of the information known about this group results from molecular-based surveys of the environment, which indicates that *Thaumarchaeota* tend to be most abundant in areas of decreased light, temperature, and pH. Our laboratory has maintained stable enrichment cultures (90-97% purity) of a thaumarchaeote – designated "*Candidatus* Nitrosoarchaeum limnia" – that was originally enriched from sediment in the San Francisco Bay-Delta and is representative of the low-salinity "ecotype" found in rivers, lakes, and estuaries worldwide. This CSP project seeks to understand how such changes in environmental conditions impact the growth and metabolism of "*Ca*. N. limnia", including temperature,

salinity, pH, and inorganic nitrogen concentration. Using the production of nitrite from ammonia oxidation to calculate the growth rate of "*Ca*. N. limnia", in preliminary experiments, we observed the highest and lowest growth at the highest (30ºC) and lowest temperature (10ºC), respectively. Increased ammonia concentration and decreased pH also enhanced nitrite production. We detected transcripts from key nitrification-related genes (amoABC, nirK) under normal growth conditions, as well as those involved in chemotaxis, motility, and stress response. These genes will be the focus of comparative analysis of transcripts obtained from growth manipulation experiments, where we can use such information to gain physiological insights into this group of low-salinity Thaumarchaeota.

Combining Molecular, Genomic, and Isotopic Techniques to Examine the Diversity and Activity of Marine Thaumarchaeota in Monterey Bay

Tolar, Bradley B.* (btolar1@stanford.edu)[1]; Reji, Linta[1]; Smith, Jason M.[2, 3]; Chavez, Francisco P.[2]; Francis, Christopher[1]

[1]Department of Earth System Science, Stanford University, Stanford, CA. [2]Biological Oceanography Group, Monterey Bay Aquarium Research Institute, Moss Landing, CA. [3]Marine Science Institute, University of California Santa Barbara, Santa Barbara, CA.

Monterey Bay is a dynamic, coastal marine ecosystem with seasonal upwelling and steep environmental gradients. Additionally, it is a highly productive ecosystem with annual blooms of phytoplankton that feed fisheries. Amidst these other organisms, microbes involved in the nitrogen cycle are abundant and active. Previous work has been done characterizing those organisms involved in nitrification in Monterey Bay: ammonia-oxidizing bacteria (AOB) and archaea (AOA), and nitrite-oxidizing bacteria (NOB). Our ongoing CSP project seeks to determine whether variability in the abundance or diversity of AOA (members of the phylum Thaumarchaeota) relates to observed changes in nitrification rates, key environmental parameters, or to other microbial communities (including other nitrogen-cycling guilds). We took advantage of a two-year time series of monthly cruises to two stations in Monterey Bay and collected DNA samples from up to 10 depths between 5-500 m. Next-generation sequencing was employed to characterize the diversity of Archaea and Bacteria 16S rRNA, as well as Eukaryotic 18S rRNA, in these samples. This massive dataset was compared against a suite of environmental variables measured at the time of sample collection, including rates of nitrification and quantitative PCR-based estimates AOA abundance. We retrieved ~100 AOA OTUs (97% similarity) from three primary clusters (termed 'ecotypes') – shallow water column A (like Nitrosopelagicus brevis), deep water column B (abundant below 100 m depth), and marine sediment (similar to Nitrosopumilus maritimus). These ecotypes showed unique distributions both seasonally and by depth, and had varied relationships with observed nitrification rates, from strong to weak. OTUs belonging to NOB seemed to covary with those belonging to AOA, with significant positive associations ($p < 0.05$) between OTUs consistently retrieved in species co-occurrence networks created with subsets of the dataset (e.g., the top 500 OTUs). Moreover, the most abundant Thaumarchaeal and NOB OTUs were found to co-occur with members of Flavobacteria and Verrucomicrobia. No significant interaction patterns were observed for AOB OTUs in our dataset. Investigating the functional implications of these interaction patterns will aid in better elucidating the ecology of nitrifiers in marine systems. Furthermore, analysis of sequenced AOA metagenomes and metatranscriptomes will allow us to determine if the diverse AOA ecotypes have different gene content or other metabolic genes correlated to nitrification.

Draft Genome Sequence of the Pathogenic White-Rot Fungus *Phellinus noxius* OVT-YTM/97

Tsang, Karen S. W.* (kswtsang@hotmail.com); Wong, Chong Kim; Kwan, Hoi Shan

School of Life Sciences, The Chinese University of Hong Kong.

The white-rot fungal pathogen *Phellinus noxius* is the underlying cause of brown root rot disease. It has a global tropical and subtropical distribution and has a host range of over 200 plant species. This pathogen resides in soils and is usually spread through root-to-root contact, causing irrevocable damage in plant hosts and if left untreated, can lead to swift deterioration of the host health within a year. In recent years, incidences of *P. noxius* infected trees have increased dramatically worldwide, and in Hong Kong, many of these trees are old and valuable with cultural or historical significance. Currently there are no curative measures for this disease, despite the development of several biocontrol methods and fungicides.

The first genome sequence of *P. noxius* OVT-YTM/97 has been assembled and was isolated from a root sample obtained from an infected *Ficus microcarpa* tree in Hong Kong. The strain found to have particularly aggressive pathogenicity and exhibited resistance to biocontrol agents. The genome was sequenced on Illumina Miseq and PacBio RSII platforms followed by *de novo* assembly using the SPAdes assembler v3.9.1 and annotated. GeneMark-ES predicted a total of 9,957 protein-coding genes present, and carbohydrate-active enzyme (CAZyme) analysis using dbCAN v5 revealed 478 CAZymes and a strong resemblance to that of necrotrophic plant pathogens with large proportions of glycoside hydrolases and carbohydrate esterase families 1 and 10, which are required for the complete breakdown of plant cell walls for infection. In addition, analyses using the Database of Fungal Virulence factors (DFVF) indicated an abundance of β (1-3)-glucan synthases and class I and II chitin synthases, which control cell wall synthesis and septa production, providing structural stability for the hyphae. The presence of these carbohydrate-active enzymes and virulence factors could explain the high virulence of the pathogen and its ability to cause catastrophic fatalities in infected trees. The results generated from this study allow us to have a deeper understanding of the mechanisms of action and the hypervirulence of this pathogen. It also provides a basis for the development of prevention and treatments methods to protect trees from this pathogen.

Exploring the Robustness and Functional Potential of Boreal Shield Lake Archaean Ocean Analogues

Tsuji, Jackson M.* (jackson.tsuji@uwaterloo.ca)[1]; Schiff, Sherry L.[1]; Wu, Lingling[1]; Venkiteswaran, Jason J.[1, 2]; Molot, Lewis A.[3]; Elgood, Richard J.[1]; Paterson, Michael J.[4]; Neufeld, Josh D.[1]

[1]University of Waterloo, 200 University Ave W, Waterloo, Ontario, N2L 3G1, Canada. [2]Wilfrid Laurier University, 75 University Ave W, Waterloo Ontario, N2L 3C5, Canada. [3]York University, 4700 Keele Street, Toronto, Ontario, M3J 1P3, Canada. [4]IISD-Experimental Lakes Area, 111 Lombard Ave Suite 325, Winnipeg MB R3B 0T4, Canada.

The early Earth ocean of the Archaean Eon (≥2.5 billion years ago) is thought to have been anoxic, with high levels of dissolved iron and low levels of sulfur. Modern lakes have been identified with similar ferruginous (i.e., high-iron and low-sulfur) conditions in their anoxic zones and have been used to explore the biogeochemistry of the Archaean ocean in a unique, whole-ecosystem manner. Study of

such lakes, termed Archaean ocean analogues, has informed debate surrounding the deposition of Archaean banded iron formations. Photosynthetic, iron-oxidizing bacteria (photoferrotrophs) related to the type strain *Chlorobium ferrooxidans* have been identified in several Archaean ocean analogues, implying that ancient microbial iron oxidation could have contributed to the mixed iron oxidation states apparent in these formations. However, a major limitation to the use of Archaean ocean analogues is their natural rarity: only four permanently anoxic analogues have been studied to date. Boreal shield lakes number in the millions globally and naturally develop ferruginous waters once anoxic. Unlike existing analogues, most boreal lakes fully reoxygenate each spring and fall, disrupting anaerobic microbial communities and corresponding biogeochemical processes. We investigated the possibility of using boreal shield lakes as alternative Archaean ocean analogues by studying the robustness and functional potential of microbial communities in two such lakes in northwestern Ontario (Canada). Using 16S rRNA gene sequencing, we identified Chlorobi, with over 99.5% sequence identity to *C. ferrooxidans*, as a dominant group in the anoxic zones of both lakes over two summers. These organisms were accompanied by potential iron-reducing bacteria, sulfur-cycling bacteria, and methanotrophs, with shared dominant genera compared to those reported in another Archaean ocean analogue. Metagenome sequencing of anoxic zone samples confirmed that functional genes involved in sulfur oxidation and reduction (sox, dsr) and methanotrophy (pmo) were at high relative abundance compared to taxonomic marker genes. Detected Chlorobi were also found to have sulfur oxidation genes, raising the possibility for concurrent phototrophic iron- and sulfur-oxidation in the lake water columns. No iron oxidation gene is known for this group. Overall, our findings highlight the consistency of microbial communities across boreal shield lakes and sampling years despite seasonal reoxygenation and show metabolically unique microbial communities potentially useful for extrapolation to early Earth ecosystems. The potential involvement of boreal shield lake anoxic zone microbial communities in iron, sulfur, and methane cycling has implications for both modern limnology and early Earth research. Boreal lake anoxia is expected to increase in prevalence and duration with climate change, making the contribution of detected communities to net lake carbon balance, or iron fertilization of surface algal blooms, potentially important at regional or global scales. Moreover, having millions of ferruginous lakes opens the possibility for new types of early Earth research, such as inferring early Earth ecosystem controls by studying how modern microbial communities shift along natural environmental gradients. Boreal shield lakes present an unexpected and potentially transformative research platform to address questions of ancient and modern significance due to their robust anaerobic microbial consortia.

A pre-print of this work is available online at https://doi.org/10.1101/054478

Searching for Microbes in the Unchartered Venomous Territories of Marine Cone Snails

Ul-Hasan, Sabah* (sul-hasan@ucmerced.edu)[1]; Bowers, Robert M.[2]; Woyke, Tanja[2]; Sistrom, Mark[1]

[1]School of Natural Sciences, University of California Merced, Merced, CA, 95343, USA. [2]DOE Joint Genome Institute, Walnut Creek, CA 94598, USA.

Californiconus californicus is one of over 800 members of Conidae, a family of venomous marine gastropods. The California cone snail is a Conidae outgroup in (1) being the only species found along the California-Baja coast and (2) exhibiting extremely generalist predatory behavior. Conidae venom chemistry and phylogenetics are extensively studied, but surprisingly little is known about *C.*

californicus ecology or microbes within venomous animals. Found across three marine ecoregions and possessing an easily dissectible venom duct, *C. californicus* serves as a dynamic model for bridging macroecology and microbiology. We ask if microbes influence venom production in their hosts, focusing on *C. californicus* as an underutilized, cross-disciplinary model. We initially examined host microbial community and venom chemistry variation in one of the three marine ecoregions (Puerto Nuevo, MX). DNA was extracted from the foot, hepatopancreas, venom duct, and eggs of five animals for 16S and 18S iTag sequencing. Animals were collected at an arbitrary coastal point during low tide, with second (0.5 km) and third locations (1.5 km) northward. A total of twenty animals were collected for corresponding venom chemistry characterization. Water, sediment, and captive *C. californicus* were also sequenced as referenced controls. We hypothesize microbial communities differ according to anatomy, life stage, medium, time of day, and year, but maintain conservation between individuals in a given location. A conserved microbial community in the venom duct may infer previously unidentified symbiotic relationships, which we aim to further investigate via antibiotic treatments in a laboratory setting.

Interspecies Communities and Signaling in Plant Associated Bacteria

Venturi, Vittorio* (venturi@icgeb.org)[1]; Bertani, Iris[1]; Moronta Barrios, Felix[1]; Caballo Ponce, Eloy[2]; Ramos, Cayo[2]

[1]International Centre for Genetic Engineering and Biotechnology (ICGEB), Trieste, Italy. [2]University of Malaga - Malaga, Spain.

It is believed that most bacteria live in constant association or in the vicinity of many different bacterial species. In addition, most bacteria are now believed to produce and respond to chemical signals in a cell-density dependent manner in a process known as quorum sensing (QS). QS results in a synchronous response of bacterial populations which confers them a form of multicellularity and enables them to adapt and survive to challenging environments. Most bacterial QS studies thus far have involved mono-species (in fact mono-strain) set up which are rather distant from what occurs in nature. It is our major interest to investigate chemical signaling in interspecies bacterial communities and the possible role of chemical signals in plant-bacteria interactions.

