U.S. DEPARTMENT OF ENERGY

2019 Genomic Sciences Program (GSP) Annual PI Meeting

Abstract Book

Sponsored by the
U.S. Department of Energy
Office of Biological and Environmental Research
Biological Systems Science Division

Department of Energy
Washington, DC 20585
February 25, 2019

Dear Colleague:

On behalf of the Biological Systems Science Division (BSSD) welcome to the 2019 Genomic Science Annual PI Meeting! It has been another very productive year and I'd like to thank you all for your continued interest in our programs and for your scientific accomplishments that keep this program at the forefront of genomic science. We have planned a full agenda highlighting the exceptional research results produced within the program over the past year. We hope that these presentations will spark fruitful discussions that can be carried into the poster sessions later today and tomorrow.

The Bioenergy Research Center program is in its 12th year and we will hear from the center directors about each BRC's research vision for the future. These Centers continue to be extremely productive and play a central role in BER's bioenergy research but our basic research goals are broader than the BRCs. We continue to solicit research in a variety of topics. Over the past year we've had an opportunity to bring in a cohort of new projects within the Systems Biology for Bioenergy portfolio and the Plant Feedstocks portfolio. These small-team and single-PI projects complement and expand research on microbial and plant species for bioenergy research beyond what we currently have under study in our larger BRCs. As a program we continue to support a broad portfolio of basic research in a variety of sizes and configurations because we've found this combination to be incredibly fruitful within the program. Nowhere is this most evident than at this annual meeting where PIs from all sizes of projects across the portfolio are able to interact, learn from and build on each other's successes.

We are continuing to reconfigure our computational biology and bioinformatics systems within JGI and KBase to create a more open access and online system for use within the program. Having a common online framework to access, analyze and share analyses of large complex omics/genomic datasets in a reproducible manner is crucial to accelerating research within our programs. These efforts will be further aided as we stand up a new National Microbiome Data Collaborative (NMDC). The NMDC will add to these existing systems and build out a computational infrastructure needed to support and enable a broad microbiome science community.

At this meeting there will be plenary and poster presentations of a variety of enabling capabilities supported within the Office of Biological and Environmental Research (BER) including the DOE Joint Genome Institute (JGI), the Environmental Molecular Science Laboratory (EMSL), the DOE Synchrotron Light and Neutron Sources and, the DOE Systems Biology Knowledgebase (KBase). Investigators from the Bioimaging research portfolio will also be in attendance beginning on Tuesday evening and will participate in the Bioimaging Research annual PI meeting immediately following this meeting. I would urge you to visit with representatives of these facilities/projects to learn more about new and/or upcoming capabilities.

To complement ongoing research we are continuing to pursue new enabling capabilities within the portfolio. This past year we have positioned instrumentation and operations support to facilitate broader access to major cryo-EM centers at SLAC (funded by NIH) and Brookhaven National Lab (funded by NY State). We've also initiated a pilot activity at ORNL to pursue neutron science applications in biology. These are exciting new capabilities that will be available to researchers and we think will present new opportunities for scientific discovery. We will keep you informed on their progress. We will also hear brief-outs from two BSSD sponsored workshops on new topics for future research.

We are extremely pleased to welcome Dr. Jonathan Schilling from the University of Minnesota as our keynote speaker for this meeting. Dr. Schilling will speak about his research on fungal metabolism. If you are familiar with Dr. Schilling's research then you know we are in for a treat. Dr. Schilling has been a leader in understanding the mechanisms of biomass breakdown by fungi and this talk is quite timely and highly relevant to this meeting.

We will also hear the latest updates on new applications being developed within KBase later today and a Hands-On training session on Tuesday. Programmatically we enabled a subset of the DOE National Laboratory programs to build out applications in KBase that will not only enable their own research but will be available to KBase users as well. We've also planned a more forward-looking breakout session on new computational approaches for modelling and data analysis being developed within the portfolio.

Finally, we are proud to once again host a plenary session featuring some of the most recent recipients of the Office of Science (SC) Early Career program awards. The Early Career program is consistently one of the most competitive programs within SC and award recipients are part of an exclusive group. We are pleased to have four awardees present at this year's meeting.

Finally, I would ask that you treat this meeting as an opportunity to view your research in the context of the larger portfolio. All funded principal investigator projects from our University portfolio and the DOE National Laboratory portfolio are present at this meeting. Please take full advantage of this opportunity to meet with your colleagues and to meet with your DOE program staff, representatives from elsewhere within the Department of Energy, and colleagues from other Federal Agencies.

Thank you again for making the program the success that it is. We look forward to another excellent meeting!

Sincerely,

Todd Anderson, Ph.D.
Director, Biological Systems Science Division, SC-23.2
Office of Biological and Environmental Research
Office of Science

Table of Contents

(Click on Title to advance to the abstract)

Paul Adams: ENIGMA: Assembly Mechanism of Subsurface Microbial Community under Stress Gradient and Adaptation of Super Phylum Patescibacteria with Genome Simplicity
PRESENTER: Daliang Ning

Paul Adams: ENIGMA: New technologies for the large-scale screening of microbial interactions
PRESENTER: Fangchao Song

Paul Adams: ENIGMA: Long-Term Continuous Monitoring Gives Insight into Patterns Between Groundwater and Weather Events
PRESENTER: Kathleen Walker

Paul Adams: ENIGMA: Using in-field bioreactors to monitor microbial community dynamic shifts with geochemical perturbations
PRESENTER: Regina Wilpiszeski

Paul Adams: ENIGMA: Building a reference-based metagenomics workflow in KBase
PRESENTER: Anni Zhang

Andrew Allen: New advances in diatom functional genomics: Gene knockout methodology and genome-wide mapping of transcription factor binding sites
PRESENTER: Andrew Allen

Andrew Allen: Subcellular compartmentalization and biochemistry of lysine biosynthesis in the model diatom
PRESENTER: Christopher Dupont

Andrew Allen: Developing episome-based, gene expression modulation by exogenous chemicals in the diatom Phaeodactylum tricornutum.
PRESENTER: Tessema Kassaw

Andrew Allen: Interactive Metabolic Pathway Visualization for Optimization of Energy Flux
PRESENTER: Zachary King

Andrew Allen: The Plastid Terminal Oxidase does not play a photoprotective role in the marine diatom Phaeodactylum tricornutum
PRESENTER: Graham Peers

Andrew Allen: Dynamic metabolic constraints of Phaeodactylum tricornutum reveal bases on organelle-specific carbon and nitrogen partitioning during nitrogen depletion
PRESENTER: Cristal Zuniga

Steve Allison: Fungal Trait Tradeoffs in a Southern Californian Grassland
PRESENTER: Charlotte Alster

Steve Allison: Physiological adaptations of leaf litter microbial communities to drought reduce decomposition rates
PRESENTER: Ashish Malik

Steve Allison: Microbial Communities Exhibit Resilient Extracellular Enzyme Activity Along a Climate Gradient
PRESENTER: Bahareh Sorouri

Ana Alonso: Investigating the Natural Variation of Pennycress Metabolome, an Emerging Crop for Aviation Biofuel.
PRESENTER: Ana Paula Alonso

Ana Alonso: Techno-economic, Energy and Greenhouse Gas Emissions Analyses of Pennycress Production and Logistics for Aviation Biofuel
PRESENTER: Ajay Shah

Daniel Amador-Noguez: In vivo thermodynamic analysis of metabolic networks and engineered pathways
PRESENTER: Daniel Amador-Noguez

James A Anderson: Genetic improvement of seed yield and oil content in field pennycress, a non-food oilseed feedstock species
PRESENTER: Ratan Chopra

Adam Arkin: KBase: The Systems Biology Knowledgebase for Predictive Biological and Environmental Research in an Integrated Data Platform
PRESENTER: Benjamin Allen

Adam Arkin: Microbiome Genome Extraction, Phylogenomics, and Metabolic Modeling of Species Interactions in KBase
PRESENTER: Dylan Chivian

Adam Arkin: New pipeline for multi-omics data integration and discovery in KBase to identify the mechanism driving metabolic/OTU dynamics in a microbiome
PRESENTER: Christopher Henry

Adam Arkin: Transcriptome Profiling and Genome-Scale Metabolic Modeling in Fungal and Plant genomes using KBase
PRESENTER: Sunita Kumari

Adam Arkin: KBase User Engagement: Leveraging Outreach and In-Reach Strategies to Build a Collaborative Community
PRESENTER: Elisha Wood-Charlson

Jose Avalos: Biosensor and Optogenetics for Systems Biology of Yeast Branched-Chain Higher Alcohol Production and Tolerance
PRESENTER: Jose Avalos

Rebecca Bart: Optimizing tradeoffs implicit during bioenergy crop improvement: Understanding the effect of altered cell wall and sugar content on sorghum-associated pathogenic bacteria
PRESENTER: Rebecca Bart

Ivan Baxter: Manipulating Carbon Flux between Mesophyll and Bundle Sheath Cells to Optimize Photosynthetic Performance
PRESENTER: Jennifer Arp

Ivan Baxter: Phenotyping for water use efficiency in C4 grasses Setaria and Sorghum
PRESENTER: Ivan Baxter

Ivan Baxter: Tissue and environment-specific synthetic gene circuits to regulate crown root growth
PRESENTER: Jennifer Brophy

Ivan Baxter: Modification of PEPC kinetics to enhance the efficiency of C4 photosynthesis
PRESENTER: Asaph Cousins

Ivan Baxter: Metabolome and transcriptome responses to a water deficit time course in the model C4 grass Setaria Viridis
PRESENTER: Allen Hubbard

Ivan Baxter: Development of Transgenic Sorghum Lines to Enhance Water Use and Photosynthetic Efficiencies
PRESENTER: Albert Kausch

Ivan Baxter: Phenomics of Stomata and Water Use Efficiency in C4 Species
PRESENTER: Andrew Leakey

Ivan Baxter: Leveraging Setaria viridis and Sorghum bicolor pan-genomic variation for gene mining and crop improvement
PRESENTER: Scott Lee

Ivan Baxter: Targeted Mutagenesis and Programmed Transcriptional Regulation in Setaria and Sorghum
PRESENTER: Yang Liu

Ivan Baxter: Improving water-use efficiency under fluctuating light
PRESENTER: Charles Pignon

Ivan Baxter: QTG-Finder: A Machine-Learning Algorithm to Prioritize Causal Genes of Quantitative Trait Loci in Plants
PRESENTER: Seung Rhee

Ivan Baxter: Computer-aided Engineering of Biomass Production under Drought in Sorghum and Setaria
PRESENTER: Seung Rhee

Ivan Baxter: Investigation of Carbonic Anhydrase to Improve Sorghum C4 Photosynthetic Efficiency
PRESENTER: Josh Rosnow

Michael Betenbaugh: Creating Multifunctional Synthetic Lichen Platforms for Sustainable Biosynthesis of Biofuel Precursors
PRESENTER: Jackson Jenkins

Michael Betenbaugh: Synthetic lichens as a sustainable platform for biosynthesis
PRESENTER: Cristal Zuniga

Crysten Blaby: QPSI: Quantitative Plant Science Initiative
PRESENTER: Crysten Blaby

Crysten Blaby: Phylogenomics-guided Approaches for Predicting and Characterizing Functions Across a Multi-functional Protein Family
PRESENTER: Miriam Pasquini

Eduardo Blumwald: Integration of Physiological Phenotyping and Cell-Type-Specific Omics Approaches to Study Individual and Abiotic Stress Combinations in Poplar
PRESENTER: Amir H. Ahkami

Eduardo Blumwald: Discovery of Novel Cis-Regulatory Elements Responsive to Salt-Stress in Hybrid Poplar for Designing Inducible Synthetic Promoters
PRESENTER: JunHyung Lee

Nicholas Bouskill: Microbial environmental feedbacks and the evolution of soil organic matter
PRESENTER: Nicholas Bouskill

Nanette Boyle: Enabling Predictive Metabolic Modeling of Diurnal Growth Using a Multi-Scale Multi-Paradigm Approach
PRESENTER: Nanette Boyle

Tom Bruns: Microbial Responses to Pyrogenic Organic Matter: Cultivating Isolates, Ageing PyOM, Tracing Gas Fluxes, and Meta-analysis
PRESENTER: Timothy Berry

Tom Bruns: Functional genomics of pyrophilous fungi – determining the fate of pyrolyzed carbon in post-fire soils
PRESENTER: Sara Calhoun

Tom Bruns: Rising From the Ashes: Succession and Metabolism of Post-Fire Fungi
PRESENTER: Monika Fischer

Tom Bruns: Microbial Community Succession and Metabolism in Post-Fire Environments
PRESENTER: Neem Patel

Daniel Buckley: The Ecophysiological Determinants of Mineral-Associated Carbon in Agricultural Soil
PRESENTER: Roland Wilhelm

William Cannon: Simulation and Experiment Uncover Complementary Regulation in the Circadian System Neurospora crassa
PRESENTER: William Cannon

Patrick Chain: Screening Fungal Genome Sequencing Data and Culture Collections to Better Understand Bacterial:Fungal Interactions
PRESENTER: Geoffrey House

Patrick Chain: Democratization of fungal highway columns as a tool to investigate bacteria associated with soil fungi
PRESENTER: Pilar Junier

Patrick Chain: Using Gel Microdroplets to Investigate Bacterial Influence on Fungal Spore Germination
PRESENTER: Julia Kelliher

Patrick Chain: Taking a Closer Look at Bacterial & Fungal Interactions in Soil Using Imaging
PRESENTER: Demostheens Morales

Patrick Chain: Bacterial Oxalotrophy as a Biocontrol Mechanism against Fungal Pathogens
PRESENTER: Fabio Palmieri

Patrick Chain: Diversity of endo- and exo-bacteria associated with soil fungi
PRESENTER: Debora Rodrigues

George Church: Parallelized in vivo Construction of a Synthetic 57-Codon E. coli Genome
PRESENTER: Maximilien Baas-Thomas

George Church: Establishing and optimizing RecT-mediated homologous recombination in bacteria beyond E. coli
PRESENTER: Gabriel Filsinger

George Church: Deciphering Microbial Structures in situ Using Subcellular RNA Sequencing
PRESENTER: Andrew Pawlowski

Luca Comai: Discovery and characterization of disease resistance loci using a unique gene copy number variant population
PRESENTER: Heloise Bastiaanse

Gloria Coruzzi: EvoNet: A Phylogenomic and Systems Biology approach to identify genes underlying plant survival in marginal, low-Nitrogen soils
PRESENTER: Gil Eshel

Hugo Cuevas: Anthracnose resistant response of the USDA-NPGS sweet sorghum [Sorghum bicolor (L.)] germplasm collection
PRESENTER: Hugo Cuevas

Gautam Dantas: Characterizing the transcriptomes of Rhodoccocus opacus strains adapted to model lignin breakdown products
PRESENTER: Winston Anthony

Gautam Dantas: Multi-omic analysis of adapted Rhodococcus opacus strains for the characterization of aromatic metabolism
PRESENTER: Tayte Campbell

Gautam Dantas: A Concerted Systems Biology Analysis of Aromatic Metabolism in Rhodococcus opacus PD630
PRESENTER: Garrett Roell

Gautam Dantas: Exploring the Aging of Lignin Breakdown Products for Biological Conversion
PRESENTER: Zeyu Shang

Brian Davison: Solvent-induced Membrane Disruption Revealed by Neutron Scattering and Molecular Simulation
PRESENTER: Brian Davison

Brian Davison: Structural Characterization of Poplar Variants Provides New Insights into Plant Cell Wall Recalcitrance
PRESENTER: Hugh O'Neill

Brian Davison: Dynamics of the Lignin Glass Transition
PRESENTER: Loukas Petridis

Kristen DeAngelis: Influence of microbial diversity on carbon use efficiency in soils
PRESENTER: Luiz Domeignoz-Horta

Kristen DeAngelis: Heavy, Wet, and Stuck in the Mud: Quantifying the Effect of Violating Common Assumptions of the 18OH2 Method of Measuring Carbon Use Efficiency
PRESENTER: Grace Pold

Evan DeLucia: Using Marginal Lands for Bioenergy Crop Production in the Rainfed Regions of Continental U.S.: Economic Potential and Environmental Benefits
PRESENTER: Sarah Acquah

Evan DeLucia: Oilcane: Metabolic Engineering and Genome Editing to Improve Energy Density, Agronomic and Conversion Performance of Sugarcane
PRESENTER: Fredy Altpeter

Evan DeLucia: Miscanthus sinensis Genome: Updates on the Genome Assembly and Analysis of the Transcriptome and Small RNA
PRESENTER: Mohammad Belaffif

Evan DeLucia: Developing a Novel Plant-Microbial Interactions Model to Predict the Impacts of Bioenergy Crops on Soil Carbon and Nitrogen Cycling
PRESENTER: Edward Brzostek

Evan DeLucia: The Nitrogen Biogeochemical Impacts of Energy Sorghum, a Potential Bioenergy Crop
PRESENTER: Mark Burnham

Evan DeLucia: Economic Incentives for Achieving Conservation and Renewable Energy Goals with the Conservation Reserve Program
PRESENTER: Luoye Chen

Evan DeLucia: Economics of Plant Oil Recovery: A Review
PRESENTER: Ming-Hsun Cheng

Evan DeLucia: Bioconversion of Pelletized Big Bluestem, Switchgrass, and Low-Diversity Grass Mixture into Sugars and Bioethanol
PRESENTER: Bruce Dien

Evan DeLucia: Evaluating the effects of the Renewable Fuel Standard on Water Quality: An Integrated Economic Assessment
PRESENTER: Kelsie Ferin

Evan DeLucia: TOWARDS QUANTIFYING THE CENTRAL DOGMA OF MOLECULAR BIOLOGY ON A GENOME-WIDE SCALE
PRESENTER: Meera Gupta

Evan DeLucia: Identification of Sorghum stem-preferred genes and their cis-regulatory elements through de novo promoter analysis
PRESENTER: Madara Hetti-Arachchilage

Evan DeLucia: Modeling Energy Sorghum Emissions for the Rainfed United States
PRESENTER: Tara Hudiburg

Evan DeLucia: Integrating transcriptomics and metabolomics in the oleaginous yeast Lipomyces starkeyi for the production of biofuels and bioproducts
PRESENTER: Sujit Jagtap

Evan DeLucia: A Role for Differential Rca Isoform Expression in C4 Bioenergy Grass Thermotolerance?
PRESENTER: Sang Yeol Kim

Evan DeLucia: Improved high-throughput quantification of nitrogen cycling genes in bioenergy crop soils
PRESENTER: Lanying Ma

Evan DeLucia: CROPSR: An Automated Platform for Complex-Genome Wide CRISPR gRNA Design and Validation
PRESENTER: Hans Müller Paul

Evan DeLucia: Genomics-Assisted Improvement of Abiotic Stress Tolerance and Yield in the Biomass Feedstock Crops Miscanthus and Saccharum
PRESENTER: Erik Sacks

Evan DeLucia: Cooperative Asymmetric Reactions Combining Photocatalysis and Enzymatic Catalysis
PRESENTER: Yajie Wang

Evan DeLucia: Toward a Dynamic Photosynthesis Model to Guide the Yield Improvement for C4 Energy Crop

PRESENTER: Yu Wang

Evan DeLucia: Population Genomics of Issatchenkia orientalis Reveals the Genetic Mechanisms of its Multi-Stress Tolerance

PRESENTER: Yasuo Yoshikuni

Mitch Doktycz: Plant-Microbe Interfaces: From gene discoveries to molecular, genetic and biochemical validations in Populus

PRESENTER: Jay Chen

Mitch Doktycz: Plant-Microbe Interfaces: Detection and characterization of signaling polypeptides in plant-microbial systems using high-performance tandem mass spectrometry

PRESENTER: Robert Hettich

Mitch Doktycz: Plant-Microbe Interfaces: Integrated omics and phytobiome analysis of the adaptation of Populus deltoides to acute progressive drought and cyclic drought

PRESENTER: Daniel Jacobson

Mitch Doktycz: Plant-Microbe Interfaces: Assembly and function of the ectomycorrhizaome of Populus trichocarpa

PRESENTER: Brian Looney

Mitch Doktycz: Plant-Microbe Interfaces: Measuring rates and consequences of horizontal gene transfer in the Populus rhizosphere

PRESENTER: Joshua Michener

Mitch Doktycz: Plant-Microbes Interfaces: Investigating microbial recruitment and colonization dynamics in the poplar microbiome

PRESENTER: Jennifer Morrell-Falvey

Mitch Doktycz: Plant-Microbe Interfaces: Transcriptional co-regulation of MYB46 and WRKY33 by ANGUSTIFOLIA modulates Arabidopsis resistance towards biotrophic and necrotrophic pathogens

PRESENTER: Wellington Muchero

Mitch Doktycz: Plant-Microbe Interfaces: Cultivating a representative Populus microbiome?