We are using beneficial bacterial endophytes which enter the plant via the rhizosphere to study interspecies interactions and dynamics of endophytic life style and multispecies community formation. This could lead to the use of endophytes as microbial products to improve plant health and sustainability in agriculture. In parallel we are studying bacterial interspecies signaling using a plant disease as a model. The olive knot disease caused by *Pseudomonas savastanoi* results in tumors/galls in olive trees; we have established that inside the tumors, together with the pathogen, other bacterial species interact and communicate with the pathogen resulting in mutual benefit and in a more aggressive disease.

A Rhizosphere-Scale Investigation of Root Effects on Wetland Methane Dynamics

Waldo, Nicholas B.* (nbwaldo@uw.edu)[1]; Fadely, Ella C.[2]; Ardissono, Robert J.[1]; Gough, Heidi L.[1]; Chistoserdova, Ludmila[3]; Beck, David A. C.[3]; Neumann, Rebecca B.[1]

[1]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA. [2]Department of Geology, Carleton College, Northfield, MN. [3]Department of Chemical Engineering, University of Washington, Seattle, WA.

Methane (CH_4) is a greenhouse gas emitted by wetlands. In the anoxic soils of wetlands, CH_4 is produced by anaerobic methanogens from acetate and hydrogen produced by anaerobic fermenters. However, much of this CH_4 is oxidized to CO_2 as it passes through oxic zones before emission to the atmosphere. A key region for microbial processes is the rhizosphere, where roots introduce organic carbon and O_2 belowground allowing aerobic heterotrophs to compete with fermenters for organic substrates and with aerobic methanotrophs for O_2. These microbial interactions control rates of CH_4 production and oxidation. This study tracked the belowground movement of organic carbon and O_2 from plants to quantitatively understand rhizosphere microbial processes key to wetland CH_4 dynamics.

Fluxes of CO_2 and CH_4 from boxes of peat containing *Carex aquatilis* were used to determine net primary productivity (NPP), belowground respiration, and total production, oxidation and emission of CH_4. We tracked O_2 concentrations around the roots using optical O_2 sensors (optodes). Twice, the plants were exposed to headspace $13CO_2$: at mid-growth and when fully grown. The plants fixed the 13C, some of which was exuded through the roots and used by microbes. We tracked the isotope ratio of emitted CO_2 and CH_4 to establish the timing and extent of 13C being respired and fermented to CH_4. Soils samples for microbial DNA analyses were collected at multiple time points after labeling using optode data guide collection from zones of differing oxygenation. Labeled (13C) DNA was separated from unlabeled DNA using ultracentrifugation to identify microbial populations that had used the root exudate carbon, providing insight into how microbial competition and substrate selection vary with root inputs of oxygen and carbon. Together, data from the experiment will elucidate the plant-microbe interactions that control rates of methane production and oxidation in the rhizosphere of wetland plants.

Metagenome-Assembled Genomes from the Beaufort Sea Reveals a Unique Arctic Ocean Microbiome

Walsh, David A.* (david.walsh@concordia.ca)[1]; Lovejoy, Connie[2]

[1]Concordia University, Montreal, Canada. [2]Laval University, Quebec City, Canada.

Arctic Ocean surface waters are fresher, more stratified, and enriched in terrestrial organic matter compared to other oceans. Given these observations, it is intriguing to speculate on the existence of endemic arctic taxa. However, evidence for AO endemism is rare and those taxa inhabiting the AO exhibit a seemingly much wider oceanographic distribution. To investigate the degree of AO endemism, we used metagenomic binning to reconstruct bacterial genomes from a metagenomic dataset from the Beaufort Sea in the western AO. Near complete genomes represented a wide phylogenetic diversity, including numerous uncultivated bacterial lineages. Comparing the arctic

genomes with globally distributed marine metagenomes revealed a diversity of biogeographical patterns, including some that appeared to support the existence of endemic AO populations. Given the unprecedented warming and freshening of the AO, identifying arctic ecotypes and their genomic traits will provide insight into how the AO microbial communities will respond to such change.

CRISPR-Cas9 Genome Editing and Gene Regulation Tools for Rapid Engineering of *Yarrowia lipolytica*

Wheeldon, Ian R.* (iwheeldon@engr.ucr.edu)[1]; Schwartz, Cory[1]; Frogue, Keith[1]; Shabbir-Hussain, Murtaza[2]; Blenner, Mark A.[2]

[1]University of California Riverside, Riverside CA USA. [2]Clemson University, Clemson SC USA.

The oleaginous yeast *Yarrowia lipolytica* has been widely studied and engineered for the renewable production of lipid-based fuels and chemicals. To facilitate the rapid development of new *Y. lipolytica* strains, we have adapted the type II CRISPR-Cas9 system from *Streptococcus pyogenes* to function in this organism. Critical to the success of the system was the design of a synthetic RNA polymerase III (Pol III) promoter that produced high expression levels of sgRNAs. The best performing sgRNA promoter was the fusion of a glycine tRNA and a truncated version of the native *Y. lipolytica* SCR1 Pol III promoter, which achieved gene disruption rates upwards of 90%. Building from this system, we designed and implemented both a CRISPRi system for gene knockdowns and a genome editing tool for markerless integration into standardized chromosomal loci. Using the CRISPRi system, we investigated the effect of repressing non-homologous end joining function (KU70, KU80, DNL4) on homologous recombination. The CRISPR-Cas9 system was then used to accomplish markerless integration of genes into the *Y. lipolytica* genome and investigate integration and expression of heterologous genes from different loci across the genome. Several genomic loci amenable to high efficiency integration were identified and expression of a fluorescent protein from these sites was quantified. These CRISPR-Cas9 based genome editing tools are the basis for a new joint project with the DNA Synthesis program at JGI, were we will create a genome-wide knockout technology for use in *Y. lipolytica*.

Genomic Encyclopedia of Bacterial and Archaeal Type Strains, Phase III and IV: The Genomes and Pangenomes of Soil and Plant-Associated and Newly Described Type Strains

Whitman, William B.* (whitman@uga.edu)[1]; Woyke, Tanja[2]; Zhou, Y.-G.[3]; Klenk, Hans-Peter[4]; Kyrpides, Nikos[2]

[1]Department of Microbiology, University of Georgia, Athens, GA. [2]DOE Joint Genome Institute, Walnut Creek, CA. [3]China General Microbiological Culture Collection Center (CGMCC), Institute of Microbiology, Chinese Academy of Sciences, Beijing, China. [4]School of Biology, Newcastle University, Newcastle upon Tyne, UK.

The Genomic Encyclopedia of Bacteria and Archaea (GEBA) project was launched by the JGI in 2007 as a pilot project to sequence about 250 bacterial and archaeal genomes of elevated phylogenetic diversity. Here, this approach was extended to type strains of prokaryotes associated with soil or

plants and their close relatives as well as type strains from newly described species [1]. Individual investigators were invited to submit DNA from any of these type strains to JGI for sequencing and annotation. Since the project began in the Fall of 2013, individual investigators proposed 852 type strains for genome sequencing, and 588 of these strains were approved. The major reason the projects were not approved was that sequencing was in progress elsewhere. Sequences for 256 genomes have been completed or are in progress. Projects approved were largely for type strains from soils, plant associated and saline soils and were contributed by investigators from 14 nations, chiefly India, Spain, United Kingdom, South Africa, and China. In addition, approval was obtained for sequencing 328 type strains provided by the China General Microbiological Culture Collection Center (CGMCC), which possesses a large collection of type strains isolated in China. Sequences for 270 of these have been completed or are in progress. Therefore, this project has significantly increased the number of genome sequences for type strains, especially among plant and soil associated species. In the Fall of 2016, an additional phase of the project (GEBA IV) was initiated to sequence pangenomes of plant and soil associated species. So far, 13 investigators have expressed the intention to contribute DNA from 282 stains for the determination of 32 new pangenomes.

References: [1] Whitman, W.B., Woyke, T., Klenk, H.-P., Zhou, Y-G., Lilburn, T.G., Beck, B.J., Kyrpides, N.C. (2015). Genomic Encyclopedia of Bacterial and Archaeal Type Strains, Phase III: the genomes of soil and plant-associated and newly described type strains. S*tandards in Genomic Sciences*, 10(1). doi:10.1186/s40793-015-0017-x.

RNA-Editing in Model Mushroom *Coprinopsis cinerea*

Xie, Yichun* (xieyichun5050@gmail.com); Chang, Jinhui; Kwan, Hoi Shan

School of Life Sciences, The Chinese University of Hong Kong, New Territories, Hong Kong.

Introduction. RNA-editing is an unusual form of post-transcriptional event which can recodes the hereditary information by changing the nucleotide sequence of RNA molecules. It diversifies the transcriptome and proteome of the organism by several means. To date, research on RNA-editing in fungi is rare. *Coprinopsis cinerea*, a typical mushroom, is one of the model organisms to study homobasidiomycetous fungi. It has a short life-cycle and is easy to be cultivated in laboratory.

Aims. To identify RNA-editing sites and editing pattern in *C. cinerea*. Also, to determine the influence of RNA-editing on transcription factors (TFs) and the expression of downstream genes.

Methods. We first sequenced and analyzed the total RNA collected from mycelium of medium cultured *C. cinerea* to predict RNA-editing candidates. The RNA was then reverse transcribed with specific primers which target on the predicted sites. cDNA prepared were sequenced separately using Sanger method. To confirm the RNA-editing event, we aligned the sequencing read to the reference genome. Expression level is measured by real-time PCR.

Results. *In silico* analysis of RNA-seq data revealed that several TFs, involving in CAZymes expression, kinases expression and other cell activities, are likely to have RNA-editing. We have identified one TF that one of the base of its mRNA is edited from A to I. This editing event changes the expression level of downstream genes. We also looked into the mechanism of the editing event. However, no ADAR gene was found. The genome of *C. cinerea* has 3 predicted ADAT genes but the function has not been confirmed.

Discussion. In this study, we uncovered one A-to-I RNA-editing site on the transcription factor of *C. cinerea*, which is the first report of RNA-editing in this model mushroom. Here, we have not analyzed the editing pattern of the whole genome or different developmental stages. RNA-editing can help to improve the diversity of gene expression products in an organism, which means that it may help the creature to survive under stress environments or provide extra diversities in protein coding genes. Stage-specific RNA-editing events and stress-stimulated RNA-editing events worth further studies.

Charactering Microbe Response to P Availability in Panama Soils by Long Term Fertilization

Yao, Qiuming* (yaoq@ornl.gov)[1]: Li, Zhou[1]; Yang, Song[1]; Wright, Joseph S.[2]; Guo, Xuan[3]; Biswas, Abhishek[1]; Tringe, Susannah G.[4]; Hazen, Terry[1, 3]; Turner, Benjamin L.[2]; Mayes, Melanie[1]; Pan, Chongle[1, 3]

[1]Oak Ridge National Laboratory, Oak Ridge, TN. [2]Smithsonian Tropical Research Institute. [3]University of Tennessee, Knoxville, TN. [4]DOE Joint Genome Institute, Walnut Creek, CA.

The ecological impact of P availability in terrestrial ecosystems was investigated by performing long-term P fertilization experiments. In previous studies, stable isotope labeling, signature molecular profiling, enzyme assay, denaturing gradient gel electrophoresis and 16S genotyping were commonly applied to characterize the impact and response on soil microorganisms. Here by the most current proteogenomic techniques, we obtained deeper information of soil microbe response on P availability at Gigante peninsula in Panama. Field soil samples from P-fertilized plots (#1 and #30) and control plots (#6 and #36) were used to perform large scale metagenomics and metaproteomics. Overall 1-3 millions of genes were identified from each metagenomics sample, and 140 enzyme commission (EC) numbers were identified as significantly different in genes. Downstream 30,000 peptides were identified by mass spectrometry for each metaproteomics sample, which were assembled to 6,000-9,000 proteins in each. Genome abundance shows significant high phosphatases, phospholipases, nucleases, pectin and hemi-cellulose degradation enzymes, proteases and peptidases which suggests the microbes in P-deficient plots were contributing main functions in P acquisition from organic forms, and degrading simple carbon and nitrogen resources. In P-rich plots, microbes show significant gene abundance in utilizing aromatic carbons, and reducing nitrogen and sulfur compounds. Metagenome assembly revealing the diverse phylogeny structure of phytases in our communities and their neighboring genes, suggesting important role of both P acquisition and soil organic matter utility from microbes. Two assembled near complete genomes out of twelve encoding phytases and phospholipases. Proteomes suggest possible molecular linkage through phytate and glucose-6-phosphate to utilize both phosphorus and carbon effectively in the system. Our enzyme assays validated phosphorus and carbon metabolic effects under different P availability conditions through microbe responses. From this study, we concluded and hypothesized under the difference of long term P availability, soil microbiome shifted its community functionalities toward adapting and balancing various nutrient transformations in P,C,N,S.

Validation and Characterization of Development-Related microRNA-like RNAs in the Mushroom-Forming Fungus *Coprinopsis cinerea*

Yuet Ting, Lau*; Lau, Amy (1155032273@link.cuhk.edu.hk); Kwan, Hoi Shan

School of Life Science, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong.

Introduction. *Coprinosis cinerea* is a model organism for the study of homobasidiomycete fungi due to its short life cycle and easy cultivation in the laboratory. The function and regulation mechanism of microRNA (miRNA) of animals and plants have been extensively studied. Although emerging studies have suggested that microRNA-like RNAs (milRNAs) are present in fungi and their biogenesis pathway might not be identical as conventional miRNA in plants and animals, milRNAs are still elusive for. The lack of studies on the role of milRNAs in fungi hinders the improvements of mushroom strains.

Aim. We aim to understand the development of mushroom-forming fungi at the molecular level, including RNA-mediated regulation.

Methods. The presence of milRNAs has been validated experimentally through Northern blotting and their expression level has been quantified using Stem-loop Reverse Transcription Real-time PCR (Q-PCR). Computational methods have been applied subsequently to characterize the roles of milRNAs in the development of mushrooms, especially during the transition from mycelium to primordium stage.

Results. We have identified, from *C. cinerea* genome, 16 putative milRNAs candidates with sizes around 21nt and a group of Dicer homologs (CC1G_00230, CC1G_03181, CC1G_13988) that are specific to mushroom-forming fungi. The PZA domain is present in the Dicer-like (DCLs) proteins of mushroom, while it is absent in all other fungal DCLs. Besides, for these 16 putative milRNAs candidates, one of them showed higher expression in mycelium stage and 15 showed higher expression in primordium stage.

Discussion. The mushroom-specific PAZ domains were grouped closely with the animal PAZs and plant PAZs and the Dicer_dimer domain sequences alone could distinguish fungal DCLs from their homologs in animals and plants. The DCLs domains of mushroom suggest Dicer genes duplicated and diversified independently in early evolution of all three multicellular kingdoms-animals, plants and fungi. A better understanding of the biogenesis pathways and regulatory mechanisms milRNAs in *C. cinerea* could improve the development of homobasidiomycete fungi, such as edible ad medicinal mushrooms, contributing to evolutionary studies of miRNA among kingdom in the near future.

Melanin Production across Species: A Fungal Comparative Genomics Case Study

Zeiner, Carolyn A.* (zeiner@bu.edu); Siletti, Cheta E.; Talbot, Jennifer M.

Department of Biology, Boston University, Boston, MA, USA.

Background/Question/Methods
Fungal diversity in soil communities may affect soil carbon (C) storage through species-specific differences in the hyphal content of melanin, a recalcitrant pigment. Microbial comparative genomics has become an increasingly powerful tool to link this diversity to ecosystem function by assessing

the functional potential of individual species based solely on their genomes. However, despite the increased availability of sequenced genomes, linking this detailed genomic data to metabolic pathways of ecological importance remains challenging due to variability in genome annotation quality and completeness. This challenge becomes even greater for fungal genomes, for which accurate gene annotation often lags behind their bacterial counterparts. Here we conducted a customized fungal comparative genomic analysis to investigate the distribution of melanin production across 29 fungal taxa. We hypothesized that functional gene abundance could predict lab-based measurements of fungal melanin content across species. Using genomic data obtained from the U.S. Department of Energy's 1000 Fungal Genomes Project, we calculated the number of genes assigned to each protein (Pfam) domain annotation in each genome. We then tested for correlation between hyphal melanin content and two sets of genomic data: (1) gene counts of Pfam domains that have been experimentally verified to be involved in melanin biosynthesis (e.g., laccases) and (2) gene counts for all Pfam domains in whole genomes. Due to incomplete Pfam annotations, both analyses required manual assignment of metabolic pathways and cellular functions to Pfam domains and their associated proteins using various resources, including JGI-provided InterProScan annotations, BLAST results from the NCBI non-redundant protein database, UniProt searches, and the literature.

Results/Conclusions
In analysis (1), we found positive correlations between hyphal melanin content and 3 proteins known to be involved in melanin biosynthesis pathways: laccase, scytalone dehydratase, and polyketide synthase. Manual annotation increased the occurrences of laccase-related Pfam domains from 0/29 genomes to 29/29 genomes. In analysis (2), we identified a total of 157 Pfam domains that positively correlated with hyphal melanin content, including domains involved in biosynthesis of phenolic melanin precursors, primary and secondary metabolite biosynthesis, cell signaling, and redox reactions. Twenty domains were negatively correlated with melanin content and were associated with DNA replication processes and stress response, suggesting a tradeoff between melanin biosynthesis and cellular growth. In these whole genome queries, manual annotation allowed us to use sparsely annotated domains that otherwise would have been discarded, resolve disparities among conflicting annotations, generate higher-quality data by replacing anomalous annotations with fungal-specific information, and assign relevant biochemical pathways. Our analysis suggests that it may be possible to predict the approximate amount of melanin in a given fungal species using genomic information alone, which could be of value in predicting soil C storage from metagenomic data. Additionally, this work highlights the need for refinements to existing functional annotation pipelines for fungal genomes, which could increase the speed and accuracy of fungal comparative genomic investigations.

ViCA: Identifying Highly Divergent Viruses Using Supervised Learning across the Homology-Composition Spectrum

Zhang, Qingpeng* (qingpeng@lbl.gov); Rivers, Adam; Tringe, Susannah G.

DOE Joint Genome Institute, Walnut Creek, CA.

Viruses are important drivers of ecosystem dynamics, nutrient turnover and disease. While the role of viruses in disease is well studied, their roles in ecosystems are becoming more apparent through increased metagenomic surveys. New techniques now open the door to surveys of single stranded DNA viruses and RNA viruses, which may be equally abundant and important in some environments. These new data sources and the growing amount of metatranscriptome and metagenome data

present the opportunity to find highly divergent new viruses. However the process of identifying divergent viruses, that by definition lack closely related representatives, presents challenges.

The identification of viruses in sequence data is a challenging subset of the larger problem of taxonomic classification. Most approaches to taxonomic classification rely on marker genes or on sequence homology. Compositional features are more commonly used for the related problem of binning metagenomic sequences into genome bins. Combining multiple types of genomic features has the potential to improve taxonomic classification. Information is organized at different scales across the genome selected for by different environmental pressures. By selecting biologically informed features at multiple scales we could bring prior information about the structure inherent in our data to the problem.

Here we present ViCA - a regularized logistic classifier trained on features across the homology-composition spectrum to classify metagenomic sequences, with a focus on identifying novel viral lineages. K-mer frequency, codon usage and the identification of protein families provide features that reduce variance. Combining data from multiple feature sets strengthens both the precision and recall of the classifier at most taxonomic levels. We have shown that the combined feature set generalizes well in classifying novel groups of viruses, and outperforms alignment-based classification. The ViCA classifier is implemented on the large-scale distributed machine learning platform Apache Spark and will be available along with analysis code.

Comparative Transcriptomic Study towards Two *Aspergillus oryzae* Strains

Zhong, Yiyi* (yiyi11zhong@gmail.com); Nong, Wenyan; Xing, Lei; Lu, Xi; Kwan, Hoi Shan

The Chinese University of Hong Kong, Hong Kong.

Aspergillus oryzae, equipped with an arsenal of secreted enzymes to break down proteins and complex polysaccharides, is critical in food processing industry, especially soy sauce fermentation. The strain RD2 is commercially used in Chinese soy sauce industry. A degenerated strain TS2 obtained during production would yield poor-quality soy sauce when used in Koji making. We aimed to conduct a comparative transcriptomic study to explain the phenotypic differences between two strains. We first cultivated RD2 and TS2 in the mixture of steamed soy beans and wheat flour, and sampled in duplicate in three Koji-fermentation stages: mycelium expansion (ME), early sporulation (ES) and mature sporulation (MS). Then the extracted RNA was sequenced in Illumina Hiseq 2000 Platform. We identified 1589 differentially expressed genes (DEGs) in ME, 686 DEGs in ES and 225 DEGs in MS. Among them, much effort has been devoted to those related to amylases and proteases. According to the result, the higher expression level of both amylases and proteases in RD2 throughout all stages directly explains the better performance of RD2 in Koji fermentation. While in the degenerated strain TS2, more genes related to glycolysis and amino acid catabolism are expressed at a higher level. We suppose that the poor-quality soy sauce produced by TS2 is probably due to the poorer hydrolase expression and the overuse of flavor compounds for energy, and RD2 could not compete with TS2. In conclusion, RD2, which possesses the metabolic saccharolytic and proteolytic activities, is a good fermentative strain in soy sauce production with the great capacity of releasing compounds for flavor formation and subsequent fermentation, while the spontaneous degeneration would cause loss of the industrially favorable phenotypes but gain of growth dominance. The transcriptomic comparison helps to identify the molecular markers related to good soy sauce fermentation, which in turn defines good Koji in a molecular way. Further, we will investigate the mechanism and function of *Aspergillus*

oryzae metabolism and identify more molecular markers that can be applied to the selection and breeding of superior soy sauce production strains.

A Tale of Transition of Sex Determination from XY to ZW in *Salix purpurea*

Zhou, Ran* (razhou@mix.wvu.edu)[1]; Carlson, Craig H.[2]; Rodgers-Melnick, Eli[1]; Schmutz, Jeremy[3]; Jenkins, Jerry[3]; Tuskan, Gerald A.[4]; Smart, Lawrence B.[2]; DiFazio, Stephen[1]

[1]Department of Biology, West Virginia University, Morgantown, WV 26506-6057. [2]Horticulture Section, School of Integrative Plant Science, Cornell University, New York State Agricultural Experiment Station, Geneva, NY 14456. [3]HudsonAlpha Institute of Biotechnology, Huntsville, AL. [4]Biosciences Division, Oak Ridge National Lab, Oak Ridge, TN.

Recent studies of the sex determining systems in multiple *Salix* and *Populus* (Salicaceae) species indicate that while *Populus* displays an XY system (heterogametic males) on chromosome 19, *Salix* possesses a ZW (heterogametic females) system on chromosome 15. However, these studies in Salix have not provided a comprehensive set of candidate genes for sex determination due to poor genome assembly in that highly polymorphic region. Here, by taking advantage of our completed shrub willow genome (*Salix purpurea*, clone ID 94006) and genotype data, we use quantitative trait locus mapping and a genome-wide association study to show that *S. purpurea* also has a ZW system on chromosome 15. Suppressed recombination, high structural polymorphism, and an abundance of transposable elements are also observed on the sex chromosome. RNA-seq identified several genes showing sex-biased expression within the sex-determining region. Additionally, our results suggest that chromosome 19 in *Salix* may still display vestigial features of the ancestral sex chromosome. Thus, we hypothesize that chromosome 19 is the ancestral sex chromosome in *Salix* and *Populus*, and there was a transition from XY on chromosome 19 to ZW on chromosome 15 in *Salix*. The occurrence of contrasting systems of sex determination in these two genera, which share high synteny across the genome, provides excellent opportunities to study the causes and consequences of the transition of sex determination systems in plants.