PRESENTER: Dale Pelletier

Mitch Doktycz: Plant-Microbe Interfaces: Accessing the hidden diversity of the poplar microbiota through targeted metagenomics and cultivation

PRESENTER: Mircea Podar

Mitch Doktycz: Plant-Microbe Interfaces: Effects of variation in host secondary chemistry across Populus genotypes on the composition of the rhizosphere microbiome

PRESENTER: Christopher Schadt

Mitch Doktycz: Plant-Microbe Interfaces: Systematic discovery and prediction of novel LuxI-type quorum sensing signals in members of the Populus microbiome
PRESENTER: Amy Schaefer

Mitch Doktycz: Plant-Microbe Interfaces: The nature of the progression of drought stress drives differential metabolomic responses in Populus deltoides
PRESENTER: Timothy Tschaplinski

Mitch Doktycz: Plant-Microbe Interfaces: Functional characterization of small secreted proteins in poplar in relation to Populus-Laccaria symbiosis
PRESENTER: Xiaohan Yang

Tim Donohue: Approaches to Increasing Terpenoid Production in Zymomonas mobilis by Improving Activity of IspG and IspH of the MEP Pathway
PRESENTER: Isabel Askenasy

Tim Donohue: The Response of Bacterial Communities to Nitrogen Fertilization Depends on Temporal and Spatial Scale
PRESENTER: Lukas Bell-Dereske

Tim Donohue: Vertical niche analysis of rhizosphere and soil fungal and bacterial communities: A deep look into sustainable biofuel cropping systems
PRESENTER: Pedro Beschoren da Costa

Tim Donohue: Advanced Biofuels of the Future: Atom–Economical or Energy–Economical?
PRESENTER: Peyman Fasahati

Tim Donohue: Modulating hemicellulose to improve bioenergy crops
PRESENTER: Sang-Jin Kim

Tim Donohue: Biomimetic Lignin Cleavage Using Small Thiols
PRESENTER: Grace Klinger

Tim Donohue: Engineering Alternative Model Yeast Species for Biofuel Production
PRESENTER: David Krause

Tim Donohue: Genetics in Zymomonas mobilis ZM4
PRESENTER: Piyush Lal

Tim Donohue: Kinetic and mechanistic insights into hydrogenolytic production of lignin monomers in a continuous flow-through system
PRESENTER: Yanding Li

Tim Donohue: Regulation of biomass yield and composition of energy sorghum stems
PRESENTER: Brian McKinley

Tim Donohue: Development of Biosensors for High-throughput Enzyme Screens
PRESENTER: Anthony Meger

Tim Donohue: Microbial Conversion of Chemically Depolymerized Lignin Into Valuable Compounds
PRESENTER: Jose Perez

Tim Donohue: Investigating Mechanisms of Soil Carbon Accrual and Protection in Bioenergy Cropping Systems
PRESENTER: Matthew Reid

Tim Donohue: The Other Lignocellulosic Leftover: Conversion Residue Valorization with a Microbial Community
PRESENTER: Matthew Scarborough

Tim Donohue: Overexpression and Metabolic Regulation of Z. mobilis MEP Pathway Enzymes
PRESENTER: Mehmet Tatli

Tim Donohue: Using Engineered Streptomyces for Production of Fatty Acids and Isoprenoids from Lignocellulosic Biofuel Conversion Residue
PRESENTER: Kurt Throckmorton

Tim Donohue: The Structure and Specificity Landscape of a Glycoside Hydrolase Family
PRESENTER: Kirk Vander Meulen

Tim Donohue: Producing Valuable Bioproducts from the Waste Residue of Biofuel Production
PRESENTER: John Wolters

Tim Donohue: Comparative Studies of Diverse Lignocellulosic Feedstocks for Microbial Biofuel Synthesis and the Impact on the Biorefinery
PRESENTER: Yaoping Zhang

John Dunbar: Principles for Successfully Engineering Microbial Community Functioning
PRESENTER: Michaeline Albright

John Dunbar: Role Of Geographic Scale In Likelihood Of Microbial Driven Variation In Litter Decomposition
PRESENTER: John Dunbar

John Dunbar: Microbial lineages linked to decomposition outcomes across multiple litter types
PRESENTER: Renee Johansen

John Dunbar: Dissolved organic carbon (DOC) abundance links to microbial and DOC composition in soil
PRESENTER: Danielle Ulrich

Mary Dunlop: High-throughput Chemical Imaging for Optimizing Biofuel Synthesis Using Synthetic Biology
PRESENTER: Mary Dunlop

Serge Edme: Genetics and Genomcis ocf Pathogen Resistance in Switchgrass
PRESENTER: Serge Edme

Edward Eisenstein: Identification and Characterization of Rust Effectors that Affect Host Immunity
PRESENTER: Ed Eisenstein

Mary Firestone: Conversion of Natural Grassland into Biofuel Crops (Switchgrass, Panicum virgatum L.) is Associated with Reduction in Methane Consumption
PRESENTER: Colin Bates

Mary Firestone: Characterizing the Multitrophic Interactions that Mediate Carbon Flow in Soil and their Responses to Drought
PRESENTER: Javier Ceja-Navarro

Mary Firestone: Unravelling Rhizosphere-microbial Interactions in the Rhizosphere of Alamo Switchgrass (Panicum virgatum) under Abiotic Stresses
PRESENTER: Na Ding

Mary Firestone: Following Soil Carbon Cycling under Reduced Precipitation in a California Annual Grassland
PRESENTER: Katerina Estera-Molina

Mary Firestone: Performance of Switchgrass 'Alamo' Population and Selections for Sustainable Production on Marginal Soils under Low-Input Management
PRESENTER: Mary Firestone

Mary Firestone: Effects of Switchgrass (Panicum virgatum L.) on Deep Carbon Pools in Marginal Lands
PRESENTER: Jialiang Kuang

Mary Firestone: Diversity of Viruses in Soil: Do Interactions with Soil Organisms Impact Carbon Flow?
PRESENTER: Alexa Nicolas

Mary Firestone: Plant-microbe and microbe-microbe interactions mediate switchgrass sustainability: following rhizosphere microbial communities during switchgrass establishment
PRESENTER: Erin Nuccio

Mary Firestone: Exploring Potential Interactions of Bacteria and Arbuscular Mycorrhizal Fungi Using Network Analyses of an Annual Grassland Soils
PRESENTER: Mengting Yuan

Mary Firestone: Molecular basis of plant-microbe-soil interactions shaping microbial communities and ecosystem function in a Mediterranean grassland
PRESENTER: Kateryna Zhalnina

Deanna Funnell-Harris: Resistance to Stalk Pathogens for Bioenergy Sorghum
PRESENTER: Deanna Funnell-Harris

Jeffrey Gardner: Saprophytic Bacterium Cellvibrio japonicus Has Selective CAZyme Requirements During Physiologically Relevant Hemicellulose Degradation
PRESENTER: Jeffrey Gardner

Ryan Gill: Automated Circuit Design in Saccharomyces cerevisiae
PRESENTER: Ye Chen

Ryan Gill: Design and Engineering of Synthetic Control Architectures for Complex Phenotypes in E. coli and Yeast
PRESENTER: Emily Freed

Ryan Gill: Design and engineering of native regulatory networks in non-model microbes
PRESENTER: Margaret Habib

Ryan Gill: CRISPR-based trackable protein engineering
PRESENTER: Eun Joong Oh

Louise Glass: m-CAFEs: The use of EcoFABs to study beneficial plant-fungus interactions
PRESENTER: Jens Heller

Louise Glass: m-CAFEs: EcoFABs: Investigating Soil Microbe-Plant Interactions in Regulated Ecosystems
PRESENTER: Lauren Jabusch

Louise Glass: m-CAFEs Applications of Targeted Killing and Editing Methods in Microbial Networks
PRESENTER: Matthew Nethery

Kirsten Hofmockel: Molecular Composition of Field Derived Microbial Necromass
PRESENTER: KIrsten Hofmockel

Bruce Hungate: Can we measure taxon-specific biochemical efficiency in natural microbial communities?
PRESENTER: Bruce Hungate

Bruce Hungate: Genomic traits predict microbial growth in culture but fail in soils, except during resource pulses
PRESENTER: Ben Koch

Bruce Hungate: Evolutionary history constrains microbial traits across environmental variation
PRESENTER: Ember Morrissey

Tiffany Jamann: Conserved Genetic Mechanisms for Biotic Stress in Sorghum
PRESENTER: Tiffany Jamann

Janet Jansson: Soil Moisture Modulates Inter-kingdom Interactions as Visualized Using a Simulated Soil Core
PRESENTER: Christopher Anderson

Janet Jansson: Determining the Soil Microbiome Community Phenotype "Metaphenome" in Response to Changing Soil Moisture and Carbon Content
PRESENTER: Janet Jansson

Janet Jansson: Probing the Relationship Between Osmolytes and Respiration in Soils - in Real Time
PRESENTER: Mary Lipton

Janet Jansson: Development and Analysis of Reduced Complexity Microbial Consortia Emerging from Native Grassland Soil Systems
PRESENTER: Ryan McClure

Janet Jansson: Untapped Viral Diversity in Global Soil Metagenomes
PRESENTER: Jason McDermott

Janet Jansson: Integrative Network Modeling for Predicting Biochemical Signatures in Complex Soil Systems
PRESENTER: Hyun-Seob Song

Kolby Jardine: Poplar Esterified Cell Wall Transformations and Metabolic Integration (PECTIN) Study
PRESENTER: Kolby Jardine

Michael Jewett: Mining the largest collection industrially-deployed Clostridium strains for highly-evolved gene variants related to improved productivity and industrial robustness
PRESENTER: Michael Koepke

Michael Jewett: Development of a Clostridia Cell-free Platform Facilitating Accelerated Engineering of Clostridia Strains
PRESENTER: Antje Kruger

Michael Jewett: Enhancing Control of Cell-free Metabolism Through pH Modulation
PRESENTER: Blake Rasor

Michael Jewett: Computational tools to predict novel chemical reactions and metabolic network behavior
PRESENTER: Keith Tyo

Thomas Juenger: Climate adaptation and sustainability in switchgrass: exploring plant-microbe-soil interactions across continental scale environmental gradients
PRESENTER: Thomas Juenger

Thomas Juenger: Microbial successions in switchgrass roots
PRESENTER: Esthr Singer

Jay Keasling: Direct Biomass Conversion through the One-Pot Ionic Liquid Technology
PRESENTER: Ezinne Achinivu

Jay Keasling: Biomass Sorghum Supply Systems for Advanced Biofuel Production
PRESENTER: Nawa Baral

Jay Keasling: Structural and Biochemical Investigation of Terminal Alkene Formation in Allylmalonyl-CoA Biosynthesis
PRESENTER: Jacquelyn Blake-Hedges

Jay Keasling: Development Of An Automated “Cells-To-Peptides” Sample Preparation Workflow For High-Throughput Quantitative Proteomic Applications
PRESENTER: Yan Chen

Jay Keasling: Heterologous Production of Lignin Modifying Enzymes in Aspergillus niger
PRESENTER: Sophie Comyn

Jay Keasling: Lignin Manipulation in Feedstocks Towards Lignin Valorization and Bioproducts
PRESENTER: Aymerick Eudes

Jay Keasling: An Automated Microfluidic System for On-Chip Genetic Engineering Processes
PRESENTER: William Gaillard

Jay Keasling: Quantitative metabolic modeling at the Joint BioEnergy Institute
PRESENTER: Hector Garcia Martin

Jay Keasling: Collaboration with the Experiment Data Depot
PRESENTER: Nathan Hillson

Jay Keasling: Developing a Targeted Mass-Spectrometry Platform for High-Throughput Characterization of Lignocellulosic Biomass Deconstruction
PRESENTER: Nicole Ing

Jay Keasling: Isopentenol Production Using the IPP-Bypass Mevalonate Pathway
PRESENTER: Taek Lee

Jay Keasling: Machine Learning for Bioenergy Sorghum Yield Prediction under Future Climate Scenarios
PRESENTER: Umakant Mishra

Jay Keasling: Applying Multi-dimensional Solid-state NMR to Explore the Nanoarchitecture of Native and Engineered Plant Cell Walls
PRESENTER: Jenny Mortimer

Jay Keasling: Importance of metabolic state for efficiency of bioproduction using non ribosomal peptide synthetase
PRESENTER: Aindrila Mukhopadhyay

Jay Keasling: Field testing of a low lignin engineering strategy in switchgrass
PRESENTER: Kavitha Satish Kumar

Jay Keasling: Drought and Bioenergy Grasses: Effects of Drought on Biomass Composition and Effect of Biomass Composition on Drought Tolerance
PRESENTER: Henrik Scheller

Jay Keasling: Development and demonstration of CRISPR/Cas9 platform for Aspergillus niger
PRESENTER: Blake Simmons

Jay Keasling: The response of Pseudomonas putida to a sorghum lignolysate
PRESENTER: Steven Singer

Jay Keasling: Investigating Lignin Modifying Enzymes and their Synergistic Effect with Ionic Liquid Pretreatment
PRESENTER: Paul Wolski

Matias Kirst: NitFix: Global-Scale Phylogenomics of the Nitrogen-fixing Clade
PRESENTER: Ryan Folk

Matias Kirst: NitFix: Phylogenomic Discovery and Engineering of Nitrogen Fixation into the Bioenergy Woody Crop Poplar
PRESENTER: Matias Kirst

Matias Kirst: NitFix: Identification of Conserved Non-Coding Sequences in Nitrogen Fixing Plants
PRESENTER: Matias Kirst

Matias Kirst: NitFix: Engineering Root Nodule Symbiosis in Populus sp.
PRESENTER: Lucas Maia

Jared LeBoldus: Towards durable resistance to Septoria stem canker and leaf spot: a molecular understanding of resistance.
PRESENTER: Jared LeBoldus

Peggy Lemaux: EPICON: From Leaves to Roots to Microbes – How Sorghum and Its Microbiome Respond to Drought
PRESENTER: Peggy Lemaux

Mary Lidstrom: Novel Insights into Rare Earth-Mediated Metabolisms, Aerobic Methanotrophy under Hypoxia, and the Role of Nitrogen Species in Methane Oxidation
PRESENTER: Ludmila Chistoserdova

Stephen Long: Biolistic Transformation of Miscanthus Species. Preliminary Studies.
PRESENTER: Joanna Cerazy-Waliszewska

Stephen Long: Improving Photosynthesis in C4 Bioenergy Crops
PRESENTER: Kher Xing Chan

Stephen Long: Overcoming Recalcitrance of Genetic Transformation in Energycane for Improving Cold Tolerance and Biomass Yield
PRESENTER: Baskaran Kannan

Stephen Long: Techno-economic Feasibility Analysis of Biodiesel and Ethanol Co-production from Lipid-producing Energycane
PRESENTER: Deepak Kumar

Stephen Long: Characterizing photosynthetic capacity of sugarcane under fluctuating lights
PRESENTER: Moonsub Lee

Stephen Long: Modular Assembly of Gene Constructs for Engineering Lipid Accumulation into Energycane
PRESENTER: Guanbin Luo

David Lowry: Identifying Physiological and Metabolic Response Mechanisms Associated with Divergence in Chilling and Freezing Tolerance Between Upland and Lowland Switchgrass Cultivars.
PRESENTER: Anne-Sophie Bohrer

David Lowry: Identification of Adaptive Fungal Pathogen Resistance Loci in Switchgrass
PRESENTER: David Lowry

David Lowry: Genetics and Environmental Contributions to Switchgrass Biofuel Traits
PRESENTER: Acer VanWallendal

David Lowry: Optimizing RNA in situ Hybridization for Stem Parenchyma Cell-specific Promoter Characterization in Energycane
PRESENTER: Jiang Wang

Chaofu Lu: Systems Biology to Improve Camelina Seed and Oil Quality Traits
PRESENTER: Chaofu Lu

Ting Lu: Designing microbial consortia with defined social interacgtions
PRESENTER: Ting Lu

Robert Martienssen: Biological Design of Lemnaceae Aquatic Plants for Biodiesel Production
PRESENTER: Robert Martienssen

Chris Marx: Using Gene Editing and an Accumulated Bioproduct as a Reporter for Genotypic and Phenotypic Heterogeneity in Growth-vs-Production for Methylobacterium extorquens Conversion of Aromatics to Butanol
PRESENTER: Tomislav Ticak

Julie Mitchell: Genome to Structure: From Multiple Sequence Alignment to Virtual Ligand Screening Using Co-Evolutionary Protein-Residue Contact-Prediction in KBase
PRESENTER: Ada Sedova

James Moran: Evaluating Biogeochemical Processes Within Microenvironments Along the Root-rhizosphere-soil Continuum
PRESENTER: James Moran

Paula Mouser: Influence of Biogeochemistry and Engineering Controls on Microbial Growth, Membrane Features, and Interactions with Shale Matrices in Engineered Energy Systems
PRESENTER: Paula Mouser

Wellington Muchero: Immune-suppressing pattern recognition receptors mediate host-driven recruitment of microbes
PRESENTER: Wellington Muchero

Krishna Niyogi: Comparative Genomic and Transcriptomic Analyses for Pathway Discovery in Chromochloris zofingiensis
PRESENTER: Fatima Foflonker

Krishna Niyogi: Single-Molecule, Whole-Transcript Sequencing of the Transcriptome of the Green Microalga Chromochloris zofingiensis to Accurately Annotate Gene Models and Identify Splice Variants
PRESENTER: Sean Gallaher

Krishna Niyogi: Lipidomics analysis of the emerging model green alga Chromochloris zofingiensis
PRESENTER: Yuntao Hu

Krishna Niyogi: Iron and Nitrogen Regulate Trophic Transitions and Metabolism in the Green Alga Chromochloris zofingiensis
PRESENTER: Tim Jeffers

Krishna Niyogi: Genome-scale Metabolic Model of Chromochloris zofingiensis, an Emerging Model Organism for Sustainable Fuel Production
PRESENTER: Alexander Metcalf

Krishna Niyogi: A Molecular Switch for Oxygenic Photosynthesis and Metabolism in the Emerging Oleaginous Model Green Alga Chromochloris zofingiensis
PRESENTER: Melissa Roth

Krishna Niyogi: Advancing the Chromochloris zofingiensis molecular toolkit
PRESENTER: Daniel Westcott

Trent Northen: The dynamic responses of biological soil crust communities revealed by wetting experiments coupled with metagenomics
PRESENTER: Marc Van Goethem