Author Index

A

Abu Salim 36
Adams, Cat 1
Agosto, Carolyn 86
Aguilar Pontes, Maria Victoria 10
Ahrendt, Steven R. 10, 68
Alexander, William G. 54
Alford, Betsy 43
Alteio, Lauren V. 11
Altendorf, Kayla 34
Alvarez, Bryan R. 60
Alvarez-Cohen, Lisa 59
Amasino, Richard M. 32
Amirebrahimi, Mojgan 87
Amos, Kirtley B. 12
Anderbehan, Fekadu 48
Andersen, Mikael 1, 55
Anderson, James 34
Anderson, Lisa A. 12
Andreopoulos, William B. 68
Andrew, Liu 89
Andrews, Brenda 2
Ané, Jean-Michel 3
Angle, Jordan C. 30, 71
Antell, Edmund 59
Ardissono, Robert J. 95
Arellano, Ana 69
Arlow, Dan 50
Au, Chun Hang 55
Ayuso-Fernandez, Ivan 13

B

Babbitt, Patsy 2
Badis, Yacine 14
Baidoo, Edward 48
Bailey, Jake 63
Baker, Brett 41
Baker, Scott E. 31, 32, 55
Baldrian, Petr 14
Balunas, Marcy J. 59
Banfield, Jillian F. 64
Bargar, John 20
Barry, Kerrie 18, 23, 46, 48, 60, 61, 79
Battaglia, Evy 31
Baybayan, Primo 21, 56
Bayram Akcapinar, Günseli 24
Beam, Jacob P. 47
Beck, David A. C. 25, 95
Becraft, Eric D. 40
Bector, Rathin 50
Behringer, Megan G. 41
Beller, Harry R. 15
Bellows, Wendy K. 40
Beman, Michael J. 16
Benoit, Isabelle 31
Bertani, Iris 94
Bertilsson, Stefan 62
Bertsch, Jon 71
Betto-Sagahon, Janeli 82
Bewick, Adam J. 68
Bianchi, Thomas 69
Billings, Andrew 11
Binder, Morgan 39
Birch, James 74
Biswas, Abhishek 16, 98
Blanchard, Jeffrey 11
Blenner, Mark A. 17, 96
Blow, Matthew J. 86
Bohrer, Gil 30
Boon, Nico 80
Borton, Mikayla A. 30
Bowen, Ben 68
Bowers, Robert M. 93
Branco, Sara 21, 61
Brawley, Susan H. 79, 87
Brem, Rachel 68
Brocchieri, Luciano 17
Broman, Elias 62
Brown, C. Titus 2
Browne, Daniel R. 18
Brown, Joseph 40
Brown, Julia M. 40
Bruckbauer, Steven T. 18
Bruns, Tom 21, 61
Brutnell, Tom 46, 87
Buell, Robin C. 65
Bukhman, Yury 54
Bükün, Bekir 43
Buluc, A. 42
Burdo, Brett 65
Bushnell, Brian 16

C

Caballo Ponce, Eloy 94
Cai, Feng 24
Campbell, Barbara J. 19
Cantu, Dario 28
Cardarelli, Emily L. 20
Cardon, Zoe G. 77
Carlson, Craig H. 102
Carlson, Joseph W. 20, 45, 87
Carver, Akiko A. 21
Casado Lopez, Sara 32
Catalan, Pilar 32
Catcheside, David 68
Chai-Crosskey, Juan 16
Chain, P. 42

Chakraborty, Shreyasee 21
Chan, Eric W. 81
Chang, Jinhui 22, 23, 55, 97
Chang, Jui-Jen 62
Chang, Peter 43
Chan, Po Lam 22
Chavez, Francisco P. 91
Cheng, Jan-Fang 10, 54
Cheng, Xuanjin 55
Chen, Jay 23, 81
Chenthamara, Komal 24, 36
Chern, Mawsheng 60
Cheung, Man Kit 55
Chin, Chen-Shan 28
Chin Mei, Chan 36
Chistoserdova, Ludmila 25, 95
Chovatia, Mansi 86
Chris, Daum 87
Chrisler, William B. 31
Ciobanu, Doina 10
Clarkson, Sonya M. 67
Clark, Tyson A. 26
Clingenpeel, Scott 60
Clum, Alicia 10, 16, 28, 68
Cole, Benjamin J. 26
Coleman-Derr, Devin 27, 72
Coleman, Maureen 75
Concepcion, Gregory T. 28
Connor, Elise R. 28
Contreras-Moreira, Bruno 32
Cook, Douglas R. 43
Coruzzi, Gloria 3
Cox, Michael M. 18
Crain, Jared 34
Cramer, Grant R. 28
Crill, Patrick M. 39
Crump, Byron C. 29
Cumming, Jonathan 57
Cunningham, Brady R. 44
Currie, Cameron R. 50

D

Daly, Rebecca A. 30, 71
Damashek, Julian 41
Dangl, Jeffery L. 27, 60
Dannebaum, Richard O. 68
Daskalov, Asen 68
Daum, Chris 48, 68
Davidson, Patrick 20
DeAngelis, Kristen 11
DeBolt, Seth 12, 84
DeGraaf, Stephanie 27, 72
DeHaan, Lee 34
de Leon, Natalia 65
Delledonne, Massimo 28
De Lorenzo, Laura 12, 84
del Rio, Tijana G. 60
Denef, Vincent 80
DeSalle, Rob 3
Des Marais, David L. 32, 70
Deutschbauer, Adam 27, 35, 49, 86
Deutsch, Samuel 33, 54, 86
Devarenne, Timothy P. 18
DeVirgilio, Jeremy 53
de Vries, Ronald P. 10, 31, 32
D'haeseleer, Patrik M. 44
DiFazio, Stephen 102
Dippel, Andrew B. 33
Dlakic, Mensur 47
Doering, Drew T. 53
Dohnalkova, Alice 31, 32
Donaho, John A. 33
Dong, Wenming 64
Dopson, Mark 62
Dorn, Kevin M. 34
Doty, Sharon L. 35, 49, 60
Dove, Nicholas C. 36
Druzhinina, Irina 24, 36
Dudycha, Jeffry L. 44
Dunn, Christophe 28
Duong, Phat 60
Dupree, Paul 48

E

Ecker, Joseph R. 28
Edger, Patrick P. 32
Edwards, Joseph A. 38, 84
Egan, Rob S. 42
Elgood, Richard J. 92
Elkins, James G. 67
Ellison, Chris 68
Emerson, Joanne B. 39
Eng, Kevin 21
Erb, Tobias J. 4
Evans, Robert 33

F

Fadely, Ella C. 95
Feltcher, Meghan E. 26, 60
Fergusson, Elizabeth A. 40
Fields, Matthew W. 66
Figueroa-Balderas, Rosa 28
Fikre, Asnake 43
Fine, Lindy M. 29
Firrincieli, Andrea 35
Fischer, Michaela 36
Flood, Beverly 63
Flores. Gilberto E. 33
Forte, Ana 36
Forth, Michael 86
Fox, Evan 59
Francis, Christopher 20, 41, 90, 91
Frank, Carolyn 35
Freitag, Michael 68
Freitas, Nancy L. 39
Frey, Serita 11
Fritts, Ryan K. 41
Frogue, Keith 96

G

Gachon, Claire M. M. 14
Gage, Joseph 65
Garvetto, Andrea 14
Gasch, Audrey P. 54
Georganas, E. 42
Gholamian, Jason 19
Gilmore, Sean P. 76
Gladieux, Pierre 68
Glass, Lousie N. 68
Gliessman, Jennifer R. 41
Gonzalez, Isai S. 60
Goodstein, David 20, 45, 48, 85, 87
Gough, Heidi L. 95
Grant, Sarah R. 60
Greenspan, Alex 43
Greenwold, Matthew J. 44
Griessemer, Marc 44
Grigoriev, Igor 10, 21, 55, 61, 68
Grimwood, Jane 34, 48, 78, 87
Grujic, Marica 36
Guay, Lisa M. 12
Gunter, Lee 23
Guo, Xuan 16, 98
Guss, Adam M. 67
Gutiérrez, Rodrigo 3

H

Haas, Fabian B. 78
Habtesselassie, Mussie Y. 73
Hall, Richard 46, 52
Hamiton, Matthew 86
Hammond, Ming 33
Hann-Soden, Christopher 68
Hansel, Colleen 4
Haridas, Sajeet 55, 68
Harmon-Smith, Miranda 54
Harnish, Rob 63
Harting, John 56
Hartman,Wyatt H. 45
Hart, Stephen C. 36
Hatas, Emily 21
Hatzenpichler, Roland 66
Hawkes, Christine V. 28
Hayes, Richard D. 45
Hazen, Terry 98
Healey, Adam 46
Heckwolf, Marlies 65
Heiner, Cheryl R. 46
Henske, John K. 76
Herfort, Lydie 29
Hernandez, Alvaro G. 46
Herring, Robert 85
Hervey, IV, William J. 16
Heyse, Jasmine 80
Hicks, Michael A. 12
Hildén, Kristiina S. 32
Hillson, Nathan 50
Hittinger, Chris T. 53, 54
Hofmeyr, S. 42
Hon, Ting 26
Horwath, William R. 45
Hossain, Shakhawat 87
Hoyt, David W. 30, 61
Huang, Chieh-Chen 62
Huang, Pu 46
Huang, Qianli 55
Hubbard, Susan 64
Huffaker, Alis 35
Hulfachor, Amanda B. 53

I

Inskeep, Dylan J. 47
Inskeep, William P. 47, 66
Ishikawa, Toshiki 48
Ivanova, N. 42

J

Jain, Rashmi 60
James, Timothy Y. 10, 68
Jawdy, Sara S. 23, 87
Jay, Zackary J. 47
Jenkins, Jerry 18, 34, 46, 48, 79, 85, 87, 102
Jeremian, Richie 40
Jing, Beibei 48
Jin, Zhao 60
Johnson, Jenifer 87
Johnson, Joel E. 39
Johnston, Michael D. 30
Joseph, Matthew 35
Joubert, Pierre M. 35, 49
Juenger, Thomas E. 70, 87

K

Kaeppler, Heidi 3, 63
Kaeppler, Shawn 63, 65
Kahraman, Abdullah 43
Kamagata, Yoichi 66
Kang, Jin 54
Kang, Justine 50
Kanke, Matthew 63
Kantarski, Traci 34
Kardish, Melissa R. 70
Kawai-Yamada, Maki 48
Keasling, Jay D. 15, 50
Kelley, John F. 33
Kennedy, Megan 87
Keren, Ray 59
Kessler, John D. 81
Khadempour, Lily 50
Khorasani, Mahsa 35
Khosravi, Claire 31
Kiene, Ronald P. 74
Kilmer, John 38
Kimbrel, Jeffrey 44, 51, 83
Kim, Gwang Hoon 14

Kim, Peter W. 50
Kim, Soo-Hyung 35
Kingan, Sarah B. 52
Klassen, Jonathan L. 59
Klenk, Hans-Peter 96
Klochkova, Tatyana A. 14
Kominek, Jacek Z. 53
Kopchinskiy, Alexey 36
Koppisch, Andrew T. 18
Korajczyk, Peter 57
Korlach, Jonas 52
Korzeniewski, Frank 55
Kowalczyk, Joanna 31
Kozubal, Mark A. 47
Krause, David J. 54
Krause, Sascha M. B. 25
Kridelbaugh, Donna 67
Kuo, Alan J. 21, 55, 61, 68
Kuo, Rita C. 68
Kuroda, Kyohei 66
Kurtzman, Cletus P. 53
Kwan, Hoi Shan 22, 23, 55, 58, 92, 97, 99, 101
Kyrpides, Nikos 42, 71, 96

L

Labbé, Jessy 81
Labonté, Jessica M. 40
LaButti, Kurt 68
Laciny, Alice 36
Lambert, Christine 21, 56
Landick, Robert 54
Lane, Todd 51, 83
Lao, Jeemeng 48
La Reau, Alex J. 56
Larsen, Peter E. 57
Larson, Steve 34
Lau, Amy 99
Lau, Rebecca 68
Lau, Yuet Ting 58
Lawrence, Jennifer E. 59
Leach, Jan E. 5
Lee, Jessica 41
Lee, Kevin 59
Lee, Patrick 66
Lee, Po-Heng 66
Lee, Yung Yung 55
Lei, Yan 86
Levy, Asaf 60
Ley, Ruth E. 60
Liao, Hui-Ling (Sunny) 61, 81
Libault, Marc 79
Lidstrom, Mary E. 25
Liechty, Zachary 38, 84
Ligeti, Jorge 51
Li, Guotian 48, 60
Li, Lei 55
Lim, Linda 36–128
Lindquist, Erika 23, 68, 79
Lin, Senjie 87
Lin, Yu-Ju (Lulu) 62
Lipton, Mary S. 5
Lipzen, Anna 23, 60, 68
Liu, Haley 21
Liu, Yina 69
Li, Wen-Hsiung 62
Li, Zhou 98
Looney, Brian 81
Lopes, Mariana 54
Lopez-Fernadez, Margarita 62
Lor, Vai 63
Louie, Katherine 68, 88
Lovejoy, Connie 95
Lubys, Arvydas 40
Lu, Mei-Yeh 62
Lundberg, Derek 12, 60, 84
Luo, Chongyuan 28
Lu, Tse-Yuan 60
Lu, Xi 101
Lynch, Jonathan 5
Lynch, Michael 41

M

MacGregor, Barbara J. 63
Magnuson, Jon 68
Mahoy, Jill 63
Mäkelä, Miia R. 32
Malmstrom, Rex R. 27
Malueg, Cera Danielle 63
Mamidi, Sujan 46
M., Anandharaj 62
Marcus, Tamara S. 39
Martinell, Brian 63
Martinez, Angel T. 13
Martin, Joel 60
Mascena, Corianna J.1 40
Matheus Carnevali, Paula B. 64
Matthews, Ben 52
Mayali, Xavier 51, 83
Mayes, Melanie 98
Mazaheri, Mona M. 65
McCombie, Richard W. 3
McDonald, Meredith 27, 60
McInerney, Michael 66
McKay, Luke J. 66
McKinlay, James B. 41
McMahon, Katherine (Trina) 6
Meier-Kolthoff, Jan P. 56
Mei, Ran 66
Melillo, Jerry 11
Melo-Perez, Flor M. 82
Merchant ,Sabeeha S. 87
Michener, Joshua K. 67
Miller, Christopher S. 30, 71
Miller, Nathan 63
Miyamoto, Amy 63

Mizrahi, Itzhak 46
Mockler, Todd 46, 87
Mock, Thomas 6
Mohammed, Zehara 43
Molot, Lewis A. 92
Mondo, Stephen J. 55, 68
Monteiro, Freddy 60
Montoya, Liliam A. 68
Morales-Cruz, Abraham 28
Moran, Mary Ann 74
Morin, Tim H. 30
Moronta Barrios, Felix 94
Morrison, Elise S. 69
Mortimer, Jenny C. 48
Muchero, Wellington 23
Mucyn, Tatiana S. 26
Mueller, Ulrich G. 70
Mugnozza, Giuseppe 35
Mukherjee, Supratim 71
Mullet, John 87
Muratore, Daniel J. 75
Murúa, Pedro 14

N

Nair, Meera 84
Narihiro, Takashi 66
Narrowe, Adrienne B. 30, 71
Nattestad, Maria 28
Navid, Ali 44
Naylor, Daniel 27, 72
Neilson, T. 42
Neufeld, Josh D. 92
Neumann, Anthony P. 73
Neumann, Rebecca B. 39, 95
Newton, Ryan J. 75
Nguyen, Bao 38, 84
Nilson, Daniel J. 83
Nobu, Masaru 66
Noirot, Philippe 57
Nong, Wenyan 55, 101
Northen, Trent R. 68, 88
Norton, Jeanette M. 73
Nowinski, Brent C. 74

O

Ogram, Andrew 69
Oh, Steve 46
Oliker, L. 42
O'Malley, Michelle A. 68, 76, 88
O'Malley, Ronan 28
Opulente, Dana A. 53
Orr, Galya 31, 32
Osborne, Todd 69
Ostrowski, Martin 14
Ouyang, Yang 73
Ovchinnikova, Galina 71
Owusu-Dommey, Akosua 39

P

Pachiadaki, Maria G. 40
Palluk, Sebastian 50
Pan, Chongle 16, 98
Paredes, Sur H. 60
Parich, Alexandra 36
Parks, Donovan H. 39
Paterson, Michael J. 92
Paver, Sara F. 75
Pawlowska, Teresa E. 7
Payet, Jérôme P. 29
Peay, Kabir 61
Pelletier, Dale A. 60
Peluso, Paul 28, 52
Peng, Xuefeng (Nick) 76
Peredo, Elena L. 77
Perez Coronel, Elisabet 16
Peris, David 54
Perroud, Pierre-François 78
Petersen, Mike 63
Petronis, Arturas 40
Pett-Ridge, Jennifer 83
Petzold, Christopher J. 15
Pham, Nikki 60
Phillips, Gregory 38
Pidatala, Ramana 48
Pingault, Lise M. 79
Piotrowski, Jeff S. 54
Pires, J. Chris 8
Plott, Christopher 34, 87
Poland, Jesse 34
Poulton, Nicole J. 40
Pourmehdi, Shadi 36
Povilaitis, Tadas 40
Prather, Kristala L. J. 12
Preston, Christina 74
Preston, Jill C. 32
Pretzer, Carina 36
Prochnik, Simon E. 14, 79, 87
Props, Ruben 80
Przylucka, Agnes 24
Puckett, Sara 59
Purdom, Elizabeth 27, 72
Purvine, Samuel 31, 32, 61

Q

Qiao, Zhenzhen 79
Qualls, Morgan 87
Quandt, Alisha C. 10

R

Rahimi, Mohammad Javad 36
Ramos, Cayo 94
Ranjan, Priya 23
Rank, David R. 28
Rasmussen, Anna 41
Ray, Jayashree 35
Reddy, T.B.K. 71
Redmond, Molly C. 81
Reji, Linta 91
Rensing, Stefan A. 78, 87

Ren, Yuqi 79
Reysenbach, Anna-Louise I. 33
Rhoades, Nicholas 33
Richardson, Tammi L. 44
Rich, Virginia I. 39
Riley, Robert 55
Rivers, Adam 100
Rives, Dyllan 17
Robinson, Errol 31, 32
Rodgers-Melnick, Eli 102
Rodrigues, Andria 15
Rodriguez, William 11
Roger, Andrew J. 8
Rojas, Alejandro J. 81
Rokas, Antonis 53
Rokhsar, Daniel 20, 42, 45, 79, 85
Ronald, Pamela 48, 60
Ross, Jessica 38
Rovinskiy, Nikolay 54
Ruiz-Dueñas, Francisco J. 13
Ruiz-Font, Angelica C. 82
Rusch, Douglas B. 47
Ryan, Elizabeth M. 26

S

Saini, Avneesh 15
Salamov, Asaf 10, 55
Saleska, Scott R. 39–128
Samo, Ty 51, 83
Sanchez Barrios, Andrea 12, 84
Sanderson, Craig 14
Sandor, Laura 87
Sandoval, Luis R. 12
Santos, Christian 38
Santos-Medellin, Christian M. 84
Sarah, Lebeis 60
Sarma, Birinchi 43
Sato, Trey K. 54
Schackwitz, Wendy 60, 68
Scharffenberger, Sam 35
Schatz, Michael C. 28
Schiff, Sherry L. 92
Schmidt, Marian 80
Schmutz, Jeremy 18, 23, 34, 46, 48, 78, 79, 85, 87, 102
Scholin, Christopher 74
Schramm, Andreas 86
Schuhmacher, Rainer 36
Schulz, Frederik 8
Schwartz, Cory 96
Sedlazeck, Fritz J. 28
Seewald, Jeffrey S. 33
Senin, P. 42
Shabbir-Hussain, Murtaza 96
Shapiro, Nicole 73
Sharon, Itai 46
Shearin, Tom 84
Shen, Qirong 24
Shen, Xing-Xing 53
Sher, Andrew W. 35
Shiller, Alan M. 81
Shinde, Shalaka 57
Shulse, Christine N. 86
Shu, Shengqiang 48, 85
Siccardi, Anthony 51, 83
Sierra, Julian Muñoz 66
Siletti, Cheta E. 99
Simmons, Blake 55
Simone, Domenico 62
Simon, Holly M. 29
Singh, Anup 50
Singleton, Caitlin M. 39
Sistrom, Mark 93
Sjogaard, Kamilla S. 86
Smart, Lawrence B. 102
Smith, Chad C. 70
Smith, Garrett J. 30
Smith, Jason M. 91
Smith, Maria W. 29
Snel, Berend 10
Solden, Lindsey M. 30
Spalding, Edgar 63
Spatafora, Joseph W. 68
Sprecher, Brittany N. 87
Sreedasyam, Avinash 46, 78, 87
Stacey, Gary 78, 87
Stajich, Jason E. 68
Stamatis, Dimitrios 71
Stefanik, Kay C. 71
Stepanauskas, Ramunas X. 40
Stevenson, Dennis 3
Stillman, Kyra 21, 60
Strittmatter, Martina 14
Suen, Garret 56, 73
Sullivan, William 68
Sundaresan, Venkatesan 38, 84
Swift, Candice L. 76, 88

T

Talbot, Jennifer M. 61, 99
Taneja, Jyoti 89
Tang, Yuhong 87
Tarn, Ann H. 90
Tarn, Jonathan 90
Tarver, Angela 54
Taylor, John 61, 68
Thelen, Michael 51
Theodorou, Michael K. 76
Thomas, Brian C. 64
Thompson, Brian P. 40–128
Tolar, Bradley B. 90, 91
Tollerson, Rodney 39–128
Torres-Jerez, Ivone 87
Tournay, Robert 35
Treusch, Alexander H. 86
Tringe, Susannah G. 14, 15, 16, 45, 60, 70, 98, 100
Tsang, Karen S. W. 92

Tseng, Elizabeth 26
Tsuji, Jackson M. 92
Tupper, Ben 40–128
Turner, Benjamin L. 98
Tuskan, Gerald A. 23, 87, 102
Tyson, Gene W. 39–128

U

Udvardi, Michael 87
Ul-Hasan, Sabah 26, 93

V

Vaillancourt, Brieanne 65
Valentine, David L. 76, 90
Vanderploeg, Henry 80
van Diepen, Linda 11
van Lier, Jules 66
Varala, Kranthi 3
Varner, Ruth K. 39
Venkiteswaran, Jason J. 92
Venturi, Vittorio 94
Vilgalys, Rytas 61, 81
Visel, Axel 27, 68
Voigt, Christopher A. 12
Von Wettberg, Eric 43
Vosshall, Leslie 52

W

Waldo, Nicholas B. 95
Walsh, David A. 95
Wang, Gaoyan 26, 86
Wang, Kai 19
Wang, Mei 87
Wang, Shuwen 34
Ward, Nicholas 69
Waters, Robert J. 26
Weber, Peter K. 51, 83
Wegrzyn, Renee 9
Wei, Tong 60
Wen-Tso Liu 66
Weston, David J. 23
Wetmore, Kelly M. 26
Wheeldon, Ian R. 96
Whitman, William B. 74, 96
Wiebenga, Ad 10, 31
Wik, Martin 39–128
Wildermuth, Mary 89
Wilkins, Michael J. 30
Williams, Kenneth H. 64
Wong, Chong Kim 92
Wong, Man Chun 55
Woodcroft, Ben J. 39–128
Woods, Daniel P. 32
Woyke, Tanja 10, 60, 86, 93, 96
Wright, Joseph S. 98
Wrighton, Kelly C. 30, 71
Wrobel, Russell L. 54
Wu, Lingling 92
Wu, Xiaofen 62
Wu, Zhuoying 66

X

Xiao, Yang 55
Xie, Yichun 97
Xing, Lei 101

Y

Yang, Song 98
Yang, Xiaohan 23
Yang, Yongil 23
Yao, Qiuming 16, 98
Yelick, Kathy 42
Ye, Lin 66
Ye, Rongzhong 45
Yoshikuni, Yasuo 27, 86
Young, Eric M. 12
Yuet Ting, Lau 99
Yu, Ke 59
Yu, Xiaolan 48
Yu, Zheng 25

Z

Zane, Matthew 87
Zargar, Kamrun 15
Zeiner, Carolyn A. 99
Zerbe, Philipp 9
Zerbs, Sarah 57
Zettel, Herbert 36
Zhang, Qingpeng 100
Zhang, Xiaofei 34
Zhong, Yiyi 101
Zhou, Lijie 59
Zhou, Ran 102
Zhou, Xiaofan 53
Zhou, Y.-G 96
Zink, Erika 31, 32

Attendees

Adam, Catherine
DOE Joint Genome Institute
CAdam@lbl.gov

Adams, Catharine
UC Berkeley
catadams@berkeley.edu

Agosto, Carolyn
DOE Joint Genome Institute
carolynagosto@lbl.gov

Aguilar, Pontes Maria Victoria
CBS-KNAW Fungal Biodiversity Centre
v.aguilar@cbs.knaw.nl