Victoria Orphan: Comparative genomic and activity-based analyses reveal widespread potential for direct extracellular electron transfer among diverse methane-oxidizing ANME archaea
PRESENTER: Grayson Chadwick

Victoria Orphan: Reactive Transport Modeling of Microbial Processes: Cell Aggregate-scale Models and Upscaling Using Pore-scale Simulations
PRESENTER: Xiaojia He

Himadri Pakrasi: Systems analysis of a fast growing N2-fixing cyanobacterium for production of advanced biofuels and nitrogen-containing petrochemical replacement compounds
PRESENTER: Anindita Bannerjee

Himadri Pakrasi: Metabolic Analysis of Cyanobacteria Carboxysome Mutant Indicates a More Flexible lux Network for Bio-manufacturing
PRESENTER: Jeffrey Czajka

Himadri Pakrasi: Genome-scale fluxome of the fast-growing cyanobacterium *Synechoccus elongatus* UTEX 2973
PRESENTER: John Irudaya Raj Hendry

Terry Papoutsakis: Syntrophic co-cultures of Clostridium organisms to produce higher alcohols and other C6-C8 metabolites
PRESENTER: Kamil Charubin

Kabir Peay: Environmental Drivers of the North American Populus Microbiome
PRESENTER: Michael Van Nuland

Gary Peter: Enhanced Resistance Pines for Improved Renewable Biofuel and Chemical Production
PRESENTER: Gary Peter

Jennifer Pett-Ridge: Taxon-Specific Growth and Mortality Rates of Bacteria in Soil Following the Rewetting of a Seasonally Dried Grassland
PRESENTER: Steve Blazewicz

Jennifer Pett-Ridge: Quantitative Stable Isotope Probing to Measure Microbial Growth and Activity
PRESENTER: Megan Foley

Jennifer Pett-Ridge: Linking Microbial Growth to Edaphic Variation Across a Mediterranean-Grassland Precipitation Gradient Using Shotgun-Metagenomic qSIP
PRESENTER: Alex Greenlon

Jennifer Pett-Ridge: Microbes Persist: Building a KBase Foundation for Viral Ecogenomics In Soil
PRESENTER: Jeffrey Kimbrel

Jennifer Pett-Ridge: The Impacts of Redox Periodicity on Microbial Community Structure and Carbon Transformations in a Wet Tropical Forest Soil
PRESENTER: Jennifer Pett-Ridge

Jennifer Pett-Ridge: Sensitivity Analysis and Methodological Improvements for Quantitative Stable Isotope Probing
PRESENTER: Ella Sieradzki

Virginia Rich: Solving Multi-disciplinary, Multi-Scale, Interactive Data Management for Heterogeneous Users: The IsoGenieDB
PRESENTER: Virginia Rich

Virginia Rich: Genome-resolved Insights Into Permafrost Thaw Processes
PRESENTER: Virginia Rich

Virginia Rich: Role of Viruses in Carbon Processing Along a Permafrost Thaw Gradient
PRESENTER: Matthew Sullivan

Virginia Rich: Integrating Metabolomics and Proteomics Reveals Biotic and Abiotic Controls on CO2 and CH4 Production in Peatlands
PRESENTER: Rachel Wilson

Erik Sacks: Introgression of Novel Disease Resistance Genes from Miscanthus into Energycane
PRESENTER: Erik Sacks

Howard Salis: Next-generation Synthetic Biology Technologies for Controlling Metabolism in Clostridia
PRESENTER: Howard Salis

Daniel Schachtman: Leaf Carbon and Nitrogen Isotope Composition in Diverse Sorghum Lines Under Differential Water and Nitrogen Treatments
PRESENTER: Asaph Cousins

Daniel Schachtman: Defining the Functional Genome Associated with Enhanced Water and Nitrogen Use Efficiency in Bioenergy Sorghum
PRESENTER: Andrea Eveland

Daniel Schachtman: Storage, Integration and Dissemination of data for a multi-disciplinary project.
PRESENTER: Phlip Ozersky

Daniel Schachtman: The Metabolome of Early Season Sorghum Plant Tissue is Predictive of End of Season Biomass
PRESENTER: Jessica Prenni

Daniel Schachtman: Traits for Development of Sustainable Energy Sorghum Biofuel Feedstock Production in Marginal Environments
PRESENTER: Daniel Schachtman

Daniel Schachtman: Sorghum root microbiome dynamics under nutrient-limited and drought conditions
PRESENTER: Susannah Tringe

Jonathan Schilling: Gene Regulatory Networks Enabling Fungi to Selectively Extract Sugars from Lignocellulose
PRESENTER: Jesus Castano

Danny Schnell: The Metabolic Basis of Low Biosynthetic Efficiency in an Oilseed
PRESENTER: Lisa Carey

Danny Schnell: Transcriptional Regulatory Network Analysis to Increase Yields in Camelina
PRESENTER: Erich Grotewold

Danny Schnell: Enhancing chloroplast-mitochondrial metabolic networks to improve photosynthesis and crop yields
PRESENTER: Danny Schnell

Danny Schnell: Strategies to Increase Photosynthetic Efficiency, CO2 Fixation and Resource Allocation in Crops.
PRESENTER: Heike Sederoff

John Sedbrook: Advancing Field Pennycress as a New Oilseed Biofuels Feedstock that does not Require New Land Commitments
PRESENTER: John Sedbrook

Rhona Stuart: Model exploration of carbon feedbacks between microalgae and heterotrophic bacteria
PRESENTER: Yiwei Cheng

Rhona Stuart: Improving Draft Metabolic Models in KBase: Tools for Importing, Comparing and Merging Metabolic Annotations
PRESENTER: Patrik D'haeseleer

Rhona Stuart: Changes to the elemental composition of the green alga Chlamydomonas reinhardtii based on variations in cultivation conditions
PRESENTER: Colleen Hui

Rhona Stuart: Chemical and Proteomic Profiling of Organic Compounds for Detecting Algal Interactions with Grazers and Commensalistic Bacteria
PRESENTER: Todd Lane

Rhona Stuart: Characterization of microalgal-bacterial symbioses at the microscale to examine association-dependent remineralization processes at the single-cell level
PRESENTER: Ty Samo

Rhona Stuart: Multitrophic considerations critical to understanding algal pond ecosystems
PRESENTER: Christopher Ward

Bob Tabita: Novel Microbial Routes to Synthesize Industrially Significant Precursor Compounds
PRESENTER: Bob Tabita

Neslihan Tas Baas: Understanding the microbial controls on biogeochemical cycles in permafrost ecosystems
PRESENTER: Neslihan Tas Baas

Cong Trinh: Harnessing robustness of Yarrowia lipolytica for effective utilization of undetoxified biomass hydrolysates
PRESENTER: Seunghyun Ryu

Cong Trinh: Understanding Exceptional Ionic Liquid Tolerance in Yarrowia lipolytica
PRESENTER: Caleb Walker

Danielle Tullman-Ercek: Employing Bacterial Microcompartments To Create Privileged Redox Pools for Biofuel Production

PRESENTER: Danielle Tullman-Ercek

Gerald Tuskan: Lignin Removal Controls Kinetics of Biomass Deconstruction by Consolidated Bioprocessing

PRESENTER: Christian Alcaraz

Gerald Tuskan: Mechanisms of Plant Cell Wall Polysaccharide Biosynthesis: A Two-Phase Model for the Non-Processive Biosynthesis of Homogalacturonan Polysaccharides by the GAUT1:GAUT7 Complex

PRESENTER: Robert Amos

Gerald Tuskan: Lignin Valorization: Biosynthesis and Bioengineering of an Ideal Lignin

PRESENTER: Fang Chen

Gerald Tuskan: Expression Quantitative Trait Nucleotide (eQTN) Mapping Reveals Transcriptional Regulatory Networks in Populus trichocarpa

PRESENTER: Jay Chen

Gerald Tuskan: Engineering CRISPR/Cas Systems for Genome Editing in Clostridium thermocellum

PRESENTER: Carrie Eckert

Gerald Tuskan: CRISPR-Tool Development in Pseudomonas putida KT2440 for High-Titer Strain Engineering and Multiplexed Approaches

PRESENTER: Jacob Fenster

Gerald Tuskan: Effectiveness of Cotreatment during Fermentation on Various Feedstocks, Impact on Different Microbes, and Changes in Biomass Physical Properties

PRESENTER: Sanchari Ghosh

Gerald Tuskan: Isolation and Rapid Domestication of Non-Model Microbial Hosts for Biofuels Production

PRESENTER: Adam Guss

Gerald Tuskan: Exploiting Nature's Knockouts: How the Pangenome Reveals New Insights into Poplar Phenotypic Variation

PRESENTER: David Kainer

Gerald Tuskan: Aboveground Effects of Beneficial Microbe Co-Culture with Populus Cell Wall Chemistry Variants

PRESENTER: Udaya Kalluri

Gerald Tuskan: Glycosylation is Vital for Industrial Performance of Hyper-Active Cellulases

PRESENTER: Brandon Knott

Gerald Tuskan: Quantitative Network AnalysisQuantitative Network Analysis in Clostridium thermocellum using 13C and 2H-tracers in Clostridium thermocellum using 13C and 2H-tracers

PRESENTER: Travis Korosh

Gerald Tuskan: Investigating the Role of Beneficial Microbes Reducing the Rust Infection Severity in Populus

PRESENTER: Jessy Labbe

Gerald Tuskan: Recombination and Segregation Genomic Patterns in Populus to Facilitate Genomic Selection

PRESENTER: David Macaya-Sanz

Gerald Tuskan: Development and Optimization of Unmanned Aerial Vehicle High-Throughput Phenotyping of Field-Grown Switchgrass

PRESENTER: Reginald Millwood

Gerald Tuskan: Manipulation of Seasonal Dormancy as a Strategy to Improve Switchgrass Biomass Yield

PRESENTER: Ali Missaoui

Gerald Tuskan: Development of Consortia of Beneficial Microbes for Switchgrass and Consequences of Their Deployment on Native Soil Microbiome

PRESENTER: Prasun Ray

Gerald Tuskan: Unraveling New Mechanisms for Lignin Catabolism in Nature

PRESENTER: Davinia Salvachua

Gerald Tuskan: The GWAS Times: A Time-Series Aware GWAS to Detect Natural Climate Adaptations in Arabidopsis and Populus

PRESENTER: Jared Streich

Gerald Tuskan: CRISPR Genome Editing in Populus

PRESENTER: Chung-Jui Tsai

Keith Tyo: Prospecting thiamine diphosphate-dependent carboligases and characterizing their promiscuity to create novel metabolic pathways from primary metabolites

PRESENTER: Bradley Biggs

Jin Wang: Tools for Easy, Fast and Accurate Quantitative Characterization of the Methanotroph-Photoautotroph Coculture

PRESENTER: Kiumars Badr

Jin Wang: Methanotroph-Photoautotroph Coculture – A Flexible Platform for Efficient Biological CO2-CH4 Co-utilization

PRESENTER: Matthew Hilliard

David Weston: Determining the genetic and environmental factors underlying mutualism within a Sphagnum peatmoss microbiome system

PRESENTER: David Weston

Jan Westpheling: Development of emerging model microorganisms: Megasphaera elsdenii for biomass and organic acid upgrading to fuels and chemicals

PRESENTER: Janet Westpheling

Ian Wheeldon: Developing the thermotolerant yeast Kluyveromyces marxianus as a host for next generation bioprocessing
PRESENTER: Ian Wheeldon

Kelly Wrighton: Scaling from genomes to methane emission: targeting critical knowledge gaps in wetland soils
PRESENTER: Adrienne Narrowe

Todd Yeates: Application of a Designed Protein Cage to Increase Pathway Flux of Cellulolysis
PRESENTER: Kevin Cannon

Todd Yeates: Cadmium Toxicity Impacts the Transcriptome and Global Mineral Homeostasis in Chlamydomonas reinhardtii
PRESENTER: Madeli Castruita

Todd Yeates: Orchestrated Response of Dunaliella for Coping with Iron Limitation
PRESENTER: Lital Davidi

Todd Yeates: Advances to Determine the Factors that Govern Stable Protein-display in the Model Gram-positive Bacterium Bacillus subtilis
PRESENTER: Jason Gosschalk

Todd Yeates: LARKSdb: a tool for proteome-wide identification of functional low-complexity domains of proteins
PRESENTER: Michael Hughes

Todd Yeates: Characterization of the Methanosaeta concilii Sheath, a Remarkably Resilient Protein Assembly
PRESENTER: John Muroski

Todd Yeates: Multivariant RNAseq reveal novel players in algae nitrogen metabolism
PRESENTER: Stefan Schmollinger

Todd Yeates: Synthetic Biochemistry: Making Biofuels and Commodity Chemicals the Cell-Free Way
PRESENTER: Saken Sherkhanov

Todd Yeates: Correlation of periodic changes in the transcriptome and proteome in the Chlamydomonas cell cycle
PRESENTER: Daniela Strenkert

Todd Yeates: Cell-free enzymatic biosynthesis of cannabinoids
PRESENTER: Meaghan Valliere

Todd Yeates: A Designed Scaffold for Near-Atomic Resolution Cryo-EM Imaging of Small Proteins
PRESENTER: Todd Yeates

Jamey Young: Systematic debottlenecking of isobutyraldehyde production in cyanobacteria
PRESENTER: Yi Ern Cheah

Jamey Young: Modeling carbon metabolism of the diatom Phaeodactylum tricornutum
PRESENTER: Amy Zheng

Philipp Zerbe: Investigating the switchgrass diterpenoid metabolic network toward improved stress tolerance and biofuel production
PRESENTER: Philipp Zerbe

Huimin Zhao: Transcriptome and metabolome analysis of the oleaginous yeast Rhodosporidium toruloides uncovered new gene targets for metabolic engineering
PRESENTER: Anshu Deewan

Huimin Zhao: A Comprehensive Genome-Scale Model for Rhodosporidium toruloides IFO0880 Accounted for Functional Genomics and Phenotypic Data
PRESENTER: Hoang Dinh

Huimin Zhao: Development of a genetic toolbox for metabolic engineering of Issatchenkia orientalis
PRESENTER: Zia Fatma

Huimin Zhao: Model-guided analysis of the role of OPI1 and RPD3 transcription factors in hexadecanol overproduction strategies
PRESENTER: Shekhar Mishra

Huimin Zhao: Metabolic flux measurement in I. orientalis and R. toruloides
PRESENTER: Yihui Shen

Huimin Zhao: Quantifying Metabolic Enzyme Variation in Non-Model Yeasts
PRESENTER: Arjuna Subramanian

Huimin Zhao: Genome-scale Metabolic Reconstruction of the Non-model Yeast Issatchenkia orientalis SD108
PRESENTER: Patrick Suthers

Huimin Zhao: Oxidative Pentose Phosphate Pathway is the Major Cytosolic NADPH Source in Rhodosporidium toruloides
PRESENTER: Tianxia Xiao

Huimin Zhao: Evaluation of Issatchenkia orientalis as a next-generation chassis for bioproduct production from lignocellulosic biomass
PRESENTER: Yasuo Yoshikuni

Jizhong Zhou: Feedback Responses of Grassland Microbial Communities to Experimental Warming, Precipitation Alternation, and Clipping
PRESENTER: Jizhong Zhou

BioImaging Abstracts

Systems Biology-Based Optimization of Extremely Thermophilic Lignocellulose Conversion to Bioproducts

Diep M. N. Nguyen[1]* (dinguyen@uga.edu), Gabriel M. Rubinstein,[1] James R. Crosby,[2] Ryan G. Bing,[2] Christopher T. Straub,[2] Alex Arzamasov,[3] Ke Zhang,[4] Ying Zhang,[4] Dmitry A. Rodionov,[3] Robert M. Kelly,[2] and **Michael W. W. Adams[1]**

[1]University of Georgia, Athens, GA; [2]North Carolina State University, Raleigh, NC; [3] Sanford-Burnham-Prebys Med. Discovery Institute, San Diego, CA; [4]University of Rhode Island, Kingston, RI

Project Goals: We are using systems biology-guided approaches to develop a non-model, microbial metabolic engineering platform based on the most thermophilic lignocellulose-degrading organism known, *Caldicellulosiruptor bescii,* which grows optimally near 80°C. This work leverages recent breakthrough improvements in the development of molecular genetic tools for this organism, complemented by a comprehensive understanding of its metabolism and physiology gained over the past decade of study in the PIs' laboratories. We are applying the latest metabolic reconstruction and modeling approaches to optimize biomass to product conversion using switchgrass as the model plant and acetone and 3-hydroxypropionate as model industrial products. The over-arching goal is to demonstrate that a non-model microorganism, specifically an extreme thermophile, can be a strategic metabolic engineering platform for industrial biotechnology using a systems biology-based approach.

Bio-processing above 70°C can have important advantages over near-ambient operations. Highly genetically modified microorganisms usually have a fitness disadvantage and can be easily overtaken in culture when contaminating microbes are present. The high growth temperature of extreme thermophiles precludes growth or survival of virtually any contaminating organism or phage. This reduces operating costs associated with reactor sterilization and maintaining a sterile facility. In addition, at industrial scales, heat production from microbial metabolic activity vastly outweighs heat loss through bioreactor walls such that cooling can be required. Extreme thermophiles have the advantage that non-refrigerated cooling water can be used if needed, and heating requirements can be met with low-grade steam typically in excess capacity on plant sites. This project is leveraging recent developments in the PIs' labs for *C. bescii* that enables the proposed effort (1-10). We are developing approaches that provide a comprehensive description of this bacterium's physiology and metabolism to inform metabolic engineering strategies, validate the models with experimental data, and demonstrate that unpretreated lignocellulose can be converted into value-added industrial chemicals at bioreactor scale. The specific aims of this research are: 1) to construct and test a robust metabolic model based on a metabolic reconstruction of *C. bescii* growing on the simple sugars, glucose and xylose, 2) to construct and test a robust metabolic model of *C. bescii* growing on complex biomass-related sugars, cellulose and xylan, 3) to optimize the production of acetone and 3-hydroxypropionate from simple sugars

guided by metabolic modeling, and 4) to demonstrate conversion of cellulose, xylan, cellulose/xylan, and the model biomass switchgrass to acetone and 3-hydroxypropionate.

In the first phase of the project, high temperature chemostat cultures are being used in conjunction with transcriptomic analysis to determine bioenergetic parameters and gene regulation patterns for *C. bescii* growth on lignocellulose-relevant sugars (e.g., glucose, xylose, arabinose, cellobiose, cellulose and xylan). In addition, a novel aspect of *C. bescii* fermentation metabolism has been identified that involve two-step enzymatic process to oxidize glyceraldehyde-3-phosphate during glycolysis (11). These efforts feed into comprehensive metabolic reconstruction and modeling analyses with the ultimate goal of optimizing *C. bescii* production of industrial chemicals from renewable feedstocks.