Ahrendt, Steven
DOE Joint Genome Institute
sahrendt0@lbl.gov

Akiva, Eyal
UC San Francisco
eyal.akiva@ucsf.edu

Alessi, Joe
Illumina
JAlessi@illumina.com

Alexander, Harriet
UC Davis
harriet.xander@gmail.com

Allen, Benjamin
Oak Ridge National Lab
allenbh@ornl.gov

Alteio, Lauren
DOE Joint Genome Institute
lalteio@umass.edu

Amos, Kirtley
University of Kentucky
bkamos3@g.uky.edu

Anantharaman, Karthik
UC Berkeley
karthik.a@berkeley.edu

Andeer, Peter
Berkeley Lab
pfandeer@lbl.gov

Andersen, Mikael
Denmark Technical University
mr@bio.dtu.dk

Anderson, Lisa
MIT
laanders@mit.edu

Anderson, Todd
U.S. Department of Energy
todd.anderson@science.doe.gov

Andreopoulos, Bill
DOE Joint Genome Institute
wandreopoulos@lbl.gov

Andrews, Brenda
University of Toronto
brenda.andrews@utoronto.ca

Attwood, Graeme
AgResearch
graeme.attwood@agresearch.co.nz

Babbit, Patsy
UC San Francisco
babbitt@cgl.ucsf.edu

Backman, Tyler
Berkeley Lab
tbackman@lbl.gov

Badis, Yacine
Scottish Association for Marine Science
Yacine.Badis@sams.ac.uk

Baldrian, Petr
Institute of Microbiology, Prague
baldrian@biomed.cas.cz

Ballon, Massie
DOE Joint Genome Institute
mlballon@lbl.gov

Barry, Kerrie
DOE Joint Genome Institute
kerrieb@gmail.com

Bartley, Laura
University of Oklahoma
lbartley@ou.edu

Bauer, Diane
DOE Joint Genome Institute
dmbauer@lbl.gov

Beller, Harry
Berkeley Lab
hrbeller@lbl.gov

Beman, J. Michael
UC Merced
mbeman@ucmerced.edu

Bhattacharya, Debashish
Rutgers University
debash.bhattacharya@gmail.com

Bhaya, Devaki
National Science Foundation
dbhaya@nsf.gov

Biswas, Abhishek
Oak Ridge National Lab
biswasa@ornl.gov

Blaby, Crysten
BNL
cblaby@bnl.gov

Blazewicz, Steve
Lawrence Livermore National Lab
blazewicz1@llnl.gov

Blenner, Mark
Clemson University
blenner@clemson.edu

Blouin, Nic
University of Wyoming
nblouin@uwyo.edu

Boeuf, Dominique
University of Hawai'i at Mānoa
dboeuf@hawaii.edu

Bonito, Gregory
Michigan State University
bonito@msu.edu

Bowen, Ben
Berkeley Lab
bpbowen@lbl.gov

Bowers, Robert
DOE Joint Genome Institute
rmbowers@lbl.gov

Brawley, Susan
University of Maine
brawley@maine.edu

Briscoe, Leah
UC Los Angeles
lbriscoep81@ucla.edu

Brocchieri, Luciano
University of Florida
lucianob@ufl.edu

Brodie, Juliet
Natural History Museum, UK
j.brodie@nhm.ac.uk

Brooks, Lauren
UC Merced
lbrooks3@ucmerced.edu

Brown, Natasha
DOE Joint Genome Institute
nzvenigorodsky@lbl.gov

Brown, Steven
Oak Ridge National Lab
brownsd@ornl.gov

Brown, Chris
UC Berkeley
ctb@berkeley.edu

Brown, C. Titus
UC Davis
ctbrown@ucdavis.edu

Browne, Daniel
Texas A&M University
dbrowne.up@gmail.com

Bruckbauer, Steven
University of Wisconsin - Madison
steven.bruckbauer@gmail.com

Brutnell, Tom
Donald Danforth Plant Science Center
tbrutnell@danforthcenter.org

Calhoun, Sara
DOE Joint Genome Institute
secalhoun@lbl.gov

Campbell, Barbara
Clemson University
bcampb7@clemson.edu

Campbell, Ashley
Lawrence Livermore National Lab
campbell87@llnl.gov

Cao, Yuanyu
University of New Hampshire
yov2@wildcats.unh.edu

Cardarelli, Emily
Stanford University
ecardare@stanford.edu

Carlson, Joe
DOE Joint Genome Institute
JWCarlson@lbl.gov

Cartwright, Amy
DOE Joint Genome Institute
amyflyone@gmail.com

Carver, Akiko
DOE Joint Genome Institute
acarver@lbl.gov

Catcheside, David
Flinders University
David.Catcheside@flinders.edu.au

Cavalier, David
Great Lakes Bioenergy Center
cavalie@msu.edu

Chakraborty, Shreyasee
Pacific Biosciences
Schakraborty@pacificbiosciences.com

Chamarande, Jordan
DOE Joint Genome Institute
jordan.chamarande@agrosupdijon.fr

Chan, Steve
Berkeley Lab
sychan@lbl.gov

Chan, Po Lam
The Chinese University of Hong Kong
KATHYPOLAM@HOTMAIL.COM

Chang, Jinhui
The Chinese University of Hong Kong
tjchangjinhui@gmail.com

Chen, Cindy
DOE Joint Genome Institute
cindychen@lbl.gov

Chen, Jay
Oak Ridge National Lab
chenj@ornl.gov

Cheng, Jan-Fang
DOE Joint Genome Institute
jfcheng@lbl.gov

Chenthamara, Komal
TU WIEN
komal.chenthamara@tuwien.ac.at

Chhatre, Vikram
University of Wyoming
vchhatre@uwyo.edu

Chiniquy, Dawn
DOE Joint Genome Institute
dmchiniquy@lbl.gov

Chistoserdova, Ludmila
University of Washington
milachis@uw.edu

Chovatia, Mansi
DOE Joint Genome Institute
mrchovatia@lbl.gov

Church, Deanna
10x Genomics
deanna.church@10xgenomics.com

Ciobanu, Doina
DOE Joint Genome Institute
dgciobanu@lbl.gov

Clark, Tyson
Pacific Biosciences
tclark@pacificbiosciences.com

Cohen, Lisa
UC Davis
ljcohen@ucdavis.edu

Cole, Jason
DOE Joint Genome Institute
jason.t.cole@lbl.gov

Cole, Benjamin
DOE-Joint Genome Institute
bjcole@lbl.gov

Coleman, Jeff
Auburn University
jjcoleman@auburn.edu

Coleman-Derr, Devin
UC Berkeley
colemanderr@berkeley.edu

Concepcion, Gregory
Pacific Biosciences
gconcepcion@pacificbiosciences.com

Connor, Elise
University of Texas at Austin
eworchel@utexas.edu

Cornelius, Joe
ARPA-E
joe.cornelius@hq.doe.gov

Coruzzi, Gloria
New York University
gloria.coruzzi@nyu.edu

Cottingham, Robert
Oak Ridge National Lab
cottinghamrw@ornl.gov

Couradeau, Estelle
DOE Joint Genome Institute
ecouradeau@lbl.gov

Cox, Michael
University of Wisconsin - Madison
cox@biochem.wisc.edu

Crowe, Sean
University of British Columbia
sean.crowe@ubc.ca

Crowell, Chase
Cornell University
crc255@cornell.edu

Crump, Byron
Oregon State University
bcrump@coas.oregonstate.edu

Cullen, Daniel
USDA Forest Products Laboratory
daniel.cullen@wisc.edu

Cullings, Ken
NASA-Ames
cullings1@earthlink.net

Cushman, John
University of Nevada, Reno
jcushman@unr.edu

Daly, Rebecca
Ohio State University
daly.130@osu.edu

Daskalov, Asen
UC Berkeley
asen.daskalov@berkeley.edu

Daum, Chris
DOE Joint Genome Institute
cgdaum@lbl.gov

Davidson, Patrick
DOE Joint Genome Institute
padavidson@lbl.gov

de Lima Matos, Juliana
UC Berkeley
jmatos@berkeley.edu

De Paoli, Henrique
DOE Joint Genome Institute
hcdepaoli@lbl.gov

de Vries, Ronald
Westerdijk Fungal Biodiversity Institute
r.devries@cbs.knaw.nl

DeGraaf, Stephanie
UC Berkeley
sldegraaf@gmail.com

Deshpande, Shweta
DOE Joint Genome Institute
sdeshpande@lbl.gov

Desiro', Alessandro
Michigan State University
adesiro@msu.edu

Desmarais, David
Harvard University
desmarais@fas.harvard.edu

D'haeseleer, Patrik
Lawrence Livermore National Lab
patrikd@gmail.com

Diamond, Spencer
UC Berkeley
sdiamond@berkeley.edu

Dijkstra, Paul
Northern Arizona Univerisity
paul.dijkstra@nau.edu

Dippel, Andrew
UC Berkeley
adippel@berkeley.edu

Dockter, Rhyan
DOE Joint Genome Institute
rbdockter@lbl.gov

Doktycz, Mitchel
Oak Ridge National Lab
doktyczmj@ornl.gov

Donaho, John
Portland State University
jdonaho@pdx.edu

Dopson, Mark
Linnaeus University
mark.dopson@lnu.se

Dorn, Kevin
Kansas State University
dorn@ksu.edu

Doty, Sharon
University of Washington
sldoty@uw.edu

Doud, Devin
DOE Joint Genome Institute
dfdoud@lbl.gov

Dove, Nicholas
UC Merced
ndove@ucmerced.edu

Drake, Meghan
Oak Ridge National Lab
drakemm@ornl.gov

Drell, Daniel
U.S. Department of Energy
Daniel.Drell@science.doe.gov

Druzhinina, Irina
TU WIEN
irina.druzhinina@tuwien.ac.at

Duffy, Kecia
DOE Joint Genome Institute
kmduffy@lbl.gov

Edwards, Joseph
UC Davis
edwards@ucdavis.edu

Egan, Rob
DOE Joint Genome Institute
RSEgan@lbl.gov

Egerton-Warburton, Louise
Chicago Botanic Garden
lwarburton@chicagobotanic.org

Eloe-Fadrosh, Emiley
DOE Joint Genome Institute
eaeloefadrosh@lbl.gov

Erb, Tobi
Max Planck Institute
toerb@mpi-marburg.mpg.de

Erbilgin, Onur
DOE Joint Genome Institute
oerbilgin@lbl.gov

Eshel, Gil
New York University
ge30@nyu.edu

Espinoza, Josh
J. Craig Venter Institute
jespinoz@jcvi.org

Evans, Robert
DOE Joint Genome Institute
rse@lbl.gov

Fagnan, Kjiersten
Berkeley Lab
kmfagnan@lbl.gov

Fergusson, Elizabeth
Bigelow Laboratory for Ocean Sciences
efergusson@bigelow.org

Fero, Michael
TeselaGen Biotechnology, Inc.
mike.fero@gmail.com

Floge, Sheri
Ohio State University
floge.1@osu.edu

Flood, Beverly
University of Minnesota
beflood@umn.edu

Forrer, Mark
Agile BioFoundry / Joint BioEnergy Institute
mark.forrer@lbl.gov

Foster, Bryce
DOE Joint Genome Institute
brycefoster@lbl.gov

Fox, Brian
University of Wisconsin - Madison
bgfox@biochem.wisc.edu

Francis, Christopher
Stanford University
caf@stanford.edu

Frank, Carolin
UC Merced
cfrank3@ucmerced.edu

Freedman, Zachary
West Virginia University
zachary.freedman@mail.wvu.edu

Fritts, Ryan
Indiana University Bloomington
rkfritts@indiana.edu

Fuchs, Bernhard
Max Planck Institute
bfuchs@mpi-bremen.de

Gabriel, Raphael
Joint BioEnergy Institute
raphaelgabriel@lbl.gov

Gach, Philip
Joint Bioenergy Institute
pcgach@lbl.gov

Galazka, Jonathan
NASA-ARC
jonathan.m.galazka@nasa.gov

Gao, Jian
Berkeley Lab
jiangao@lbl.gov

Gifford, Scott
UNC Chapel Hill
sgifford@email.unc.edu

Gilbert, David
DOE Joint Genome Institute
degilbert@lbl.gov

Glavina del Rio, Tijana
DOE Joint Genome Institute
tglavina@lbl.gov

Goode, Jason
Illumina
jgoode@illumina.com

Goodstein, David
DOE Joint Genome Institute
dmgoodstein@lbl.gov

Gordon, Sean
DOE Joint Genome Institute
seangordon07@gmail.com

Goudeau, Danielle
DOE Joint Genome Institute
dgoudeau@lbl.gov

Gouran, Hossein
DOE Joint Genome Institute
hgouran@lbl.gov

Graham, Danielle
DOE Joint Genome Institute
dgraham@lbl.gov

Granitsiotis, Michael S.
DOE Joint Genome Institute
msg@lbl.gov

Gravuer, Kelly
Arizona State University
klgravuer@ucdavis.edu

Green, Beverley
University of British Columbia
brgreen@mail.ubc.ca

Greenspan, Alex
UC Davis
greenspan@ucdavis.edu

Grettenberger, Christy
UC Davis
christygrettenberger@gmail.com

Griesemer, Marc
Lawrence Livermore National Lab
griesemer1@llnl.gov

Grigoriev, Igor
DOE Joint Genome Institute
ivgrigoriev@lbl.gov

Grossman, Arthur
Carnegie Institution for Science
arthurg@stanford.edu

Guay, Lisa
Massachusetts Institute of Technology
guayl@mit.edu

Haas, Fabian
University of Marburg
fabian.haas@biologie.uni-marburg.de

Hall, Richard
Pacific Biosciences
rhall@pacificbiosciences.com

Hansel, Colleen
Woods Hole Oceanographic Institution
chansel@whoi.edu

Haridas, Sajeet
DOE Joint Genome Institute
sharidas@lbl.gov

Harmon-Smith, Miranda
DOE Joint Genome Institute
mlharmon-smith@lbl.gov

Harris, Steven
University of Nebraska
sharris2@unl.edu

Hartman, Wyatt
DOE Joint Genome Institute
whhartman@lbl.gov

Hatas, Emily
Pacific Biosciences
ehatas@pacb.com

Hayes, Richard
DOE Joint Genome Institute
rdhayes@lbl.gov

He, Guifen
DOE Joint Genome Institute
ghe@lbl.gov

Healey, Adam
HudsonAlpha Institute
ahealey@hudsonalpha.org

Hedlund, Brian
University of Nevada, Las Vegas
brian.hedlund@unlv.edu

Heiner, Cheryl
Pacific Biosciences
cheiner@pacificbiosciences.com

Hentrich, Klaus
TTP Labtech
klaus.hentrich@ttplabtech.com

Hess, Matthias
UC Davis
mhess@ucdavis.edu

Hill, Tom
Colorado State University
Thomas.Hill@colostate.edu

Hill, Ryan
DOE Joint Genome Institute
rhill@lbl.gov

Hirsch, Ann
UC Los Angeles
ahirsch@ucla.edu

Hoi Shan, Kwan
Chinese University of Hong Kong
hoishankwan@cuhk.edu.hk

Hoover, Cindi
DOE Joint Genome Institute
cahoover@lbl.gov

Huang, Pu
Donald Danforth Plant Science Center
phuang@danforthcenter.org

Hudson, Matthew
University of Illinois
mhudson@illinois.edu

Hundley, Hope
DOE Joint Genome Institute
HNTice@lbl.gov

Huntemann, Marcel
DOE Joint Genome Institute
mhuntemann@lbl.gov

Inskeep, William
Montana State University
binskeep@montana.edu

Ivanova, Natalia
DOE Joint Genome Institute
nnivanova@lbl.gov

Iwai, Kosuke
Joint BioEnergy Institute
KosukeIwai@lbl.gov

Jacobson, Daniel
Oak Ridge National Lab
jacobsonda@ornl.gov

Jain, Rashmi
UC Davis
rsjain@ucdavis.edu

Jarett, Jessica
DOE Joint Genome Institute
jkjarett@lbl.gov

Jenkins, Jerry
HudsonAlpha Institute
jjenkins@hudsonalpha.org

Jing, Beibei
Joint BioEnergy Institute
bjing@lbl.gov

Joachimiak, Andrzej
Argonne National Laboratory
andrzejj@anl.gov

Johnson, Jenifer
DOE Joint Genome Institute
jkaplan@lbl.gov

Johnson, Matthew
Woods Hole Oceanographic Institution
mattjohnson@whoi.edu

Joubert, Pierre
University of Washington
pierrj@uw.edu

Juneja, Punita
Bayer
punita.juneja@bayer.com

Jungbluth, Sean
DOE Joint Genome Institute
sjungbluth@lbl.gov

Kang, Don
DOE Joint Genome Institute
ddkang@lbl.gov

Kang, Julie
DOE Joint Genome Institute
jjk@lbl.gov

Karaoz, Ulas
Berkeley Lab
ukaraoz@lbl.gov

Karpowicz, Steven
Tau Biosciences
skarp@chem.ucla.edu

Katz, Andrew
DOE Joint Genome Institute
andrewkatz@lbl.gov

Ke, Jing
DOE Joint Genome Institute
jingke@lbl.gov

Kegg, Lisa
DOE Joint Genome Institute
lrkegg@lbl.gov

Kennedy, Megan
DOE Joint Genome Institute
mckennedy@lbl.gov

Khadempour, Lily
University of Wisconsin - Madison
khadempour@wisc.edu

Kim, Peter
Joint BioEnergy Institute
pwkim@sandia.gov

Kim, Jinwoo
UC San Diego
jik152@eng.ucsd.edu

Kim, Jay
UC Santa Cruz
jay.wj.kim@gmail.com

Kimbrel, Jeff
Lawrence Livermore National Lab
kimbrel1@llnl.gov

Kingan, Sarah
Pacific Biosciences
skingan@pacb.com

Kirton, Edward
DOE Joint Genome Institute
ESKIRTON@LBL.GOV

Klassen, Jonathan
University of Connecticut
jonathan.klassen@uconn.edu

Klein, Anita
University of New Hampshire
anita.klein@unh.edu

Kohler, Annegret
INRA Nancy-Lorraine
annegret.kohler@inra.fr

Kominek, Jacek
University of Wisconsin - Madison
jkominek@wisc.edu

Kosina, Suzanne
DOE Joint Genome Institute
smkosina@lbl.gov

Krause, David
University of Wisconsin
dkrause2@wisc.edu

Kresovich, Stephen
Clemson University
skresov@clemson.edu

Kuftin, Andrea
DOE Joint Genome Institute
ankuftin@lbl.gov

Kuo, Alan
DOE Joint Genome Institute
akuo@lbl.gov

Kuo, Rita
DOE Joint Genome Institute
rckuo@lbl.gov

Kyrpides, Nikos
DOE Joint Genome Institute
nckyrpides@lbl.gov

La Reau, Alex
University of Wisconsin - Madison
alareau@wisc.edu

LaButti, Kurt
DOE Joint Genome Institute
klabutti@lbl.gov

Lail, Kathleen
DOE Joint Genome Institute
klail@lbl.gov

Larsen, Peter
Argonne National Laboratory
plarsen@anl.gov

Lau, Rebecca
Berkeley Lab
rlau@lbl.gov

Lau, Amy
The Chinese University of Hong Kong
1155032273@link.cuhk.edu.hk

Laudencia-Chingcuanco, Debbie
USDA-ARS, Western Regional Research Center
debbie.laudencia@ars.usda.gov

Lawrence, Jennifer
UC Berkeley
jelawrence@berkeley.edu

Leach, Jan
Colorado State University
Jan.Leach@ColoState.edu

LeBoldus, Jared
Oregon State University
jared.leboldus@science.oregonstate.edu

Lee, Janey
DOE Joint Genome Institute
jlee2@lbl.gov

Lee, Juna
DOE Joint Genome Institute
jel@lbl.gov

Leebens-Mack, Jim
University of Georgia
jleebensmack@uga.edu

Lemaux, Peggy
UC Berkeley
lemauxpg@berkeley.edu

Levdokymenko, Kateryna
University of British Columbia
ievdokymenko@gmail.com

Levy, Asaf
DOE Joint Genome Institute
alevy@lbl.gov

Lewald, Kyle
DOE Joint Genome Institute
kmlewald@lbl.gov

Lewis, Jared
University of Chicago
jaredlewis@uchicago.edu

Li, Guotian
Joint BioEnergy Institute
guotianli@lbl.gov

Liang, Liyuan
Pacific Northwest National Laboratory
liyuan.liang@pnnl.gov

Liao, Junhui
UIUC
jliao12@illinois.edu

Liao, Hui-Ling
University of Florida
sunny.liao@ufl.edu

Lin, Yu-Ju
Academia Sinic
superolulu@gmail.com

Lin, Junyan
DOE Joint Genome Institute
junyanlin@lbl.gov

Lipton, Mary
Pacific Northwest National Laboratory
mary.lipton@pnnl.gov

Lipzen, Anna
DOE Joint Genome Institute
alipzen@lbl.gov

Liu, Lifeng
DOE Joint Genome Institute
lifeng.liu@lbl.gov

Lopez Garcia de Lomana, Adrian
Institute for Systems Biology
alomana@systemsbiology.org

Lopez Meza, Sarynna
Great Lakes Bioenergy Center
lopezsa5@msu.edu

Lopez Peredo, Elena
Marine Biological Laboratory
elperedo@mbl.edu

Lopez-Fernandez, Margarita
Linnaeus University
margarita.lopezfernandez@lnu.se

Louie, Katherine
DOE Joint Genome Institute
kblouie@lbl.gov

Lovejoy, Connie
Laval University
Connie.lovejoy@bio.ulaval.ca

Lu, Mei-Yeh
High Throughput Genomics Core,
Academia Sinica, Taiwan
meiyehlu@sinica.edu.tw

Lynagh, Peter
UC Davis
pglynagh@ucdavis.edu

Lynch, Jonathan
Penn State University
jpl4@psu.edu

Macaya-Sanz, David
West Virginia University
david.macayasanz@mail.wvu.edu

MacGregor, Barbara
Marine Sciences Dept., UNC - Chapel Hill
bmacgreg@unc.edu

Madera, Mary
UC Berkeley
mary.madera2@gmail.com

Makhalanyane, Thulani
University of Pretoria
thulani.makhalanyane@up.ac.za

Malmstrom, Rex
DOE Joint Genome Institute
rrmalmstrom@lbl.gov

Malueg, Cera
University of Wisconsin - Madison
cdmalueg@wisc.edu

Marcus, Tamara
University of New Hampshire
tsm1004@wildcats.unh.edu

Mars, Kristin
PacificBiosciences
kmars@pacb.com

Martin Joel
DOE Joint Genome Institute
j_martin@lbl.gov

Martin, Francis
INRA
francis.martin@inra.fr

Martin Gracie
UDSA ARS
graciebensonmartin@gmail.com
Mascena, Corianna
Bigelow Laboratory for Ocean Sciences
cmascena@bigelow.org

Matheus Carnevali, Paula
UC Berkeley
pmatheus@berkeley.edu

Mayali, Xavier
Lawrence Livermore National Lab
mayali1@llnl.gov

Mazaheri, Mona
University of Wisconsin
mmazaheri@wisc.edu

McKain, Michael
Donald Danforth Plant Science Center
mrmckain@gmail.com

McKay, Luke
Montana State University
luke.mckay@montana.edu

McMahon, Trina
University of Wisconsin
kdmcmahon@wisc.edu

Mei, Ran
University of Illinois at Urbana-Champaign
ranmei2@illinois.edu

Meng, John
DOE Joint Genome Institute
xianweimeng@lbl.gov

Meng, Xiandong
DOE Joint Genome Institute
xiandongmeng@lbl.gov

Meredith, Laura
University of Arizona
laurameredith@email.arizona.edu

Michener, Joshua
Oak Ridge National Lab
michenerjk@ornl.gov

Mihaltcheva, Sirma
DOE Joint Genome Institute
smihaltcheva@lbl.gov

Mock, Thomas
U East Anglia
T.Mock@uea.ac.uk

Mondo, Stephen
DOE Joint Genome Institute
sjmondo@lbl.gov

Montoya, Liliam
UC Berkeley
liliam_m007@berkeley.edu

Morrison, Elise
University of Florida
emorrison@ufl.edu

Mou, Xiaozhen (Jen)
Kent State University
xmou@kent.edu

Mouncey, Nigel
DOE Joint Genome Institute
nmouncey@lbl.gov

Mueller, Ulrich
University of Texas at Austin
umueller@austin.utexas.edu

Mukherjee, Supratim
DOE Joint Genome Institute
supratimmukherjee@lbl.gov

Na, Hyunsoo
DOE Joint Genome Institute
hna@lbl.gov

Nagarajan, Harish
Genomatica
hnagarajan@genomatica.com

Narrowe, Adrienne
University of Colorado Denver
adrienne.narrowe@ucdenver.edu

Nath, Sangeeta
DOE Joint Genome Institute
snath@lbl.gov

Nath, Nandita
DOE Joint Genome Institute
nnath@lbl.gov

Navid, Ali
Lawrence Livermore National Lab
navid1@llnl.gov

Naylor, Daniel
UC Berkeley
dtnaylor@berkeley.edu

Necula, Simona
DOE Joint Genome Institute
sfnecula@lbl.gov

Neumann, Anthony
University of Wisconsin - Madison
apneumann2@wisc.edu

Ng, Vivian
DOE Joint Genome Institute
vng@lbl.gov

Nielsen, Torben
DOE Joint Genome Institute
torben@lbl.gov

Niyogi, Krishna
UC Berkeley
niyogi@berkeley.edu

Norman, Dirk
Wisconsin Energy Institute/
Great Lakes Bioenergy Research Center
dirk.norman@wisc.edu

Northen, Trent
DOE Joint Genome Institute
trnorthen@lbl.gov

Northrup, Dan
Booz Allen
northrup_daniel@bah.com

Norton, Jeanette
Utah State University
jeanette.norton@usu.edu

Nowinski, Brent
University of Georgia
nowinski@uga.edu

Nuccio, Erin
Lawrence Livermore National Lab
nuccio1@llnl.gov

Olsen, Christian
Pacific Biosciences
colsen@pacificbiosciences.com

O'Malley, Ronan
DOE Joint Genome Institute
romalley@lbl.gov

Osborne, Jeff
Manchester University
jposborne@manchester.edu
Otani, Hiroshi
DOE Joint Genome Institute
hotani@lbl.gov

Owiti, Judith
UC Berkeley
jowiti@berkeley.edu

Pace, Norman
University of Colorado
Norman.Pace@colorado.edu

Paez-Espino, David
DOE Joint Genome Institute
adpaezespino@lbl.gov

Pan, Chongle
Oak Ridge National Lab
panc@ornl.gov

Pangilinan, Jasmyn
DOE Joint Genome Institute
jlpangilinan@lbl.gov

Pasa-Tolic, Liliam
Pacific Northwest National Laboratory
ljiljana.pasatolic@pnnl.gov