References

1. Lee, L. L., et al. (2018) Comparative biochemical and structural analysis of novel cellulose binding proteins (Tāpirins) from extremely thermophilic *Caldicellulosiruptor* species. *Appl. Environ. Microbiol.* (doi: 10.1128/AEM.01983-18)
2. Williams-Rhaesa, A. M., et al. (2018) Engineering redox-balanced ethanol production in the cellulolytic and extremely thermophilic bacterium, *Caldicellulosiruptor bescii*. *Metab. Engineering Commun.* **7,** e00073 (doi: 10.1016/j.mec.2018.e00073)
3. Zeldes, B.M. et al. (2018) A synthetic enzymatic pathway for extremely thermophilic acetone production based on the unexpectedly thermostable acetoacetate decarboxylase from *Clostridium acetobutylicum*. *Biotechnol. Bioeng.* 115:2951-2961(doi: 10.1002/bit.26829)
4. Williams-Rhaesa, A.M. et al. (2018) A xylose inducible promoter expands the genetic tools for the biomass-degrading, extremely thermophilic bacterium *Caldicellulosiruptor besci*. *Extremophiles* (doi: 10.1007/s00792-018-1023)
5. Conway, J.M. et al. (2018) Novel multi-domain, multi-functional glycoside hydrolases from highly lignocellulolytic *Caldicellulosiruptor* species. *AIChE J.* 64:4218-4228
6. Conway, J.M. et al. (2018) Parsing *in vivo* and *in vitro* contributions to microcrystalline cellulose hydrolysis by multi-domain glycoside hydrolases in the *Caldicellulosiruptor bescii* secretome. *Biotechnol. Bioeng.* 115:1-15.
7. Williams-Rhaesa, A., et al. (2017) Genome stability in engineered strains of the extremely thermophilic, lignocellulose-degrading bacterium *Caldicellulosiruptor bescii*. *Appl Environ Microbiol.* 83 (doi: 10.1128/AEM.00444-17)
8. Zurawski, J.V., et al. (2017) Bioavailability of carbohydrate content in natural and transgenic switchgrasses for the extreme thermophile *Caldicellulosiruptor bescii*. *Appl. Environ. Microbiol.* 83:e00969-17 (doi: 10.1128/AEM.00969-17)
9. Conway, J.M., et al. (2017) Functional analysis of the Glucan Degradation Locus (GDL) in *Caldicellulosiruptor bescii* reveals essential roles of component glycoside hydrolases in plant biomass deconstruction. *Appl. Environ. Microbiol.* 83:e01828-17 (doi: 10.1128/AEM.01828-17)
10. Lee, L.L., et al. (2018) Genus-wide assessment of lignocellulose utilization in the extremely thermophilic *Caldicellulosiruptor* by genomic, pan-genomic and metagenomic analysis. *Appl. Environ. Microbiol.* e02694-17 (doi: 10.1128/AEM.02694-17)
11. Scott, I. et al. (2019) The thermophilic biomass-degrading bacterium *Caldicellulosiruptor bescii* utilizes two enzymes to oxidize glyceraldehyde-3-phosphate during glycolysis. *J Biol. Chem.* (in revision)

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under Award Number DE-SC0019391

Global Models for Metabolism and Transcriptional Regulation in *Caldicellulosiruptor* Strains

Ying Zhang[1]*(yingzhang@uri.edu), Ke Zhang,[1] Alex Arzamasov,[2] Dmitry A. Rodionov,[2] James R. Crosby,[3] Ryan G. Bing,[3] Diep M. N. Nguyen,[4] Gabriel M. Rubinstein,[4] Robert M. Kelly,[3] and **Michael W. W. Adams[4]**

[1]University of Rhode Island, Kingston, RI; [2]Sanford-Burnham-Prebys Med. Discovery Institute, San Diego, CA; [3]North Carolina State University, Raleigh, NC; [4]University of Georgia, Athens, GA;

Project Goals: We are using systems biology-guided approaches to develop a non-model, microbial metabolic engineering platform based on the most thermophilic lignocellulose-degrading organism known, *Caldicellulosiruptor bescii,* which grows optimally near 80°C. This work leverages recent breakthrough advances in the development of molecular genetic tools for this organism, complemented by a deep understanding of its metabolism and physiology gained over the past decade of study in the PIs' laboratories. We are applying the latest metabolic reconstruction and modeling approaches to optimize biomass to product conversion using switchgrass as the model plant and acetone and 3-hydroxypropionate as products. The over-arching goal is to demonstrate that a non-model microorganism, specifically an extreme thermophile, can be a strategic metabolic engineering platform for industrial biotechnology using a systems biology-based approach.

The thermophilic cellulolytic bacterium *Caldicellulosiruptor bescii* has extensive and highly diversified carbohydrate utilization machinery. A detailed reconstruction of this machinery including uptake mechanisms, biochemical transformations and transcriptional regulation is of key importance for the scope of our DOE-sponsored project involving metabolic engineering of *Caldicellulosiruptor* species for unpretreated lignocellulose conversion to bioproducts. Accurate functional assignment of carbohydrate utilization genes is challenging due to substantial variations of the respective pathways between species including frequent non-orthologous gene displacements and functionally divergent paralogs. A global phylogenomic analysis of 13 *Caldicellulosiruptor* and other thermophilic species in Clostridia positions our target organism at a distant branch from other widely studied model organisms. Therefore, homology-based annotations in existing genomic databases are often incomplete and imprecise. To address this challenge we combine a subsystems-based approach to pathway analysis (implemented in the SEED genomic platform) with *in silico* reconstruction of metabolic networks and transcriptional regulatory networks.

We have obtained a first draft of metabolic reconstruction of *C. bescii* that contains 530 metabolic genes, 787 metabolites (non-unique) and 695 metabolic reactions. The draft reconstruction is represented in a YAML format and curated with support of the PSAMM

software [1] to incorporate all published experimental data and inferences about enzyme activities and substrate specificities. Further, differential gene expression patterns on various carbohydrates will be integrated into the simulation of metabolic activities. To reconstruct transcriptional regulons in thirteen available *Caldicellulosiruptor* genomes we identified previously known regulons (e.g., arabinose regulon AraR) based on ortholog mapping and applied *ab initio* prediction of novel regulons (e.g., fucose regulon FucR) using available transcriptomics data for *C. besci* and *C. saccharolyticus* grown on different mono- and polysaccharides. Reconstructed regulons provide an additional layer of genome context, helping to significantly improve the accuracy of functional annotations and metabolic reconstruction. For instance, the inferred AraR regulon includes a novel functional variant of arabinose isomerase gene, which is non-orthologous to previously characterized enzymes.

Our analysis also revealed substantial differences in sugar catabolic pathways between *Caldicellulosiruptor* and other previously studied bacteria. The repertoire of transporters and regulators involved in sugar catabolism in *Caldicellulosiruptor* demonstrate the most prominent differences in comparison with other taxa. For example, by analyzing the Rex regulon, an NADH-sensing transcriptional regulator, among the thirteen *Caldicellulosiruptor* genomes, we identified conserved core regulon including [FeFe]- and [NiFe]-hydrogenases, various ferredoxin-dependent oxidoreductases and enzymes involved in central carbon metabolism. Using the systems-based approach, we will further identify specificities of these carbon utilization enzymes and provide new information to further complete the reconstruction of metabolic and regulatory networks.

This study illustrates the power of the subsystems-based approach for comparative genomic reconstruction of metabolic and regulatory networks and this will be further extended for the assessment of biofuel production by *C. bescii*.

References

1. Steffensen JL, Dufault-Thompson K, Zhang Y (2016) PSAMM: A Portable System for the Analysis of Metabolic Models. PLoS Comput Biol 12. doi:10.1371/journal.pcbi.1004732.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under Award Number DE-SC0019391

ENIGMA: MAGI: A Method For Metabolite, Annotation, And Gene Integration

Onur Erbilgin[1], **Benjamin P. Bowen[1,3,*] (bpbowen@lbl.gov),** Oliver Rübel[2], Katherine B. Louie[3], Matthew Trinh[1], Markus de Raad[1], Tony Wildish[3,4], Daniel W. Udwary[3,4], Cindi A. Hoover[3], Samuel Deutsch[1,3], **Trent R. Northen[1,3,*],** A.P. Arkin[1,5] and P.D. Adams[1,5]

[1]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory; [2]Computational Research Division, Lawrence Berkeley National Laboratory; [3]Joint Genome Institute, Lawrence Berkeley National Laboratory; [4]National Energy Research Scientific Computing Center, Lawrence Berkeley National Laboratory; [5]University of California, Berkeley, CA.

http://enigma.lbl.gov

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies use a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

Abstract: Metabolomics is a widely used technology for obtaining direct measures of metabolic activities from diverse biological systems. However, it is limited by ambiguous metabolite identifications. Furthermore, interpretation is limited by incomplete and inaccurate genome-based predictions of enzyme activities (*i.e.* gene annotations). Metabolite, Annotation, and Gene Integration (MAGI) attempts to addresses these challenges by generating metabolite-gene associations via biochemical reactions based on a score between probable metabolite identifications and probable gene annotations. This is calculated by a method that emphasizes consensus between metabolites and genes via biochemical reactions. To demonstrate the potential of this method, we applied MAGI to integrate sequence data and metabolomics data collected from *Streptomyces coelicolor* A3(2), an extensively characterized bacterium that produces diverse secondary metabolites. Our findings suggest that coupling metabolomics and genomics data by scoring consensus between the two increases the quality of both metabolite identifications and gene annotations in this organism. MAGI also made biochemical predictions for poorly annotated genes that were consistent with the literature. This limited analysis suggests the potential using metabolomics data to improve annotations in sequenced organisms and also providing testable hypotheses for specific biochemical functions. MAGI is freely available for academic use both as an online tool at https://magi.nersc.gov and with source code available at https://github.com/biorack/magi.

This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

ENIGMA: The selective pressures on the microbial community in a metal-contaminated aquifer

Hans K. Carlson[1,*], Morgan N. Price[1], Mark Callaghan[1], Alex Aaring[2], Romy Chakraborty[2], Hualan Liu[1], Jennifer V. Kuehl[1], Adam M. Deutschbauer[1,4,*], **Adam P. Arkin**[1,3] and **Paul D. Adams**[1,3]

[1]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA; [2]Earth and Environmental Sciences Area, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA; [3]Department of Bioengineering, University of California, Berkeley, CA 94720, USA; [4]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA

http://enigma.lbl.gov

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies mission is to support the development of laboratory and computational tools that link the molecular functions within individual microbes to the integrated activities of microbial communities as they interact with their environment. Our goal is to understand how human activity associated with energetic processes - in particular, contamination and climate change - is affecting the ecology of critical soil, groundwater, and aquifer systems.

Abstract: In many environments, toxic compounds restrict which microorganisms persist. However, in complex mixtures of inhibitory compounds, it is challenging to determine which specific compounds cause changes in abundance and prevent some microorganisms from growing. We focused on a contaminated aquifer in Oak Ridge, Tennessee, U.S.A. that has large gradients of pH and widely varying concentrations of uranium, nitrate and many other inorganic ions. In the most contaminated wells, the microbial community is enriched in the *Rhodanobacter* genus. *Rhodanobacter* abundance is positively correlated with low pH and high concentrations of uranium and 13 other ions and we sought to determine which of these ions are selective pressures that favor the growth of *Rhodanobacter* over other taxa. Of these ions, low pH and high UO_2^{2+}, Mn^{2+}, Al^{3+}, Cd^{2+}, Zn^{2+}, Co^{2+}, and Ni^{2+} are both (a) selectively inhibitory of a *Pseudomonas* isolate from an uncontaminated well versus a *Rhodanobacter* isolate from a contaminated well, and (b) reach toxic concentrations (for the *Pseudomonas* isolate) in the *Rhodanobacter*-dominated wells. We used mixtures of ions to simulate the groundwater conditions in the most contaminated wells and verified that few isolates aside from *Rhodanobacter* can tolerate these 8 ions. These results clarify which ions are likely causal factors that impact the microbial community at this field site and are not merely correlated with taxonomic shifts. Furthermore, our general high-throughput approach can be applied to other environments, isolates and conditions to systematically help identify selective pressures on microbial communities. We are currently extending our approach to identify selective nutrients and carbon sources that influence community composition at our field site. We are leveraging the results of these efforts to classify isolates and determine genetic mechanisms of adaptation to field relevant selective pressures.

References/Publications:

1. Carlson, H. K.; Price, M. N.; Callaghan, M.; Aaring, A.; Chakraborty, R.; Liu, H.; Kuehl, J. V.; Arkin, A. P.; Deutschbauer, A. M. The ISME Journal 2018, 1–13. doi:10.1038/s41396-018-0328-1

Funding Statement: This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies (http://enigma.lbl.gov), a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

ENIGMA: Capturing the Diversity of Subsurface Microbiota – Choice of Carbon Source for Microcosm Enrichment and Isolation of Groundwater Bacteria

Xiaoqin Wu[1], Sarah Spencer[4], Eric J. Alm[4], Jana Voriskova[1], **Romy Chakraborty[1*]** (**RChakraborty@lbl.gov**), A.P. Arkin[2,3] and P.D. Adams[2,3]

[1]Department of Ecology, Earth and Environmental Sciences Area, Lawrence Berkeley National Laboratory, Berkeley, CA; [2]**Lawrence Berkeley National Lab, Berkeley**; [3]University of California at Berkeley; [4] Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA

http://enigma.lbl.gov

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies use a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

Abstract: Improved and innovative enrichment/isolation techniques that yield relevant isolates representing the diversity of environmental microbial communities would significantly advance exploring the physiology of ecologically important taxa in ecosystems. Traditionally, simple organic carbon (C) or yeast extract is used as C source in culture medium for microbial enrichment/isolations. In natural environment however, composition and properties of available natural organic C influences growth and assembly of microbial communities.
In this study, we fed 8 different organic C to subsurface groundwater microbes from Oak Ridge Reservation Field Research Center (ORR-FRC) background site for a 30-day incubation period to investigate the response of indigenous microbial communities to different C. Filtered groundwater was the cultivation 'medium' and was amended with simple organic C (glucose, acetate, benzoate, oleic acid, and cellulose) that are either traditionally used as C source in bacterial culture medium or proxies of C in natural environments; naturally occurring undefined complex C (bacterial cell lysate and sediment-derived natural organic matter (NOM)); as well as a vitamin mixture, commonly used ingredient in culture medium. Our results clearly indicated that natural complex C substrates fared better in enriching diverse bacterial strains compared to other C sources. Microcosms amended with small organic C (glucose, acetate, benzoate, or oleic acid) showed significantly lower biodiversity, dominated by only a few phyla such as *Proteobacteria* and *Bacteroidetes* which are commonly isolated and already have diverse representative isolates. Microcosms amended with natural complex C (cell lysate or NOM) displayed significantly higher biodiversity, three phyla (*Verrucomicrobia*, *Planctomycetes*, and *Armatimonadetes*) that are poorly represented in culture were abundantly enriched. Further, NOM and cell lysate amended enrichments led to cultivation of many unique bacterial strains

across 4 phyla, 13 orders, and 39 genera. Several of these isolates have only 85-90% similarity (by 16S rDNA sequencing) to published strains.

ENIGMA: Management and Analysis of ENIGMA Data using KBase

John-Marc Chandonia[*,1] (JMChandonia@lbl.gov), Pavel S. Novichov[1], Alexey E. Kazakov[1], Adam P. Arkin[1,2], and Paul D. Adams[1,2]

[1]Lawrence Berkeley National Lab, Berkeley; [2]University of California at Berkeley

http://enigma.lbl.gov

Project Goals: ENIGMA (Ecosystems and Networks Integrated with Genes and Molecular Assemblies) uses a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

The overarching goal of data management in ENIGMA is to enable integration of data from complex high-throughput field and laboratory studies towards multi-scale, predictive and systems level understanding of complex microbial community structure, function and evolution. To that end, we must ensure that all data are shared internally between ENIGMA teams, and also with scientists involved in other DOE programs in order to facilitate collaborative research and increase the overall pace of scientific discovery.

To enable more powerful data management and analysis, we have established a strategic collaboration with DOE Systems Biology Knowledgebase (KBase), a fast growing computational infrastructure that provides a unique solution for sharing both data and computational modules/pipelines, conforms to the SC digital data management policy, and supports private data, data sharing and provenance.

KBase has proven to be extremely valuable as an organizational tool for ENIGMA. All ENIGMA data are linked in a series of narratives from a single "master narrative" that is shared with all ENIGMA researchers. We have also helped to test and prototype KBase tools for managing organizations, using these ENIGMA narratives as a starting point.

Several large ENIGMA datasets that are amenable to analysis by KBase tools are stored as objects in KBase. These include the genomes of 297 sequenced isolates from ENIGMA campaigns, amplicon data from 16S surveys of 100 ENIGMA wells, and metagenomic shotgun reads. We use tools and workflows available in KBase to perform various types of bioinformatic analyses, including genome and metagenome assembly, QA/QC, and annotation; RNA-Seq differential expression, etc. We have applied a suite of KBase tools to analyze metagenomic samples obtained from three sediment cores across several projects (ENIGMA Core Pilot and Sediment Core Projects). In addition to being stored in KBase, public ENIGMA datasets are also deposited in appropriate repositories such as NCBI and MG-RAST. Other data that cannot currently be modeled in KBase are stored in accessible locations such as our Fitness Browser (fit.genomics.lbl.gov), or MAGI (magi.nersc.gov). These datasets are currently linked to ENIGMA narratives, and will be ported to KBase objects when such data can be modeled in KBase.

The ENIGMA data management team collaborates with other ENIGMA researchers to port tools and new types of data developed within ENIGMA into KBase. These include the metagenomics tools developed

by the Alm Lab (see poster by Anni Zhang) and metabolic analysis tools such as MAGI (see poster by Ben Bowen). We are also integrating into KBase a tool for functional and taxonomic profiling of shotgun metagenomic datasets, developed by the data management team in collaboration with the Arkin Lab. We monitor new data types and tools as they are introduced by ENIGMA team members, in order to update our priorities appropriately. We prioritize data types and computational tools for integration into KBase according to metrics such as scientific impact, costs, and feasibility of integration.

The ENIGMA data management team has also found that approximately 2/3 of ENIGMA data types would be well represented in KBase using "Generic" data objects that contain experimental measurements, which could be linked in KBase to non-Generic data types that represent biological or environmental objects (e.g., isolates and samples). We prototyped Generic data types, uploaders, graphing tools, search tools, and ontologies in KBase, and are currently collaborating with the KBase project to harden and deploy these technologies for use by ENIGMA team members as well as all other KBase users.

ENIGMA is a Scientific Focus Area Program at Lawrence Berkeley National Laboratory and is supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

ENIGMA: Novel Bio-Signatures and Activity in Fractionated Groundwater from Uncontaminated and Contaminated Sites

H.J. Smith*[1], **M.W. Fields[1,5*]** (matthew.fields@biofilm.montana.edu), L.M. Lui[2,5], I.R. Miller[1], D.C. Joyner[3,5], F. von Netzer[4], A.D. Putt[3], T.C. Hazen[3,5], A.P. Arkin[2,5], and P.D. Adams[2,5]

[1]Montana State University, Bozeman; [2]Lawrence Berkeley National Laboratory, Berkeley; [3]University of Tennessee, Knoxville; [4]University of Washington, Seattle; [5]ENIGMA (http://enigma.lbl.gov/)

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies use a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

Recent work has shown the existence of ultra-small bacteria (100-300 nm) in groundwater but no work has confirmed *in situ* activity, and the ultra-small bacteria could display very different mass transport and activity distributions in porous media flow. Therefore, field observation and characterization are needed to determine the functional role and activity of ultra-small bacteria that could specialize in specific activity and distributions as well as potential metabolic interactions. Moreover, microorganisms that can alter size in response to nutrient levels need to be differentiated from microorganisms that remain small. Total cell numbers, translationally-active cell numbers with bioorthogonal non-canonical amino acid tagging (BONCAT), and microbial activity (^{3}H-Leucine incorporation) were investigated for both low biomass uncontaminated and contaminated groundwater (both 0.2 and 0.1 μm fractions). In addition, metagenomics characterization was performed for the 0.2 and 0.1 μm fractions of both uncontaminated and contaminated groundwater.

In a recent injection experiment of emulsified vegetable oil, the phyla *AC1*, *Gemmatimonadetes*, *OP8*, *WS2*, *WS3*, and *WWE1* displayed notable increases in the 0.1 μm fraction from five shallow wells downgradient of the injection well. These results suggested that both novel and previously observed OTUs could be in the 0.1 μm fraction. However, few cultivars exist for these bacterial phyla.