Paver, Sara
University of Chicago
sara.paver@gmail.com

Pawlowska, Teresa
Cornell University
tep8@cornell.edu

Pelletier, Dale
Oak Ridge National Lab
pelletierda@ornl.gov

Peluso, Paul
Pacific Biosciences
ppeluso@pacificbiosciences.com

Peng, Ze
DOE Joint Genome Institute
zpeng@lbl.gov

Peng, Yi
DOE Joint Genome Institute
ypeng@lbl.gov

Peng, Xuefeng (Nick)
UC Santa Barbara
xpeng@engineering.ucsb.edu

Pennacchio, Len
DOE Joint Genome Institute
LAPennacchio@lbl.gov

Perrier, Rene
DOE Joint Genome Institute
raperrier@lbl.gov

Pett-Ridge, Jennifer
Lawrence Livermore National Lab
pettridge2@llnl.gov

Pham, Jonathan
Novozymes, Inc.
jph@novozymes.com

Pick, Thea
Joint BioEnergy Institute
trpick@lbl.gov

Pingault, Lise
University of Oklahoma
lise.pingault@ou.edu

Pires, Chris
University of Missouri
piresjc@missouri.edu

Poorten, Thomas
UC Davis
tpoorten@ucdavis.edu

Props, Ruben
University of Michigan
Ruben.Props@ugent.be

Raje, Manasi
Joint BioEnergy Institute
manasiraje@lbl.gov

Ran, Zhou
West Virginia University
razhou@mix.wvu.edu

Ranjan, Priya
University of Tennessee
ranjanp@ornl.gov

Rank, David
Pacific Biosciences
drank@pacificbiosciences.com

Reddy, Tatiparthi
DOE Joint Genome Institute
tbreddy@lbl.gov

Redmond, Molly
University of North Carolina Charlotte
mcredmond@uncc.edu

Reiss, Rebecca
New Mexico Tech
Rebecca.Reiss@nmt.edu

Reysenbach, Anna-Louise
Portland State University
bwar@pdx.edu

Riley, Robert
DOE Joint Genome Institute
rwriley@lbl.gov

Rodriguez, David
Thermo Fisher Scientific
david.rodriguez2@thermofisher.com

Roger, Andrew
Dalhousie
andrew.roger@dal.ca

Rojas, Alejandro
Duke University
alejo.rojas@duke.edu

Rokhsar, Dan
DOE Joint Genome Institute
dsrokhsar@lbl.gov

Romero, Juan
Thermo Fisher Scientific
juan.romero@thermofisher.com

Roux, Simon
DOE Joint Genome Institute
sroux@lbl.gov

Ruiz-Font, Angelica
National Polithecnic Institute/MEX
convenios22@hotmail.com

Rusch, Douglas
Center for Genomics and Bioinformatics
drusch@indiana.edu

Ryan, Elizabeth
DOE Joint Genome Institute
eryan@lbl.gov

Salamov, Asaf
DOE Joint Genome Institute
aasalamov@lbl.gov

Samo, Ty
Lawrence Livermore National Lab
samo1@llnl.gov

Sanchez Barrios, Andrea
University of Kentucky
asanchez-barrios@uky.edu

Sandor, Laura
DOE Joint Genome Institute
LCSandor@lbl.gov

Santos, Christian
UC Davis
cmsantosm@ucdavis.edu

Schachtman, Daniel
University of Nebraska Lincoln
daniel.schachtman@unl.edu

Schlapfer, Joelle
DOE Joint Genome Institute
jschlapfer@lbl.gov

Schmidt, Jennifer
UC Davis
jenschmidt@ucdavis.edu

Schmutz, Jeremy
HudsonAlpha Institute
jschmutz@hudsonalpha.org

Schulz, Frederik
DOE Joint Genome Institute
fschulz@lbl.gov

Sevim, Volkan
DOE Joint Genome Institute
vsevim@lbl.gov

Shahab, Atif
DOE Joint Genome Institute
ashahab@lbl.gov

Shapiro, Nicole
DOE Joint Genome Institute
nrshapiro@lbl.gov

Sharma, Aditi
DOE Joint Genome Institute
aditisharma@lbl.gov

Sheik, Cody
University of Minnesota Duluth
cssheik@d.umn.edu

Shu, Shengqiang
DOE Joint Genome Institute
sqshu@lbl.gov

Shulse, Christine
DOE Joint Genome Institute
cnshulse@lbl.gov

Frederik Schulz
DOE Joint Genome Institute
fschulz@lbl.gov

Sieber, Christian
DOE Joint Genome Institute
csieber@lbl.gov

Simirenko, Lisa
DOE Joint Genome Institute
lsimirenko@lbl.gov

Simister, Rachel
University of British Columbia
rlsimister@gmail.com

Simone, Domenico
Linnaeus University
domenico.simone@lnu.se

Singan, Vasanth
DOE Joint Genome Institute
VRSingan@lbl.gov

Singer, Esther
DOE Joint Genome Institute
esinger@lbl.gov

Singh, Kanwar
DOE Joint Genome Institute
ksingh@lbl.gov

Sistrom, Mark
UC Merced
msistrom@ucmerced.edu

Sjøgaard, Kamilla
University of Southern Denmark
kamillasjogaard@biology.sdu.dk

Smart, Larry
Cornell University
lbs33@cornell.edu

Smirnova, Tatyana
DOE Joint Genome Institute
tsmirnova@lbl.gov

Smit, Mariya
Oregon Health and Science University
smitm@ohsu.edu

Soares, Zaramela Livia
UC San Diego
lzaramela@ucsd.edu

Song, Zewei
University of Minnesota
songzewei@outlook.com

Spittle, Kristi
PacificBiosciences
kspittle@pacificbiosciences.com

Sprecher, Brittany
University of Connecticut
brittany.sprecher@uconn.edu

Sreedasyam, Avinash
HudsonAlpha Institute
asreedasyam@hudsonalpha.org

Starr, Evan
UC Berkeley
starr.evan@berkeley.edu

Stoddard, Colby
QS2 / DARPA BTO
colby.stoddard@qs-2.com

Stuart, Rhona
Lawrence Livermore National Lab
stuart25@llnl.gov

Swenson, Tami
Berkeley Lab
tlswenson@lbl.gov

Swift Candice
UC Santa Barbara
cswift@umail.ucsb.edu

Taghavi, Safiyh
Gusto Global
safiyh.taghavi@gmail.com

Taneja, Jyoti
UC Berkeley
jyotitaneja@berkeley.edu

Tarn, Jonathan
UC Santa Barbara
jon.sy.tarn@gmail.com

Tatum, Tootie
DOE Joint Genome Institute
oltatum@lbl.gov

Taylor, John
UC Berkeley
jtaylor@berkeley.edu

Thomas, Scott
University of Nevada, Las Vegas
thomass3@unlv.nevada.edu

Tolar, Bradley
Stanford University
btolar1@stanford.edu

Tom, Lauren
Joint BioEnergy Institute
ltom@lbl.gov

Tomatsu, Andy
DOE Joint Genome Institute
atomatsu@lbl.gov

Treseder, Kathleen
UC Irvine
treseder@uci.edu

Tringe, Susannah
DOE Joint Genome Institute
sgtringe@lbl.gov

Trivedi, Pankaj
Colorado State University
pankaj.trivedi@colostate.edu

Tsang, Karen
The Chinese University of Hong Kong
kswtsang@hotmail.com

Tschaplinski, Timothy
Oak Ridge National Lab
tschaplinstj@ornl.gov

Tseng, Elizabeth
Pacific Biosciences
etseng@pacb.com

Tsuji, Jackson
University of Waterloo
jmtsuji@uwaterloo.ca

Turner, Ray
DOE Joint Genome Institute
crturner@lbl.gov

Tuskan, Gerald
Oak Ridge National Lab
tuskanga@ornl.gov

Udwary, Daniel
Berkeley Lab
dwudwary@lbl.gov

Ul-Hasan, Sabah
UC Merced
sul-hasan@ucmerced.edu

Unruh, Sarah
University of Missouri
unruh.saraha@gmail.com

U'Ren, Jana
University of Arizona
juren@email.arizona.edu

van der Lelie, Daniel
Gusto Global
daniel.vanderlelie@gmail.com

Varghese, Neha
DOE Joint Genome Institute
njvarghese@lbl.gov

Venturi, Vittorio
ICGEB
venturi@icgeb.org

Vesth, Tammi
Technical University of Denmark
tcve@bio.dtu.dk

Visel, Axel
DOE Joint Genome Institute
avisel@lbl.gov

Vogel, John
DOE Joint Genome Institute
jpvogel@lbl.gov

Voriskova, Jana
Berkeley Lab
jvoriskova@lbl.gov

Waage, Carey
DOE Joint Genome Institute
cvwaage@lbl.gov

Wakatsuki, Soichi
SLAC National Accelerator Laboratory
soichi.wakatsuki@stanford.edu

Waldo, Nick
University of Washington
nbwaldo@uw.edu

Walsh, David
Concordia University
david.walsh@concordia.ca

Wang, Zhiyong
Carnegie Institution for Science
zywang24@stanford.edu

Wang, Zhong
DOE Joint Genome Institute
zhongwang@lbl.gov

Wang, Mei
DOE Joint Genome Institute
mwang@lbl.gov

Wang, Bing
DOE Joint Genome Institute
bwang16@lbl.gov

Wang, Gaoyan
DOE Joint Genome Institute
gaoyanwang@lbl.gov

Wang, Yuan
The Samual Roberts Noble Foundation
ywang@noble.org

Ward, Christopher
Lawrence Livermore National Lab
ward56@llnl.gov

Ward, Naomi
University of Wyoming
nlward@uwyo.edu

Ware, Doreen
CSHL / USDA-ARS
ware@cshl.edu

Watson, Andrea
University of Chicago
andreawatson@uchicago.edu

Webb, Eric
University of Southern California
eawebb@usc.edu

Wegrzyn, Renee
DARPA
renee.wegrzyn@darpa.mil; reneelumm@gmail.com

Wheeldon, Ian
UC Riverside
iwheeldon@engr.ucr.edu

Whitman, William
University of Georgia
whitman@uga.edu

Whitman, Thea
University of Wisconsin - Madison
twhitman@wisc.edu

Wing, Rod
University of Arizona
rwing@mail.arizona.edu

Wisecaver, Jennifer
Vanderbilt University
jen.wisecaver@vanderbilt.edu

Wong, Kristine
DOE Joint Genome Institute
kristinewong@lbl.gov

Woyke, Tanja
DOE Joint Genome Institute
twoyke@lbl.gov

Xie, Yichun
Chinese University of Hong Kong
xieyichun5050@gmail.com

Yadon, Denise
DOE Joint Genome Institute
dyadon@lbl.gov

Yan, Juying
DOE Joint Genome Institute
juyingyan@lbl.gov

Yang, Yongil
Oak Ridge National Lab
yi_yang@ornl.gov

Yang, Xiaohan
Oak Ridge National Lab
yangx@ornl.gov

Yao, Qiuming
Oak Ridge National Lab
yaoq@ornl.gov

Yoo, Shinjae
BNL
sjyoo@bnl.gov

Yoshikuni, Yasuo
DOE Joint Genome Institute
yyoshikuni@lbl.gov

Yoshinaga, Yuko
DOE Joint Genome Institute
yyoshinaga@lbl.gov

Zaini, Paulo
UC Davis
pazaini@ucdavis.edu

Zane, Matthew
DOE Joint Genome Institute
mczane@lbl.gov

Zaretzki, Jed
Novozymes, Inc.
jze@novozymes.com

Zeiner, Carolyn
Boston University
zeiner@bu.edu

Zeng, Zhirui
Stanford University
zrzeng@stanford.edu

Zerbe, Philipp
UC Davis
pzerbe@ucdavis.edu

Zhalnina, Kateryna
Berkeley Lab
kzhalnina@lbl.gov

Zhang, Qizhi
Argonne National Laboratory
qzhang@anl.gov

Zhang, Zheyun
DOE Joint Genome Institute
zheyunzhang@lbl.gov

Zhang, Qingpeng
DOE Joint Genome Institute
qingpeng@lbl.gov

Zhang, Feng
Institut National de la Recherche Agronomique
feng.zhang@inra.fr

Zhao, Jean
DOE Joint Genome Institute
zyzhao@lbl.gov

Zhong, Yiyi
CUHK
yiyi11zhong@gmail.com

Zhou, Jizhong
University of Oklahoma
jzhou@ou.edu

Zhu Jin
Monsanto
jin.1.zhu@monsanto.com

Zimmerman, Naupaka
University of San Francisco
nzimmerman@usfca.edu

jgi.doe.gov

JGI JOINT GENOME INSTITUTE
A DOE OFFICE OF SCIENCE USER FACILITY

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY
Office of Science

17-JGI-3515