We then assessed microbial numbers and activities for unstimulated field samples (uncontaminated and contaminated). Total cell numbers (0.2 μm filter) for uncontaminated groundwater (GW271) were 1.3×10^6 +/- 4.4×10^5 cells/ml, whereas, for contaminated groundwater the total cell abundances were 7.3×10^5 +/- 3.4×10^4 and 4.6×10^5 +/- 8.4×10^4 cells/ml for FW115-24 and FW106, respectively. Abundances of smaller cells (0.1 μm filter) were highest for the uncontaminated groundwater 6.40×10^4 +/- 2.1×10^4 cells/ml, while abundances for the contaminated wells were 6.7×10^2 +/- 1.1×10^1 and 6.3×10^2 +/- 6.5×10^1 cells/ml for FW115-24 and FW106, respectively. The results demonstrated that cell numbers for the 0.2 μm fraction were approximately an order of magnitude higher for the uncontaminated

compared to the contaminated groundwater (10^6 v. 10^5). Cell numbers for the small fraction (0.1 μm fraction) were also at least an order of magnitude higher for the GW271 compared to the contaminated groundwater (10^4 v. 10^2).

For uncontaminated groundwater activity, the small cell fraction (0.1 μm) made up almost 20% of total BONCAT activity (per ml), and the small cell fraction had roughly 3-fold greater activity on a per cell basis. When uncontaminated groundwater was compared to the contaminated, there was a drastic reduction in the BONCAT activities and the contaminated groundwater was between 172-769-fold less. Additionally, the rate of leucine incorporation (^{3}H-leucine) on a per cell basis for the 0.2 μm fraction in pristine groundwater was 172 and 8,000 times greater than the contaminated groundwater (FW115 and FW106, respectively). While the overall activity for contaminated wells was low (0.5-2.0 ng C/ml/d), between 25 and 57% of the total activity was from the 0.1 μm fraction. Moreover, for the tested groundwater (uncontaminated and contaminated), the 0.1 μm fraction had higher activity on a per cell basis than the 0.2 μm fraction. Overall, for both size fractions, activity was lower (both per volume and per cell) in contaminated groundwater compared to uncontaminated groundwater. From 21 separate metagenomic assemblies (GW271, FW106, FW115-24), we currently have over 40 genomes with >98% completeness and <2% contamination. We are comparing the presence of these genomes across our samples and previous metagenomics data from the ORR test site. In addition, we are analyzing the potential functional roles of these novel microorganisms. Known species that have been classified as ultramicrobacteria are enriched in the 0.1μm fraction and predictions for potential biochemical capacity are in progress.

This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

ENIGMA: Characterization of Microbial Strains from Contaminated Groundwater and from Contaminated Sediments Using Environmental Concentrations of Metals at the Oak Ridge Reservation

X. Ge[1]*(xiaoxuan.ge25@uga.edu), M.P. Thorgersen[1], F.L. Poole II[1], E.L. Majumder[2], G.M. Zane[2], K.B. De León[2], J. Moon[3], C.J. Paradis[4], F. von Netzer[5], D.A. Stahl[5], J.M. Chandonia[6], P.S. Novichkov[6], A.M. Deutschbauer[6], J.D. Wall[2], T.C. Hazen[4], **M.W.W. Adams[1],** A.P. Arkin[6] and P.D. Adams[6]

[1]University of Georgia, Athens, GA; [2]University of Missouri, Columbia, MO; [3]Oak Ridge National Laboratory, Oak Ridge, TN; [4]University of Tennessee, Knoxville, TN; [5]University of Washington, Seattle, WA; [6]Lawrence Berkeley National Laboratory, Berkeley, CA

http://enigma.lbl.gov

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies use a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

The goal is to identify mechanisms that enable the growth of microbial communities in nitrate-, metal-contaminated and molybdenun-limited environments. A strategy was employed to isolate microorganisms from the ORR contaminated environment using environmentally relevant enrichment conditions. Strains of interest were then selected from the isolates using high throughput growth and biochemical assays. Mechanisms of resistance are being characterized in the strains of interest using genome sequencing, RB-TnSeq, metallomics and metabolomics approaches, in order to understand the role that they play in communities within this special ORR contaminated environment.

Abstract: The contamination plumes surrounding the S-3 ponds at Oak Ridge Reservation (ORR) are an extreme environment that has elevated concentrations of nitrate (up to 233 mM) and multiple metals including Al (up to 20.7 mM), Mn (up to 3.1 mM), U (up to 576 μM), and Ni (up to 157 μM). We hypothesize that microorganisms that live in this hostile environment have nitrate reduction and metal resistance mechanisms with unique properties that enable them to survive here. The goal of this project is to isolate microorganisms under environmentally relevant conditions from ORR contaminated groundwater and sediment, and to characterize the nitrate reduction and metal resistance properties of these isolates. To this end, comprehensive elemental analysis of over 50 elements was performed on two sediment cores, one outside and one within the contamination plumes, revealing differences in concentrations and populations of metals between the two locations. A mix of multiple metals approximating the contaminated environment was then used in a series of high-throughput enrichments to isolate microorganisms from both ORR groundwater and sediment. In depth characterization of selected isolates have uncovered a diversity in pH preference, carbon source utilization, metal tolerance, and nitrate utilization properties between the isolates. This includes metal resistant isolates capable of concentrating multiple metals on the cell surface, and an isolate, XG196, that is capable of

efficiently reducing nitrate when growing in low pM concentrations of molybdenum, a metal required in the catalytic center of nitrate reductase. Further in-depth characterization of the isolates is underway using multiple avenues including the development of a genetic system for MT124, construction of RB-TnSeq libraries for MT58 and other isolates, and genome sequencing of the strains. Also, compilation of the sequencing and characterization data in KBase will aid in further analysis of the isolates for pathways of interest. Finally, to fully understand the role and impact of these isolates at ORR, a cloud-based charting and mapping system called Google Fusion Tables was created by integration of ENIGMA's previous 100-survey topography, geochemical, biological and isolate-specific data streams. This could be used to rapidly analyze isolates, sequencing, and/or geochemical data using custom geospatial maps in a very user-friendly way.

This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies (http://enigma.lbl.gov), a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

ENIGMA: Dissecting Microbial Nitrogen Cycling In The Subsurface Using Tailored Reactor Schemes

Kristopher A. Hunt[1*] **(hunt0362@uw.edu),** Anne E. Otwell[1], Jacob J. Valenzuela[2], Heidi J. Smith[3], Caitlin M. Gionfriddo[4], Regina L. Wilpiszeski[4], Dwayne A. Elias[4], Matthew W. Fields[3], Nitin S. Baliga[2], **David A. Stahl**[1*]**(dastahl@uw.edu)**, Adam P. Arkin[5] and Paul D. Adams[5]

[1]University of Washington, Seattle; [2]Institute for Systems Biology, Seattle, Washington; [3]Montana State University, Bozeman; [4]Oak Ridge National Lab, Oak Ridge, Tennessee; [5]Lawrence Berkeley National Laboratory, Berkeley, California

http://enigma.lbl.gov

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies use a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

The unresolved interplay of biotic and abiotic factors drives material and energy transformations in the subsurface environment. Understanding has been limited by the lack of appropriate model systems to simulate the complex matrix of attached and unattached components of the subsurface. Therefore, the ENIGMA consortium has invested in three types of reactor systems that are designed to simulate different features of the subsurface. Simulations of these features, including fluid dynamics as well as temporal and spatial heterogeneity of processes and community composition, are being used to enrich previously uncharacterized activities and quantify the interplay of governing principles.

Abstract:

Microbial activity in the subsurface is driven by a complex interplay between abiotic and biotic processes and is limited by various nutrients over differing spatial and temporal scales. The dynamic nature of the subsurface environment has made deciphering mass and energy transformations by these activities challenging and thus they remain poorly understood (For more information on nitrogen cycling work see "Metabolomics and Transcriptomics for Environmental Systems Biology: Molecular Mechanisms of Reduced Sulfur Caused Growth Inhibition of Field-Isolated Nitrate-Reducing Bacterium" by Majumder et. al.). Specifically, many governing factors of the nitrogen cycle remain to be identified or characterized, including chemodenitrification and intermediate cycling. The Field Research Center (FRC), near Oak Ridge National Lab in Tennessee, provides a contextually relevant site for examining the effects of nitrogen cycle perturbation in field. Historical activities at the FRC led to the deposition of large quantities of heavy metal laden nitric acid that has resulted in a spatially and temporally dynamic environment of varying pH (4-7) and nitrate concentrations (µg/l to g/l). By comparing microbial activity and community structure across the environments within the FRC and

laboratory simulations performed across ENIGMA, we aim to identify and quantify microbial processes that govern nitrogen cycling in the subsurface.

Physical, chemical, and biotic dynamics of the subsurface create a diverse collection of environments that may select for microbial activities. Environmental simulations using a single reactor type miss key components required to reproduce field phenomena. However, many of the dynamics of the subsurface can be separated using three primary reactor designs: a planktonic-based chemostat, a fluidized bed reactor (FBR), and a packed bed reactor (PBR). Planktonic-based chemostats simulate the pore water of the subsurface, support steady state kinetic analyses, and when implemented in the field capture the impact of dynamic events and perturbations of the incoming community structure. However, planktonic reactors are subject to washout when a population is unable to grow at the operating dilution rate. Packed bed reactors most directly simulate the complex subsurface environment (flow through porous media) and are capable of replicating *in situ* flow rates and providing surfaces for attached populations that may grow slower than that required for planktonic populations. However, these packed bed systems also may be complicated by channeling resulting from cell growth and synthesis of microbial products. Fluidized bed reactors simulate a more homogeneous subsurface environment where particle-attached growth is exposed to a uniform fluidized local microenvironment. These general concepts are being used across ENIGMA to recapitulate environmental phenomena in a variety of reactor constructs, including packed bed reactors, fluidized bed reactors, lysimeters, and CDC-type reactors containing sediment coupons (For more details see "Using in-field bioreactors to monitor microbial community dynamic shifts with geochemical perturbations" by Wilpiszeski et. al.). Major environmental parameters, including shear and impact of changing nutrient availability or nutrient transients can be isolated by comparing these reactor systems to better understand the driving forces of community assembly and mass and energy transformations in the subsurface.

This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

ENIGMA: Core Values: Large-Scale Analysis of Environmental Constraints on Microbial Community Assembly, Activity, and Dispersal in Groundwater and Sediment from a Contaminated Subsurface Aquifer

L.M. Lui*[1] (*lmlui@lbl.gov*), H.J. Smith[2], F. von Netzer[3], K.B. De León[4], E.L-W. Majumder[5], J.V. Kuehl[1], F. Song[1], A. Sczesnak[1,6], T. Nielsen[1], M.P. Thorgesen[7], X. Ge[7], F.L. Poole[7], B.P. Bowen[1], S.M. Kosina[1], C.J. Paradis[8], K.F. Walker[9], K.A. Lowe[10], D.C. Joyner[9], M. Rodriquez, Jr.[10], B. Adams[9], D. Williams[9], J-W. Moon[11], J.D. Van Nostrand[12], D. Ning[12], Y. Fu[12], W. Shi[12], Y. Li[12], D.J. Curtis[12], Y. Fan[12], L. Wu[12], R. Tian[12], G.M. Zane[4], A.B. Aaring[1], X. Wu[1], A.E. Kazakov[1], J-M. Chandonia[1], P.S. Novichkov[1], P.J. Walian[1], R. Chakraborty[1], M.W.W. Adams[7], J. Zhou[12], T.R. Northen[1], J.D. Wall[4], D.A. Stahl[3], D.A. Elias[10], T.C. Hazen[9,10], M.W. Fields[2], **A.P. Arkin[1,6]**, P.D. Adams[1]

[1]Lawrence Berkeley National Lab; [2]Montana State University, Bozeman; [3]University of Washington; [4]University of Missouri, Columbia; [5]Scripps Research Institute; [6]University of California, Berkeley; [7]University of Georgia; [8]Los Alamos National Laboratory; [9]University of Tennessee, Knoxville; [10]Oak Ridge National Lab; [11]National Minerals Information Center, US Geological Survey; [12]University of Oklahoma

http://enigma.lbl.gov

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies use a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

Abstract: Subsurface microorganisms play important roles in mediating major biogeochemical cycles, but only in recent years have studies shed light on their population structure, biogeography, and metabolisms present. The Department of Energy ENIGMA Scientific Focus Area seeks to map the causal interactions that constrain microbial community assembly and dispersal in chemically and physically complex environments. Recently, we initiated a study of microbial communities in the shallow subsurface of a contaminated aquifer at the Oak Ridge Field Research Center, a site of nuclear weapon development during the Manhattan Project.

We hypothesize that strong gradients of pH, heavy metals, nitrate, and other contaminants at the site influence the distribution, structure, and activity of microbial communities. To study community assembly mechanisms, we performed large-scale analysis of two sediment cores and associated groundwater for which we produced depth-index data sets of physical, chemical, bulk biological and sequencing measurements. One core is considered uncontaminated (466 cm), and the other core (815 cm) is contaminated by chemicals from man-made processes. We divided the cores into ~23 cm segments for processing, resulting in 56 segments which allowed us to do a finer-grained analysis of the vertical transect as compared to other subsurface studies.

The contaminated core is much less diverse as 250 exact sequence variants (ESVs) from 16S amplicon sequencing account for 50% of observed reads as compared to 660 from the uncontaminated core, suggesting strong selective pressure from contamination. Initial analysis of ESV location suggests that there is little mixing and dispersal along a core. In general, there is little overlap in ESVs between the two

cores (~300-350 meters apart). Many of the chemical and physical measures are strongly dependent on depth and highly colinear; these colinear groups are strongly predictive of the ESV group distribution and their measured activity although not all observed groups are well predicted. Currently, we are processing shotgun metagenomics data to compare taxonomy and genes to our other data. From the water and sediment data, we have over 50 genomes with >98% completeness and <2% contamination. We are tracking these genomes through the vertical length of the core and analyzing the potential functional roles of these organisms. This study integrates over 12 measures of microbial community composition, activity, and environmental controls to provide new insights into subsurface microbial communities.

Funding statement: ENIGMA is a Scientific Focus Area Program at Lawrence Berkeley National Laboratory and is supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

ENIGMA: Metabolomics and Transcriptomics for Environmental Systems Biology: Molecular Mechanisms of Reduced Sulfur Inhibition of Field-Isolated Nitrate-Reducing Bacterium

E. L.-W. Majumder[3*](emajumder@scripps.edu), A.O. Otwell[4,5], A.V. Carr[4], M. Watkins[4], S. Turkarslan[4], D.C. Vuono[5], D.A. Stahl[5], **G. Siuzdak[3]**, N.S. Baliga[4], A.P. Arkin[1,2], and P.D. Adams[1]

[1]Lawrence Berkeley National Lab, Berkeley; [2]University of California at Berkeley; [3]The Scripps Research Institute, La Jolla, CA; [4]Institute for Systems Biology, Seattle, WA; [5]University of Washington, Seattle

http://enigma.lbl.gov

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies use a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

This work explores the field observation of exclusion of nitrate and sulfate reducing bacteria on the molecular level in the field isolate *Intrasporangium calvum* while demonstrating the incorporation of multi-omic monitoring in environmental systems.

Abstract:
The mutual exclusion of sulfate-reducing bacteria (SRB) and nitrate-reducing bacteria (NRB) from discrete depths was observed in a sediment bore hole experiment based on the taxa and the respiration activity measured at the Oak Ridge Field Research Center (FRC) (Pilot 2017 experiment). Targeted isolations were carried out from various regions of the FRC, and a nitrate-reducing bacterium that aligned with *Intrasporangium calvum* was isolated. Characterization of growth of the isolate on nitrate-reducing conditions noted growth inhibition when reduced-sulfur containing molecules were added to the medium. Both cysteine and sulfide inhibited growth, but with distinct phenotypes. The mechanisms of this growth inhibition were investigated using various physiological assays coupled with transcriptomics and metabolomics. Data analysis is still in progress, but we currently have one proposed mechanism. From global pathway mapping of overlaid metabolomics and transcriptomics data, we saw that branched chain amino acid biosynthesis was downregulated in the cysteine inhibited samples. To confirm that cysteine inhibition of branched chain amino acid biosynthesis was responsible for the cysteine samples growth phenotype, we performed a feed-in experiment with branched-chain amino acids and found the addition rescued the growth inhibition phenotype. We are continuing experiments in order to test other observed dysregulated pathways (genes and metabolites) based on transcriptomics and metabolomics data. The data are revealing on a molecular level how the end-products of SRB respiration inhibit and therefore exclude NRB in a field-relevant bacterium (non-model organism system). In other ENIGMA posters presented here, we are using reactor studies for larger scale view of these processes (see Hunt et al. 'Dissecting microbial nitrogen cycling in the subsurface using tailored reactor schemes' and 'Using in-field bioreactors to monitor microbial community dynamic shifts with geochemical perturbations' by Wilpiszeski et al.).

Funding Statement: This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies (http://enigma.lbl.gov), a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

ENIGMA: Assembly Mechanism of Subsurface Microbial Community under Stress Gradient and Adaptation of Super Phylum *Patescibacteria* with Genome Simplicity

D. Ning[1]* (ningdaliang@ou.edu), R. Tian[1], J.D. Van Nostrand[1], L. Wu[1], P. Zhang[1], W. Shi[1], L. Wu[1], Y. Zhang[1], Y. Yang[1], D.J. Curtis[1], Z. He[1], M.B. Smith[2], A.M. Rocha[3], C.S. Smillie[2], S.W. Olesen[2], C.J. Paradis[4], J.H. Campbell[3,9], J.L. Fortney[4], T.L. Mehlhorn[3], K.A. Lowe[3], J.E. Earles[3], J. Phillips[3], S.M. Techtmann[4], D.C. Joyner[4], D.A. Elias[3], K.L. Bailey[3], R.A. Hurt Jr.[3], S.P. Preheim[2], M.C. Sanders[2], J. Yang[2], M.A. Mueller[3], W.A. Lancaster[5], B.J. Vaccaro[5], F.L. Poole II[5], M.W. Fields[7], E.J. Alm[2], T.C. Hazen[3,4], M.W.W. Adams[5], P.D. Adams[6,8], A.P. Arkin[6,8], J. Zhou[1,6]

[1]**The University of Oklahoma**; [2]Massachusetts Institute of Technology; [3]Oak Ridge National Laboratory; [4]University of Tennessee; [5]University of Georgia; [6]Lawrence Berkeley National Laboratory; [7]Montana State University; [8]University of California at Berkeley; [9]Northwest Missouri State University.

http://enigma.lbl.gov/

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies mission is to support the development of laboratory and computational tools that link the molecular functions within individual microbes to the integrated activities of microbial communities as they interact with their environment. Our goal is to understand how human activity associated with energetic processes - in particular, contamination and climate change - is affecting the ecology of critical soil, groundwater, and aquifer systems.

Community assembly mechanism under a stress gradient is a fundamental question in ecology. Classical stress gradient theory discussed the roles of different deterministic forces under different degrees of stresses, however we still know little about the relative importance of various stochastic and deterministic processes along stress gradients. The groundwater in the Oak Ridge Integrated Field Research Challenge site (FRC, Oak Ridge, TN) has large geochemical gradients and has been comprehensively surveyed, providing a rare opportunity to examine ecological processes and drivers shaping subsurface microbial diversity. Groundwater samples were taken from 98 wells that covered the geochemical diversity across the site. The 16S rDNA were sequenced for all samples by Illumina MiSeq and metagenomic shotgun sequencing were performed for 12 representative samples by Illumina HiSeq. We applied various approaches to disentangle the ecological processes controlling community assembly. Since obvious limitation of previous methods, we developed two new approaches to infer community assembly mechanisms. The first one is a general framework to estimate ecological stochasticity with a new index, normalized stochasticity ratio (NST). Another one is a quantitative framework to estimate relative influence of major ecological processes, respecting the fact that different microorganisms can be under different assembly mechanisms. We tested the index with simulated communities and demonstrated substantial improvement of quantitative

and qualitative performance of the new approaches. We applied these new approaches to investigate how ecological stochasticity and different ecological processes varied along stress gradients at FRC. The results suggested a decrease of ecological stochasticity with the increase of environmental stress, consistent with multivariate analysis. The influence of heterogeneous selection sharply increased around 3 fold from low-stress to high-stress wells, related to chemical stresses imposed by abnormal pH, nitrate, carbon limitation, uranium, and some other metals. Dispersal limitation appeared as the most influential process (42% on average), corresponding to the strong limitation of microbial migration in groundwater system. Dispersal limitation becomes less important under higher stresses, in accordance with the connectivity of groundwater among highly contaminated wells, and thus significantly correlated with pH and some metals in the supernatant rather than pellet-associated metals. In addition to community assembly mechanism variation, we also observed members from an uncultivated superphylum *Patescibacteria*, which was found prevalent in subsurface environments, but the mechanisms for its prevalence are not known. The genomic features and metabolisms of this super phylum were investigated through genome-resolved metagenomics analysis. While the members of *Patescibacteria* had reduced genomes (~1 Mbp) exclusively, our results demonstrated they retained functions essential to growth and reproduction such as genetic information processing. Surprisingly, they have sharply reduced redundant and unessential functions of metabolisms, cellular activities and stress responses. The *Patescibacteria* have ultra-small cell sizes and simplified membrane structures including flagellar assembly, transporters and two-component systems. Despite the lack of CRISPR viral defense, the bacteria could have alternative strategies to evade predation such as lacking phage receptors in membranes, which may explain the lack of phage-related proteins detected in the genomes. By establishing the linkages between bacterial features and the groundwater environmental conditions, we noticed that the adaptation of *Patescibacteria* to the environment could drive the features of a reduced genome, ultra-small cell size and lack of CRISPR viral defense. Altogether, the contamination stress gradient significantly switched the microbial community assembly mechanism and the special environment of FRC groundwater led to adaption of the superphylum *Patescibacteria* with genome simplicity.

This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies (http://enigma.lbl.gov), a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

ENIGMA: New technologies for the large-scale screening of microbial interactions

Markus de Raad[1,*], **Fangchao Song[1,*](fsong@lbl.gov)** Tim Veth[1], Leqian Liu[2], Estelle Couradeau[1], Jiahao Shi[1], Adam R. Abate[2], Adam M. Deutschbauer[1], Trent R. Northen[1,3,†], Adam P. Arkin[1,4,†], Paul D. Adams[1]

[1]Environmental Genomics and Systems Biology, Lawrence Berkeley National Laboratory, Berkeley, CA 94720; [2]University of California at San Francisco, San Francisco, CA 94158; [3]Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720; [3]University of California at Berkeley, Berkeley, CA 94720

http://enigma.lbl.gov

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies use a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

Abstract: Environmental microbial community is normally composed of thousands of species and affected by a wide variety of biotic and abiotic factors, which make untangling all interactions an herculean task. A frequently used strategy to reduce the complexity is, similar like untangling a ball of yarn, to start at a single point, a single interaction between two microbes. Although this can be very insightful, the implications of this single interaction on the total community may be impossible to predict. A reversed approach, where a large number of microbes are screened on their interactions would be a better way to identify possible relationships. Here we present two novel platforms for the large-scale screening of microbial interactions. In one method, acoustic printing is used to create paired colonies in agar filled multi-well plates, which allows to make visual observations of interactions such as changes in colony size, shape, texture and pigmentation. Metabolite usage and exchange during the interactions are screened by mass spectrometry. In another method, taking advantages of droplet generation rate of 20 million per hour, the microbes are coupled into small drops, enriched, barcoded, and eventually sequenced in one pool. The composition, abundance, and cell interactions are revealed through the high throughput sequencing. Both new techniques allow us obtain thousands to millions interactions in days, which will make mapping out microbial community structure possible.

Funding Statement: This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies (http://enigma.lbl.gov), a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

ENIGMA: Long-Term Continuous Monitoring Gives Insight into Patterns Between Groundwater and Weather Events

Kathleen F. Walker[1*](kfitzge4@vols.utk.edu), K.A. Lowe[2], B.G. Adams[3], E.R. Dixon[4], D.C. Joyner[4], M. Rodriguez, Jr.[2], M.W. Fields[6], D.A. Elias[2] and **T.C. Hazen**[1,2,3,4], A.P. Arkin[5], and P.D. Adams[5].

[1]Bredesen Center, University of Tennessee, Knoxville, [2]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, [3]Department of Earth and Planetary Sciences, University of Tennessee, Knoxville, [4]Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, [5]Lawrence Berkeley National Laboratory, Berkeley, CA, [6]Center for Biofilm Engineering, Montana State University, Bozeman, MT,

http://enigma.lbl.gov/research/environmental-microbiology/

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies use a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

The main goals of this project are to use long-term continuous monitoring to understand the fluctuations seen in the microbial communities and geochemistry of groundwater over time, including diurnal, seasonal and annual time scales.

Abstract: To take an environmental systems approach to studying the microbial communities found in groundwater and sediment, we began long-term continuous monitoring as a way to understand how ecosystem phenomena were playing a role in these microscopic environments. To do this, we deployed two down-well groundwater monitoring units (In-Situ AquaTroll600) and set up two meteorological stations (HOBO U30) within contaminated and uncontaminated research field sites at Y-12 National Security Complex in Oak Ridge, Tenn.
After monitoring the groundwater and weather at these sites, we have seen interesting trends within the data. There are natural diurnal fluctuations in the groundwater elevation, relative humidity and air temperature. We also see increases in groundwater elevation and changes in groundwater geochemistry (dissolved oxygen, pH, conductivity) following heavy rain events. However, there is a distinct delay between the detected rainfall and a rise in the groundwater elevation. We will begin monthly sampling of groundwater in the spring to detect the fluctuations in microbial community and geochemistry in these sites. Our ENIGMA team members are developing a sampling plan for sediment coring in the coming months. With the continued use of these units and the collection of groundwater and sediment, we are hoping to capture the extended, especially annual and seasonal, relationships between all of these parameters and their effect on the microbial community. This data is also being used to guide future scientific focus on meteorological events of interest, seasonal focus for sampling and geochemical variations over time.

Funding Statement: This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies (http://enigma.lbl.gov), a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

ENIGMA: Using in-field bioreactors to monitor microbial community dynamic shifts with geochemical perturbations

Regina L. Wilpiszeski[1*] (wilpiszeskir@ornl.gov), Caitlin M. Gionfriddo[1*], Ann M. Wymore[1], Ji-Won Moon[1], Kenneth A. Lowe[1], Mircea Podar[1], Dominique C. Joyner[1], Craig C. Brandt[1], Anthony V. Palumbo[1], Matthew W. Fields[2], Terry C. Hazen[1,3], Judy D. Wall[4], Nitin S. Baliga[5], David A. Stahl[6], Michael W.W. Adams[7], Farris Poole III[7], Romy Chakraborty[8], Yupeng Fan[9], Joy D. Van Nostrand[9], Jizhong Zhou[9], **Dwayne A. Elias[1+]**, Adam P. Arkin[8,10], Paul D. Adams[8,10]

[1]**Oak Ridge National Lab**; [2]Montana State University; [3]University of Tennessee; [4]University of Missouri; [5]Institute for Systems Biology; [6]University of Washington; [7]University of Georgia; [8]Lawrence Berkeley National Lab; [9]University of Oklahoma;[10]University of California at Berkeley

http://enigma.lbl.gov

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies use a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

Abstract

Subsurface microbial communities mediate the transformation and fate of redox-sensitive compounds including organic matter, metals and radionuclides. Few studies have explored how changing geochemical conditions influence the composition of groundwater microbial communities. As part of the ENIGMA Environmental Ark campaign, we used 1L in-field bioreactors receiving background and contaminated well water from the Department of Energy site at the Oak Ridge Reservation, TN to test the effect of abiotic forces on microbial community structure. (For additional reactor studies see "Dissecting microbial nitrogen cycling in the subsurface using tailored reactor schemes" by Hunt et. al., and "Metabolomics and Transcriptomics for Environmental Systems Biology: Molecular Mechanisms of Reduced Sulfur Caused Growth Inhibition of Field-Isolated Nitrate-Reducing Bacterium" by Majumder et. al.) Planktonic and biofilm microbial communities were initialized with background water to establish communities in triplicate control reactors and triplicate test reactors. All were fed filtered water from the background site for 18 days. On day 18, three reactors were switched to filtered water from a contaminated well, enriched in total dissolved solids relative to the

background site, particularly chloride, nitrate, uranium, and sulfate. Biological and geochemical data were collected throughout the experiment, including planktonic and biofilm DNA for 16S rRNA amplicon sequencing, cell counts, total protein, anions, cations, trace metals, organic acids, bicarbonate, pH, Eh, DO, and conductivity. We observed significant shifts in both planktonic and biofilm microbial communities receiving contaminated water. This included a loss in diversity, especially amongst members of the *Bacteroidetes*, *Acidobacteria*, *Chloroflexi*, and *Betaproteobacteria*, but enrichment in the Fe(III)-reducing *Ferribacterium* and parastitic *Bdellovibrio*. These shifted communities were more similar to the contaminated well community, suggesting that geochemical influences on microbial community structure are substantial.

This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

ENIGMA: Building a reference-based metagenomics workflow in KBase

An-Ni Zhang[1*] (anniz44@mit.edu), Shijie Zhao[1] and Eric J. Alm[1], A.P. Arkin[2,3] and P.D. Adams[2,3]

[1]Massachusetts Institute of Technology, Cambridge; [2]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory; [3]University of California, Berkeley, CA.

Project Goals: ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies uses a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

Here we aim to develop a powerful and user-friendly tool for the analysis of genetic and evolutionary traits by integrating metagenomics and genomes. The tools developed will be embedded into the KBase platform and demonstrated using ENIGMA data.

Abstract: As sequencing becomes less expensive, researchers are turning from 16S rRNA surveys to shotgun sequence metagenomics in order to add new levels of functional and phylogenetic resolution to their sequence-based analyses. Metagenomic data harbors additional layers of data on population structure, strain dynamics, and genome evolution that cannot be inferred from 16S alone. Powerful and user-friendly tools for the analysis of these data are not yet widely available. We believe population genetic and evolutionary data analysis tools made available via the KBase platform will have an outsize impact on environmental microbiology research.

Here we report five new functions that we will add to the KBase environment to catalyze metagenomic data analysis. First, we have built a standard and comprehensive set of reference genomes to which metagenomic reads can be compared. We have designed the pipeline to compare metagenomes to references and tested it in our samples. We are now building the estimators of strain level diversity, and even inference of strain genomes. We will design the tools to study within-population genome rearrangements and mutations. Finally, we will design a new statistical approach that would allow data generated by different researchers using different protocols to be compared on equal footing.

In this work, we are collaborating with the ENIGMA data management team (John-Marc Chandonia) and the KBase team (Dylan Chivian). The ENIGMA data we are initially using are Pseudomonas genomes from Lauren M. Lui (Arkin Lab).

This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231.

New advances in diatom functional genomics: Gene knockout methodology and genome-wide mapping of transcription factor binding sites

Smith, S.R.,[1] McCarthy, J.K.,[1] Niu, D.,[1] Tan, M.,[1,2] Moosburner, M.,[1,2] Schulberg, A.,[1,2] and **Allen, A.E.[1,2*] (aallen@jcvi.org)**

[1]J. Craig Venter Institute, San Diego, CA, [2]University of California, San Diego, Scripps Institution of Oceanography, La Jolla, CA

Diatoms, which comprise 20-40% of all phytoplankton, thrive throughout the world's oceans. They are single-celled eukaryotes, drawing down atmospheric CO_2 and producing oxygen by photosynthesis. Phytoplankton biomass anchors the base of the ocean food chain, and their storage of excess energy as lipids drives the algal biofuel and food industries. To build biomass and lipids diatoms take up bicarbonate and nitrogen, in the form of nitrate. The uptake and assimilation of nitrate by diatoms is a tightly-coordinated, energy-demanding, multi-protein, multi-organelle effort. To describe the passage and transformation of nitrate through the cell, transgenic lines for twelve primary, nitrate-related genes have been produced in the marine model-diatom, *Phaeodactylum tricornutum* by CRISPR-Cas9 knockout methods. The genes chosen were identified by annotation and prior transcriptomic, time-course expression studies. CRISPR knockouts of two genes encoding outer membrane nitrate transporters (Phatr3_J26029 and J54560), tentatively described by comparative analysis as low and high sensitivity nitrate transporters, show distinctly different growth profiles. The CRISPR knockout line for nitrate reductase (Phatr3_J54983), which catalyzes the first step nitrate assimilation, does not grow on nitrate, whereas its growth ammonium is unchanged. Changes in growth and/or fluorescence (RFUs) have been tracked for the individual CRISPR knockouts for the coupled enzymes: GSII and GOGAT in the chloroplast (Phatr3_J51092 and J24739) and the GSIII and GOGAT(m) in the mitochondria (Phatr3_J22357 and J20342); here, phenotypic changes, while apparent, are more subtle. Within the chloroplast, growth phenotypes for CRISPR-knockout lines for two nitrite reductases (Phatr2_J12902 and Phatr3_EG02286) indicate potential diel roles in nitrite reduction. Previous research suggests that diatoms may store nitrate in excess of their immediate needs; CRISPR-knockout lines for two putative vacuolar nitrate transporters (Phatr3_EG01952 and J28245) have been shown to grow more slowly than WT cells. Experiments in progress may to determine if sluggish growth is directly related to nitrate uptake into the vacuole. In bacteria and plants, external and intra-cellular sensing of nitrate abundance enable cells to regulate transcriptomic responses; a CRISPR-knockout line for a nitrate sensing kinase (Phatr3_J21961), identified by homology in *P. tricornutum,* points to possible interaction with one of the outer membrane nitrate transporters.

Additionally, diatoms change patterns of gene expression as they adjust to environmental conditions, such as shifting nutrient availability or light levels. Transcription factors (TF) bind specific DNA sequences to regulate the expression of nearby genes, making them important components of this cellular response. In the model diatom *Phaeodactylum tricornutum,* little is known about which genes are regulated by specific TFs, as only a few of the 212 annotated TFs have been functionally investigated. Improved knowledge of TFs and their binding sites (TFBS) is essential to construct a systems level model of gene expression and is also valuable for enhancing the molecular toolbox for genetic engineering of diatoms. Using a new, low-cost,

high-throughput technique called DAP-seq (DNA affinity purification sequencing) we are mapping transcription factors binding sites genome-wide. Initial rounds of DAP-seq have been successfully completed, on 30 TFs selected on the basis of their transcriptional sensitivity to nutrient perturbations (Fe, N). We will expand this effort to include all annotated TFs in the *P. tricornutum* genome, greatly expanding functional knowledge of DNA regulatory elements and transcriptional regulation.

Supported by DE-SC0018344, Biological and Environmental Research

Subcellular compartmentalization and biochemistry of lysine biosynthesis in the model diatom Phaeodactylum tricornutum

Bielinski, V.,[1] Dupont, C.L.,[1] **Allen, A.E.[1,2*](aallen@jcvi.org)**

[1]J. Craig Venter Institute, San Diego, CA, [2]University of California, San Diego, Scripps Institution of Oceanography, La Jolla, CA

Diatoms and other chlorophyll c (Chl c)-containing algae dominate photosynthetic eukaryotic communities in most sunlit aquatic environments, yet our knowledge concerning overall cellular metabolism in photosynthetic eukaryotes is largely based on work from a few well-studied model green algae and vascular plants. Thus, there is a genuine need for model organism development for Chl c lineage microalgae. A critical challenge in the development of models of cellular metabolism for Chl c algae is the characterization and incorporation of these metazoan and bacterial genes and pathways into traditional green algae models. One example of this is the diatom ornithine urea cycle (OUC), a metabolic module previously found only in metazoans. Analysis of available diatom genomes suggests the presence of a DAP-dependent lysine biosynthesis pathway utilizing aminotransferase activity for production of LL-DAP precursor similar to that in land plants. *Phaeodactylum tricornutum* employs a chimeric lysine biosynthetic pathway in that appears to be 1) a hybrid of recently acquired bacterial genes with the canonical plant pathway, 2) begins and ends in the mitochondria and chloroplast, 3) terminated by a diaminopimelate decarboxylase (DAPDC) and, 4) ubiquitous in diatoms. We have validated the recombinant protein product of the gene Pt21592 as a DAP decarboxylase that produces L-lysine *in vitro* and possesses a unique substrate promiscuity profile.

Supported by DE-SC0018344, Biological and Environmental Research

Developing episome-based, gene expression modulation by exogenous chemicals in the diatom *Phaeodactylum tricornutum*.

Tessema Kassaw[1], Graham Peers[1], and **Andrew Allen**[2]
[1]Colorado State University, Fort Collins, CO; [2]J. Craig Venter Institute, La Jolla, CA

Project Goals: Overall goal - Reprogram metabolic networks using *in vivo* synthetic modules to increase the flux of energy and carbon into biofuel precursors. Goal 1) Profiling the transcriptome, proteome and metabolome to investigate cell responses to physiologically relevant conditions. Goal 2) Identify and manipulate key factors involved in the control of inorganic C assimilation, photosynthetic efficiency and regulation of lipid accumulation. Goal 3) Forward genetic library generation, screening and genotyping. These approaches complement our development of *Phaeodactylum* genome reconstruction /modeling and our development of novel synthetic genomic tools to achieve our overall goal of increasing productivity.

CRISPR/Cas9 technology has been successfully implemented for targeted genome editing and gene expression changes across a wide range of eukaryotic organisms. This project aims at engineering inducible CRISPRi machinery into an artificial chromosome within the model diatom *Phaeodactylum tricornutum*. The current state of the art for inducible gene expression in diatoms is based on endogenous promoters that respond to different environmental conditions. For instance, in *P. tricornutum*, the most commonly used inducible promoters are the light dependent pLHCF1 and the nitrate dependent pNR promoters. However, the use of these endogenous promoters could have unwanted consequences on cell physiology. This calls for the development of new transcriptional control systems. In this project, we are developing a library of orthogonal, inducible systems which are based on naturally occurring, heterologous chemically responsive transcription factors and cognate promoters. We identified and adapted six synthetic expression systems that have proven effective in other eukaryotic systems for use in *P. tricornutum*. These include the *10xN1/NEV* (4-hydroxytamoxifen inducible) system, the *XVE/OlexA* (β-estradiol inducible) system, the *pOp/LhGR* (dexamethasone inducible) system, the *AlcR/AlcA* (ethanol inducible) system, the *Tet-ON* (tetracycline or doxycycline inducible) system, and *DIG/pUAS* (Digoxin or Digoxigenin inducible) systems. Since the six genetic circuits are modular, genetic parts such as transcription factors, promoters, and terminators can be replaced to create a library of novel inducible systems. We will first evaluate the ability these inducible systems to induce the production of YFP reporter protein and adapt their use to induce gene silencing using the CRISPRi system. Developing externally controlled CRIPSRi technology for efficient and scalable disruption of gene expression in diatoms will provide powerful tools to study gene function, to enable HTP genetic screens and to modify specific pathways to produce high-value metabolites.

This research is supported by the Office of Biological and Environmental Research in the DOE Office of Science – grant # DE-SC0008595

Interactive Metabolic Pathway Visualization for Optimization of Energy Flux

Zachary King[1]* (zaking@ucsd.edu), Avi Kumar[1], Christian M. Metallo[1], **Andrew E. Allen[2]**, and **Karsten Zengler[1]**

[1] University of California, San Diego; [2] J. Craig Venter Institute, San Diego

The goal of this project is to gain new insights into the photosynthetic eukaryote *Phaeodactylum tricornutum* using a systems biology approach. To analyze and contextualize the omics data and predictive metabolic models developed for the project, new interactive, data-rich, web-based visualizations are being developed. The Escher pathway visualization tool (https://escher.github.io) will be used as a starting point, and it will be extended to represent compartmentalization in the diatom, to visualize ^{13}C isotope tracing data, to visualize time series data, and to give non-computational users the ability to explore and modify models.

To meet the visualization and data analysis needs of genome-scale models (GEMs), such as the diatom reconstruction, we developed Escher (https://escher.github.io), a web application for visualizing data on biological pathway maps. Through this project, a new version of Escher is being developed that improves the performance, reliability, and extensibility of the metabolic pathway maps. The new version includes visual representation of cellular compartments, and it is the basis of three visualization applications that will enable systems-level optimization of *Phaeodactylum tricornutum*.

First, a new visualization package called Escher-Trace has been developed through a collaboration with Prof. Christian Metallo's research group at UCSD. Escher-Trace makes ^{13}C isotope tracing data more accessible by improving data presentation and streamlining analysis. Users can upload containing isotope tracing data in common file formats (JSON, CSV) to visualize them on the Escher metabolic map. Once loaded, users can automatically generate publication quality graphs of mass isotopomer distribution, metabolite abundance, mole percent enrichment, and kinetic labeling and abundance. Escher-trace corrects for natural isotope abundance of ^{13}C labeled datasets and provides full control over data normalization and sample/condition grouping. Next, Escher-Trace will be incorporated with the metabolic pathway map of *Phaeodactylum tricornutum* to analyze new ^{13}C tracing datasets.

Second, we are developing methods for time-series visualization on the Escher map, particularly for gene expression data. The resulting map will allow animated playback of datasets through time with precise controls and the ability to export publication-ready figures.

Finally, a web-based collaboration platform is under development that will allow non-computational users to expand metabolic models with new heterologous pathways. This platform will take advantage of Escher and the interactive flux simulator Escher-FBA (https://sbrg.github.io/escher-fba) for rapid prototyping of new diatom strain designs.

Supported by DE-SC0018344, Biological and Environmental Research

The Plastid Terminal Oxidase does not play a photoprotective role in the marine diatom *Phaeodactylum tricornutum*.

Graham Peers[1*](graham.peers@colostate.edu), Denis Jallet,[1] Maxwell Ware,[1] and **Andrew Allen**[2]
[1]Colorado State University, Fort Collins, CO; [2]J. Craig Venter Institute, La Jolla, CA

Project Goals: Overall goal - Reprogram metabolic networks using *in vivo* synthetic modules to increase the flux of energy and carbon into biofuel precursors. Goal 1) Profiling the transcriptome, proteome and metabolome to investigate cell responses to physiologically relevant conditions. Goal 2) Identify and manipulate key factors involved in the control of inorganic C assimilation, photosynthetic efficiency and regulation of lipid accumulation. Goal 3) Forward genetic library generation, screening and genotyping. These approaches complement our development of *Phaeodactylum* genome reconstruction /modeling and our development of novel synthetic genomic tools to achieve our overall goal of increasing productivity.

Algae, cyanobacteria and plants require a dynamic photoprotective system to maintain high photosynthetic efficiencies in dynamic light environments. These systems need to be rapidly induced in excess light to reduce the incidence of photooxidative stress. Conversely, they need to be turned off to maintain high rates of photochemistry in lower light. The Plastid Terminal Oxidase (PTOX) has been shown to play an important role in photoprotection in the green alga Chlamydomonas. PTOX accepts electrons from the photosynthetic electron transport chain and donates them to reduce oxygen to water. This alleviates excitation pressure and increases fitness during growth in excess light. We investigated the role of PTOX in the marine diatom Phaeodactylum tricornutum by creating a double knockout of 2 PTOX isoforms using a CRISPR-Cas9/HDR approach. The mutants displayed a distinct rise in chlorophyll a fluorescence following intense illumination, suggesting it plays a role in dark chlororespiration. We found no changes in fitness between the WT and KO strains in a variety of stress conditions including excess constant light, nitrogen limitation, carbon limitation, or variable light. We also found no changes in light-dependent oxygen consumption (LDOC) between the WT and KOs in any of these conditions. The rates of LDOC remained low in all culture conditions, which is contrast to previous observations of cyanobacteria and green algae. Other processes, such as non-photochemical quenching or perhaps mitochondria-plastid coupling, appear to be the predominant forms of photoprotection in this diatom. This work suggests that increasing energy fluxes to carbon fixation through engineering PTOX levels is not viable in this species.

This research is supported by the Office of Biological and Environmental Research in the DOE Office of Science – grant # DE-SC0008595

Dynamic metabolic constraints of *Phaeodactylum tricornutum* reveal bases on organelle-specific carbon and nitrogen partitioning during nitrogen depletion

Juan D. Tibocha-Bonilla[1], Cristal Zuniga[1*] (crzuniga@eng.ucsd.edu), **Andrew E. Allen[2]**, and **Karsten Zengler[1]**

[1]University of California, San Diego; [2]J. Craig Venter Institute, San Diego

The goal of this project is to gain new insights into the photosynthetic eukaryote *Phaeodactylum tricornutum* using a systems biology approach. We will integrate various omics data into a metabolic modeling framework, to systematically identify and quantify the partitioning of carbon and nitrogen among cellular metabolism, cross-talk between organelles, and productive photosynthetic electron flow.

Diverse conditions, i.e. growth during day and night, and compartmental cellular organization require phototrophs to shift their proteome demands and therefore adjust their metabolism and biomass composition during the course of growth. The complex interplay between energy and carbon metabolism and its dynamics in phototrophs is still not fully understood. Constraint-based modeling is a systems biology tool that takes advantage of experimental data, such as uptake rates and biomass composition, for successful prediction of growth phenotypes. Currently, lack of time-course biomass composition data has restricted prediction accuracy, by forcing the models to assume that the biomass remains constant. Here, we used experimentally determined metabolomics data to determine biomass composition constrains for a genome-scale metabolic model of the diatom *P. tricornutum*. We found that time course constraints can be the sole driving force that causes the metabolic network to exhibit certain behaviors such as time-specific secretion rates, cross talk of organelles by the activation of the mitochondria, as well as activation of specific metabolic pathways. A growth rate sensitivity analysis of time-course flux distributions enabled identifying the main metabolite affecting growth. We also found that simulation results are sensitive to the accurate experimental quantification of some amino acids. Surprisingly growth sensitivity predictions are independent of biosynthetic cost and connectivity of metabolites.

Supported by DE-SC0018344, Biological and Environmental Research

Fungal Trait Tradeoffs in a Southern Californian Grassland

Charlotte J. Alster[1*] (charlotte.alster@uci.edu), Sydney I. Glassman[1,2], **Steven D. Allison**[1], **Adam C. Martiny**[1], **Jennifer B.H. Martiny**[1], and **Kathleen K. Treseder**[1]

[1]University of California, Irvine, CA; [2]University of California, Riverside, CA

Project Goals: Integrating functional responses of soil microorganisms with genomic information is a fundamental challenge in understanding decomposition dynamics, especially with global change altering our predictions. Trait-based approaches offer a way to overcome this challenge by connecting organismal phenotypic responses with genomic information. In this project we focus on linking fungal trait responses under temperature and moisture manipulations to explore how fungal species and their resulting ecosystem functions are structured by tradeoffs among traits. Specifically, we measured growth yield, resource acquisition, and drought tolerance traits to identify potential tradeoffs in resource allocation under environmentally stressed conditions. We will incorporate these fungal trait responses into the DEMENT model to better predict decomposition rates under drought conditions.

Abstract Text:

Fungi play a critical role in decomposition in terrestrial ecosystems, with major consequences for ecosystem processes and biogeochemical cycling. Yet, little is known about how fungal traits will respond to climate change and how that will impact terrestrial biogeochemical cycling. To better understand the mechanisms involved in fungal stress tolerance and connect these traits to functional differences to carbon and nitrogen cycling, we measured fungal trait responses under climate stress from several diverse strains of fungi in Southern California. Because microbial communities and their resulting functions are structured by tradeoffs among traits, we asked the driving question of what tradeoffs exist for fungal traits subjected to environmental change.

To test this question, we designed an experiment that manipulated moisture and temperature and measured three fungal traits: growth yield, resource acquisition, and drought tolerance. We chose these traits because they are relevant traits for decomposition and valuable overall indicators of microbial growth and survival. In order to isolate the effects of the different fungi, we sterilized grassland litter, and then inoculated the litter with 15 different fungal strains (Ascomycota and Basidiomycota) isolated from the same site. These microsocms were then subjected to one of nine different moisture and temperature combinations in the lab for five weeks. We measured CO_2 production rate, fungal hyphal biomass, and four different extracellular enzymes at the end of the incubation period to calculate the fungal traits. Growth yield was calculated as the inverse of mass specific respiration (a function of fungal hyphal biomass and CO_2 production rate), resource acquisition was determined through the extracellular enzyme assays, and drought tolerance was calculated as the relationship between fungal hyphal biomass and moisture level.

Our analyses revealed some distinct fungal trait tradeoffs. First, we found that as growth yield increases, there is less enzyme activity ($P < 0.01$; $R^2 = 0.64$), indicating a tradeoff between resource acquisition and growth yield (Fig. 1). Second, we also found a tradeoff between resource acquisition and drought tolerance (Fig. 1), with more enzyme activity correlating with less drought tolerance ($P < 0.001$; $R^2 = 0.73$). These results suggest that metabolic investment into resource acquisition is made at the expense of growth and investment towards drought tolerance. Third, we did not find a tradeoff between growth yield and drought tolerance (Fig. 1), since lower growth yield was correlated with less drought tolerance ($P < 0.001$; $R^2 = 0.75$); this perhaps indicates that growth yield is coupled with traits selecting for drought tolerance. Our work simultaneously improves mechanistic understanding of fungi in ecosystems and advances trait-based modeling, which is part of additional work from our group under this funding opportunity, in order to further predict fungal impact on biogeochemical cycling.

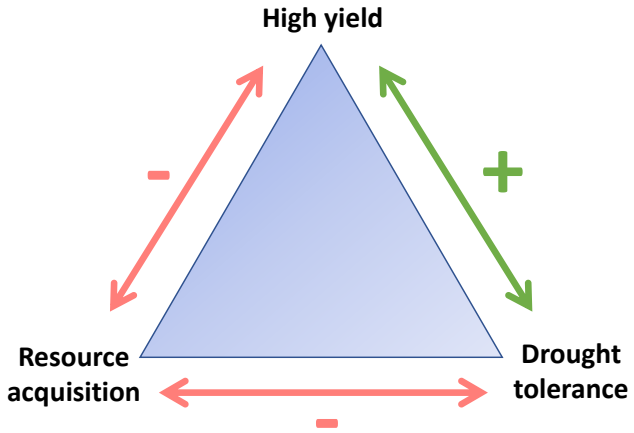


Figure 1. Schematic indicating the positive and negative relationships between high yield, resource acquisition, and drought tolerance traits observed in our study for soil fungal species.

Physiological adaptations of leaf litter microbial communities to drought reduce decomposition rates

Ashish A. Malik[1*] (a.malik@uci.edu), Tami Swenson[2], Claudia Weihe[1], Eric Morrison[1], **Jennifer B. H. Martiny**[1], **Eoin L. Brodie**[2,3], **Trent Northen**[2,4] and **Steven D. Allison**[1]

[1]University of California, Irvine; [2]Lawrence Berkeley National Laboratory, Berkeley; [3]University of California, Berkeley; [4]DOE Joint Genome Institute, Walnut Creek

Project Goals: Predictions of carbon cycle processes under environmental change are highly uncertain. Integration of molecular data with process rate measurements remains a challenge, thereby making it difficult to link microbial diversity and carbon cycling processes. We aim to uncover the genomic and cellular mechanisms that underlie microbial traits important for carbon cycling under drought. Traits related to drought tolerance, resource acquisition, and growth yield will be measured with molecular and biochemical approaches. In addition, we will quantify the consequences of these tradeoffs for litter carbon loss. The goal is to implement modeling approaches that use trait relationships as a mechanistic basis for predicting how microbial communities and carbon cycling processes will respond to drought, a critical environmental change driver.

Drought is known to reduce organic matter decomposition in terrestrial ecosystems by causing a decline in microbial growth. This decline is partially attributed to increased microbial investment in physiological adaptations to tolerate or avoid drought stress. However, we still lack a detailed understanding of the stress adaptations that enable microbial decomposers to survive and thrive in such environments. Here we present metatranscriptomic and metabolomic insights into the physiological response of in situ plant leaf litter microbial communities to long term drought and pulse wetting in Californian grass and shrub ecosystems. Litter bags with native litter were deployed mid-summer and *in situ* wetting was performed at the end of the dry season to simulate commencement of seasonal precipitation. Legacy of reduced precipitation (long-term simulated drought treatments) leads to decreased decomposition rates in grass but not in shrub litter, and overall decomposition rates were higher in grass than shrub litter. Taxonomic and functional responses of active communities to pulse wetting and subsequent drying were smaller compared to the legacy effect of drought and vegetation type. Fungi dominated both litter communities. In shrub communities, functional diversity was higher than in grass communities but there were only small differences in gene expression profiles in response to either pulse wetting or long-term drought treatment. However, in grassland drought and ambient communities had significantly distinct functional profiles, and genes for membrane transporters and flagellar motility were upregulated in response to wetting while those coding for biomolecular repair were downregulated. The most significant discernable physiological adaptations to drought across both litter

communities were production or uptake of various compatible solutes like trehalose and ectoine as well as inorganic ions to maintain cellular osmotic balance. Drought communities also increased expression of genes for capsular and extracellular polymeric substance synthesis possibly as a mechanism to retain water. The magnitude of stress tolerance traits was much bigger in grass than shrub communities. These stress adaptations highlight the metabolic tradeoffs with growth that lead to reduced decomposition under drought, thereby directly linking microbial physiology to ecosystem function.

Figure: Impact of drought legacy on microbial physiology. Frequency of significant transcript indicators in the upper level functional groups (level 1 in KEGG Subsystems classification). These indicators were unique to communities in either ambient or reduced precipitation treatments across both grass and shrub litter types and therefore represent global functional indicators of drought or control ecosystems.

Microbial Communities Exhibit Resilient Extracellular Enzyme Activity Along a Climate Gradient

Bahareh Sorouri[1*] (bsorouri@uci.edu), **Steven D. Allison**[1]

[1]University of California, Irvine, USA

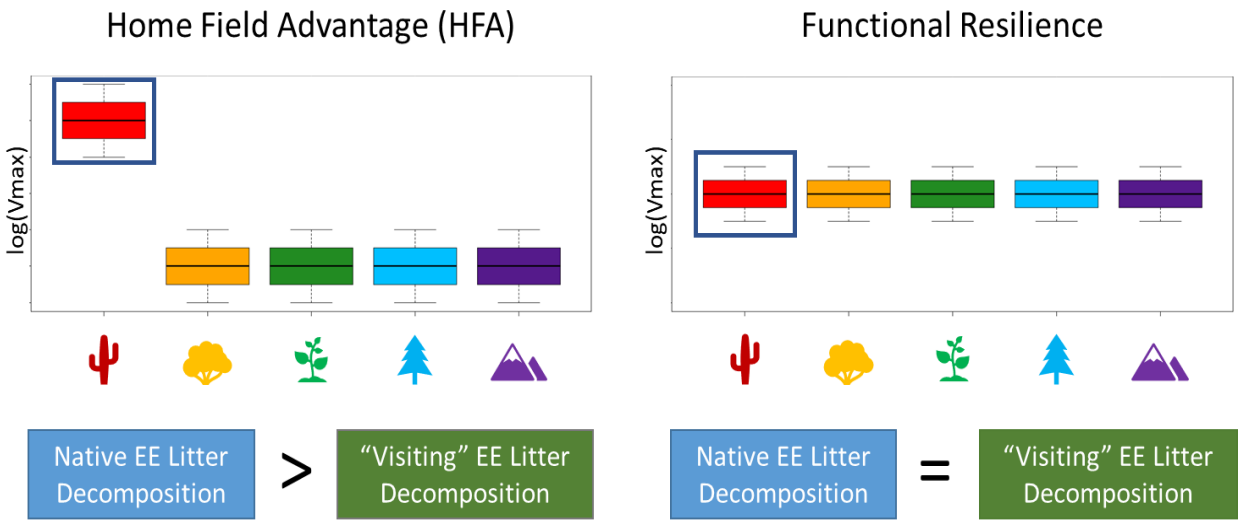
http://allison.bio.uci.edu/projects/a-trait-based-framework-for/

Project Goals: The effects of climate change on litter decomposition rates and carbon cycling are uncertain. Environmental microbes significantly contribute to both processes; however, the physiological constraints faced by the microbes from environmental stress is not fully understood. We aim to understand the cellular mechanisms and biochemical activities of environmental microbes under stressful environmental conditions simulating future climate change. By analyzing microbial community extracellular enzyme activity, we can infer the effects of changing climate on litter decomposition and thereby carbon cycling. The goal is to test the dependency of litter decomposition on microbial communities, and to predict the effects of future climate change on microbial activity.

Understanding the consequences of environmental change for the microbial regulation of carbon and nutrient cycling is a critical need. We collected data on extracellular enzyme activities from an 18-month experiment where microbial communities were reciprocally transplanted along a Southern California elevation gradient with inverse variation in temperature and precipitation. The microbial communities were from desert, scrubland, grassland, pine-oak, and sub-alpine ecosystems. By simulating different temperature and precipitation changes for ecosystem-specific microbial communities, this study was designed to predict microbial community response and resilience to future environmental changes. Our first hypothesis proposed a "home field advantage" (HFA) of enzyme activity and litter decomposition, where microbial communities in their home ecosystem outperform microbial communities transplanted from other ecosystems. Our alternate hypothesis states that transplanted microbial communities will exhibit functional resilience by producing enzymes and decomposing litter at the same rate as native microbial communities when exposed to temperature and precipitation changes. Microbial community extracellular enzyme activities were evaluated with analysis of variance. Enzyme datasets yielding a significant ($p < 0.05$) site by microbial community interaction were further analyzed with Tukey post hoc comparisons within each site to compare the extracellular enzyme activities associated with different microbial communities. Our results partially supported the microbial community resilience hypothesis with respect to extracellular enzyme activity. Significant interaction effects did not support our HFA hypothesis and instead indicated a home field disadvantage, where the transplanted microbial community displayed higher extracellular enzyme activity than the native community. Our findings suggest that transplanted microbes are resilient to climate change, which may be due to their previous adaptations to extreme temperatures and drought. These enzymatic

patterns are consistent with litter decomposition rates, which also did not exhibit home field advantage.

Figure 1: Conceptual diagram of home field advantage (HFA) and functional resilience hypotheses. With HFA, extracellular enzyme (EE) activities of native microbial communities should be significantly greater than those of transplanted microbial communities. Alternatively, our functional resilience hypothesis states that the enzyme activities of all microbial communities are the same, regardless of microbial origin. The x-axis indicates the desert, scrubland, grassland, pine-oak, and sub-alpine microbial communities, respectively. Bold lines within the boxplot are the replicate averages for each microbial community. Boxplots encompassed in a rectangle denote the home field microbial community (desert in this case).



Investigating the Natural Variation of Pennycress Metabolome, an Emerging Crop for Aviation Biofuel.

Cintia Arias[1], Tyler Swanson[1], Fan Yang[2], Erich Grotewold[3], Ajay Shah[2], and **Ana Paula Alonso[1]*** (Anapaula.Alonso@unt.edu),

[1]University of North Texas, Denton, TX; [2]The Ohio State University, Columbus, OH; [3]Michigan State University, East Lansing, MI

Project Goals: The main objective of this research is to apply a biodesign strategy for improving oil content in a promising alternative source of jet-fuel, pennycress. To advance towards this goal, we are: 1) Investigating pennycress natural variation to identify candidate genes and biomarkers associated with oil accumulation and fatty acid composition; 2) Identifying targets to improve oil content and composition, and 3) Establishing metabolic engineering targets and develop community resources.

Bioenergy crops, which have potential for jet fuel production and do not compete with food crops, are urgently needed. Our strategy to address this fundamental challenge is to develop pennycress (*Thlaspi arvens*e), a member of the Brassicaceae, as a bioenergy crop, taking advantage of its ability to produce seed oil that is ideally suited as a renewable source of biodiesel and aviation fuel[1]. Moreover, pennycress performs well on marginal land, has a short maturity time and grows off-season, serving as a winter cover crop, and complementing the production of commodity food crops. Essential pennycress molecular and genetic resources can be rapidly developed by leveraging the fully sequenced genome and research tools of its close relative *Arabidopsis thaliana*. Furthermore, a draft genome[2] for *Thlaspi arvense*, as well as transcriptome[3] and metabolome[4] information have been made recently available. However, for this plant to become an economically viable and sustainable source of jet fuel, molecular and genetic resources need to be developed, and integrated with multi-variable techno-economical analyses to guide strategies for increase oil production through breeding and/or genetic manipulation. These are the gaps that this project intends to fill.

The goal of this part of the project is to find biomarkers positively or negatively correlating with oil content to boost the breeding and/or metabolic engineering process. For this purpose, we studied the natural variation in the metabolome of 11 pennycress accessions from around the world. Biomass components–oil, protein, starch and cell wall–were sequentially extracted and quantified from two different stages of embryos development. The quantification of intracellular metabolites such as sugars, amino acids, organic acids and phosphorylated compounds was achieved by liquid chromatography tandem mass spectrometry as previously described[4-6]. The present poster focuses on integrating the results of these analyses and attempts to highlight the correlations between certain intracellular metabolites and oil composition and accumulation. Ongoing experiments are expanding this study to a total of 20 pennycress accessions. The long-term goal of this research is to use the biomarkers that we identified to enhance oil production in pennycress using rational breeding and/or metabolic engineering.

References

1. Fan, J.Q., Shonnard, D.R., Kalnes, T.N., Johnsen, P.B. & Rao, S. A life cycle assessment of pennycress (Thlaspi arvense L.) -derived jet fuel and diesel. *Biomass & Bioenergy* **55**, 87-100 (2013).
2. Dorn, K.M., Fankhauser, J.D., Wyse, D.L. & Marks, M.D. A draft genome of field pennycress (Thlaspi arvense) provides tools for the domestication of a new winter biofuel crop. *DNA Res* **22**, 121-31 (2015).
3. Dorn, K.M., Fankhauser, J.D., Wyse, D.L. & Marks, M.D. De novo assembly of the pennycress (Thlaspi arvense) transcriptome provides tools for the development of a winter cover crop and biodiesel feedstock. *Plant Journal* **75**, 1028-1038 (2013).
4. Tsogtbaatar, E., Cocuron, J.C., Sonera, M.C. & Alonso, A.P. Metabolite fingerprinting of pennycress (Thlaspi arvense L.) embryos to assess active pathways during oil synthesis. *Journal of Experimental Botany* **66**, 4267-4277 (2015).
5. Cocuron, J.C., Anderson, B., Boyd, A. & Alonso, A.P. Targeted metabolomics of *Physaria fendleri*, an industrial crop producing hydroxy fatty acids. *Plant and Cell Physiology* **55**, 620-633 (2014).
6. Cocuron, J.C., Tsogtbaatar, E. & Alonso, A.P. High-throughput quantification of the levels and labeling abundance of free amino acids by liquid chromatography tandem mass spectrometry. *Journal of Chromatography A* **1490**, 148-155 (2017).

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0019233.

Techno-economic, Energy and Greenhouse Gas Emissions Analyses of Pennycress Production and Logistics for Aviation Biofuel

Seyed Hashem Mousavi-Avval[1], Ajay Shah[1*](shah.971@osu.edu), Erich Grotewold[2], and **Ana Paula Alonso**[3]

[1]The Ohio State University, Wooster, OH; [2]Michigan State University, East Lansing, MI; [3]University of North Texas, Denton, TX

Project Goals: The main objective of this research is to apply a biodesign strategy for improving oil content in a promising alternative source of jet-fuel, pennycress. To advance towards this goal, we are: 1) Investigating pennycress natural variation to identify candidate genes and biomarkers associated with oil accumulation and fatty acid composition; 2) Identifying targets to improve oil content and composition, and 3) Establishing metabolic engineering targets and develop community resources.

Alternative biobased feedstocks for the production of renewable jet-fuel (RJF) is gaining traction in recent years due to the concerns associated with adverse environmental effect of fossil fuel use in aviation industry and concerns over the energy security[1]. To be a viable alternative, RJF should be economically competitive, have lower environmental footprint, and provide a net energy gain[2]. In addition, it should be producible in large quantities without reducing food supplies[3]. Until now several alternative feedstocks, including canola, camelina, soybean, carinata, and jatropha, have been considered as potential sources for the production of RJF[1]; however, commercial production of RJF from renewable feedstocks is still encumbering due to competition of feedstocks with food resources, high cost of production, and use of limited resources and land. Pennycress (*Thalaspi arvense*) is a winter annual crop with 25-36%w/w oil content[4-5]. It can be planted as cover crop in corn-soybean rotation in Midwestern US to provide both economic opportunity and ecosystem services. In addition, pennycress oil, with high content of unsaturated fatty acids, has acceptable quality for conversion to RJF[6-7].

The goal of this part of the project is to analyze techno-economics, energy use and greenhouse gas emissions of pennycress based agronomic and supply systems to establish targets for pennycress oil metabolic engineering. To accomplish this goal, we evaluated the technical feasibility by estimating the resources (land, infrastructure, machineries, fuel, labor, consumables) requirements for the pennycress production (planting, fertilizer and pesticide applications), harvest and post-harvest logistics (grain handling, transportation, drying and storage) to provide feedstock for a RJF plant in Ohio with production capacity of 5 million gallons per year (~19 million liter per year). Then, we estimated the costs, energy use and greenhouse gas emissions associated with these resources and infrastructure use. We estimated that annual harvest of 90,000-115,000 metric tons (t) of pennycress seeds from 41,000-63,000 hectares of land is needed to meet the biorefinery pennycress seed demand. Cost for production and logistics was estimated to be ~200 $/t, which is ~50% less than the cost of camelina production in Oregon[8]. The energy use and greenhouse gas emissions during the production and logistics of pennycress feedstock were ~3,300 MJ/t, and ~400 $kgCO_2eq/t$, respectively. The cost, energy use and greenhouse gas emissions for pennycress production and logistics were found to be highly sensitive to the pennycress seed yield, which is currently ranging from

1,200 to 2,500 kg/ha[1,5], and oil content. Although achieving these baseline estimates is promising, further reduction in total production cost, energy requirement and greenhouse gas emissions would be possible by improving the pennycress seed yield, oil content and level of unsaturated fatty acids. Improving these characteristics will advance large-scale production of pennycress as a sustainable feedstock for the development of RJF biorefineries[9].

References

1. Fan, J., Shonnard, D. R., Kalnes, T. N., Johnsen, P. B. & Rao, S. A life cycle assessment of pennycress (Thlaspi arvense L.) -derived jet fuel and diesel. Biomass and Bioenergy 55, 87–100 (2013).
2. Hill, J., Nelson, E., Tilman, D., Polasky, S. & Tiffany, D. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. PNAS 103, 11206–11210 (2006).
3. Farrell, A. E., Plevin, R. J., Turner, B. T., Jones, A. D., O'hare, M. & Kammen, D. M. Ethanol Can Contribute to Energy and Environmental Goals. Science. 311, 506–508 (2006).
4. Moser, B. R., Knothe, G., Vaughn, S. F. & Isbell, T. A. Production and Evaluation of Biodiesel from Field Pennycress (Thlaspi arWense L.) Oil. (2009).
5. Carr, P. M. Potential of fanweed and other weeds as novel industrial oilseed crops. p. 384-388. In: J. Janick and J.E. Simon (eds.), New crops. Wiley, New York. (1993).
6. Phippen, W. B. & Phippen, M. E. Soybean Seed Yield and Quality as a Response to Field Pennycress Residue. Crop Sci. 52, 2767 (2012).
7. Dose, H. L., Eberle, C. A., Forcella, F. & Gesch, R. W. Early planting dates maximize winter annual field pennycress (Thlaspi arvense L.) yield and oil content. Ind. Crops Prod. 97, 477–483 (2017).
8. Jaeger, W. K. & Siegel, R. Economics of oilseed crops and their biodiesel potential in Oregon's Willamette Valley. Spec. Rep. 1081 18–30 (2008).
9. Phippen, W. B. & Phippen, M. E. Seed Oil Characteristics of Wild Field Pennycress (Thlaspi arvense L.) Populations and USDA Accessions. Poster Present. Adv. Ind. Crop. Annu. Meet. Washington, DC. Available at: http//www.wiu.edu/pennycress/current-experiments/Pennycress seed-oil 2013.pdf (Date accessed: 2019-01-12) (2013)

This research was supported by the DOE Office of Science, Office of Biological and environmental Research (BER), grant no. DE-SC0019233.

In vivo thermodynamic analysis of metabolic networks and engineered pathways

Tyler B. Jacobson[1], Paul A. Adamczyk[1], David M. Stevenson[1], Travis Korosh[1], Daniel G. Olson[2], Lee R. Lynd[2], Jennifer Reed[1], and **Daniel Amador-Noguez*[1]** (amadornoguez@wisc.edu)

[1]University of Wisconsin-Madison, Madison, WI; [2]Dartmouth College, Hanover, NH

Project Goals: This project will integrate advanced mass spectrometry, computational modeling, and metabolic engineering to develop and experimental-computational approach for the *in vivo* genome-scale determination of Gibbs free energies (ΔG) in metabolic networks suitable for high-throughput thermodynamic profiling of engineered organisms and emerging model systems.

Thanks to continuing breakthroughs in genome editing and metabolic engineering, the range of industrially useful organisms that can transform renewable biomass resources into biofuels or bioproducts will continue to expand. Thus, there is an increasing need for genome-scale, high-throughput tools to characterize metabolic capabilities of engineered organisms, identify new routes for bioproduct synthesis, and uncover the foundational principles that drive these complex biological systems.

Thermodynamic analysis of metabolic networks has emerged as a powerful new tool for pathway design and metabolic engineering whose full potential still remains to be realized. Thermodynamics constrains the kinetics of biochemical reactions and determines enzyme efficiency in metabolic pathways. Specifically, a pathway with a strong thermodynamic driving force will achieve larger flux with less total enzyme than a pathway closer to thermodynamic equilibrium. It will also be less susceptible to product feedback inhibition, thereby reaching higher final titers. Thermodynamic analysis can therefore provide unique insights in synthetic pathway design by pinpointing the enzymes whose expression will have the largest effect on flux, by identifying kinetic and thermodynamic bottlenecks in engineered pathways, or by predicting the most efficient metabolic route for product biosynthesis while ruling-out unfavorable ones. Although the usefulness of thermodynamic analysis in pathway engineering is now widely recognized, we currently lack an experimental framework for high-throughput thermodynamic profiling of metabolic networks. This project will develop an experimental-computational approach for the *in vivo* genome-scale determination of Gibbs free energies (ΔG) in metabolic networks suitable for high-throughput thermodynamic profiling of engineered organisms and emerging model systems. This project will result in the construction of experimentally-derived models that quantitatively define trade-offs between energy efficiency of biosynthetic pathways and their overall catalytic rates.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Early Career Research Program under Award Number DE-SC0018998.

Genetic improvement of seed yield and oil content in field pennycress, a non-food oilseed feedstock species

Ratan Chopra[1*] (rchopra@umn.edu), Katherine Frels[1], M. David Marks[1], John Sedbrook[2] and **James A. Anderson**[1]

[1]University of Minnesota, St Paul, MN; [2]Illinois State University, Normal, IL

Project Goals: Our goal is to identify and characterize lines having traits that will improve pennycress efficiency and utility as a biofuel feedstock species and make the seed easier for producers to handle, namely: 1) Increased seed size and 2) Increased seed oil content. To advance towards this goal, we are: 1) Investigating the genetic control of these traits in wild-germplasm collections using quantitative trait loci (QTL) and association mapping; 2) Identifying and characterizing EMS-induced lines for these traits using high-throughput NIRS and seed analyzer screening in combination with genomics tools; 3) Generate CRISPR-Cas9 knockouts in the genes known and found to regulate these traits in Brassica species.

Pennycress is a winter-hardy cover crop that provides ecosystem services such as reduced soil erosion and nutrient loss in between fall corn harvest and spring soybean planting. Unlike traditional cover crops, field pennycress produces a mature oilseed in late spring, allowing farmers to harvest two cash crops in one year. Wild-derived pennycress lines have been shown to yield 1,500 kg/ha on average (Mitich, 1996). Pennycress seeds contain approximately 33% oil by weight, and the oil is an excellent biofuels feedstock. After pressing, the remaining pennycress meal can be used as animal feed. However, despite these environmental and economic benefits, pennycress is currently limited by its small seed size (1 mg/seed), which can complicate planting, harvesting, and handling of the seed. Increasing pennycress seed size would also increase the efficiency of oil extraction. In addition to improving the seed size, increased oil content in the seed would improve the economics of growing and processing of pennycress as a biofuels feedstock. With previous USDA NIFA funding, we developed several EMS-induced pennycress mutant lines exhibiting key domestication traits such as reduced seed pod shatter, earlier flowering, and improved fatty acid profiles (Chopra et al., 2018a; McGinn et al., 2018). We have also developed and demonstrated the utility of pennycress Agrobacterium-mediated plant transformation and CRISPR-Cas9 genome editing, generating pennycress lines with undetectable levels of erucic acid in seed oil (McGinn et al., 2018). With pennycress domestication well under way, our new goal is to identify, characterize, and introgress into breeding lines traits that will improve pennycress efficiency and utility as a biofuel feedstock species and make the seed easier for producers to handle.

The goals of this project are to identify the genetic controls that positively or negatively regulate seed size and seed oil content of pennycress. For this purpose, we have started developing necessary populations required to perform QTL mapping in F_5 or F_6 generations. We currently are propagating $F_{3:4}$ plants for seed increase. We have more than 15,000 M_3 mutant families derived through EMS mutagenesis and are available to screen for these traits. To effectively characterize these populations, we are optimizing and calibrating phenotypic tools such as a Marvin seed analyzer and NIRS (Chopra et al., 2018b) for high-throughput screening of the pennycress seeds. Preliminary analysis of EMS populations has helped us identify a number of lines showing variation in seed oil content and seed size. These lines were planted again to verify their inheritance of the trait. With previous USDA NIFA funding, we have generated several lines for reduced seed fiber content using EMS mutagenesis and CRISPR-Cas9 editing. Interestingly, *transparent testa 8* (*tt8*) alleles from these events have shown ~10% increase in total seed oil content; we are currently verifying these results in field trials. This poster will highlight the approaches we are using to identify lines and corresponding genetic markers for the above-mentioned traits. The long-term goal of this research is to use different genetic tools to identify mutations and natural variants along with molecular markers to enhance seed size and seed oil content of pennycress for introgression into elite breeding lines.

This research was supported by the Plant Feedstock Genomics for Bioenergy: A Joint Research Funding Opportunity Announcement USDA, DOE, grant no. DE-FOA-0001857.

KBase: The Systems Biology Knowledgebase for Predictive Biological and Environmental Research in an Integrated Data Platform

Adam P. Arkin[1], **Robert Cottingham**[3], **Chris Henry**[2], Benjamin Allen*[3] (allenbh@ornl.gov), Jason Baumohl[1], Jay Bolton[1], Shane Canon[1], Stephen Chan[1], John-Marc Chandonia[1], Dylan Chivian[1], Mikayla Clark[3], Paramvir Dehal[1], Meghan Drake[3], Janaka N. Edirisinghe[2], José P. Faria[2],, Annette Greiner[1], Tianhao Gu[2], James Jeffryes[2], Marcin P. Joachimiak[1], Roy Kamimura[1], Keith Keller[1], Vivek Kumar[5], Sunita Kumari[5], Miriam Land[3], Sebastian Le Bras[1], Zhenyuan Lu[5], Akiyo Marukawa[1], Sean McCorkle[4], Cheyenne Nelson[1], Dan Murphy-Olson[2],, Erik Pearson[1], Gavin Price[1], Priya Ranjan[3], William Riehl[1], Boris Sadkhin[2], Samuel Seaver[2], Alan Seleman[2], Gwyenth Terry[1], James Thomason[5], Doreen Ware[5], Pamela Weisenhorn[2], Elisha Wood-Charlson[1], Shinjae Yoo[4], Qizhi Zhang[2], Diane Zheng[1]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA; [2]Argonne National Laboratory, Argonne, IL; [3]Oak Ridge National Laboratory, Oak Ridge, TN; [4] Brookhaven National Laboratory, Upton, NY; [5] Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

http://kbase.us

Project Goals: The Department of Energy Systems Biology Knowledgebase (KBase) is a knowledge creation and discovery environment designed for both biologists and bioinformaticians. KBase integrates a large variety of data and analysis tools, from DOE and other public services, into an easy-to-use platform that leverages scalable computing infrastructure to perform sophisticated systems biology analyses. KBase is a freely available and developer extensible platform that enables scientists to analyze their own data within the context of public data and share their findings across the system.

The U.S. Department of Energy (DOE) supports biological and environmental research to investigate the complex interactions within biological systems and the processes that shape soil, water, and ecological dynamics of our biosphere. KBase is a an open-source software and data platform funded by DOE to enable sharing, integration, and analysis of many types of data associated with microbes, plants, and their communities using scalable computing infrastructure. This extensive community resource facilitates large-scale analyses of biological systems and is designed to accelerate scientific discovery, improve reproducibility, and foster open collaboration.

KBase offers a suite of scientific applications to enable users to build sophisticated analytical workflows and share their findings. The nearly 200 apps in KBase offer diverse scientific functionality across the realms of comparative genomics, communities analysis, metabolic modeling, and transcriptomics. Additionally, a number of tools and services within KBase have been co-developed with developers at the DOE Joint Genome Institute. Users can build and share sophisticated workflows through a combination of chaining together multiple analysis tools, writing scripts for automation, and using batch processing, all within notebook-style *Narrative*s that contain the employed data and tools. Developers can build, test, register, and

deploy new or existing software as KBase apps using the Software Development Kit, thereby extending the platform's scientific capabilities.

Recently developed features allow for greater organization of collaborative projects and increased depth of discovery within massive datasets. Projects, laboratories, and even whole institutions can organize their users and associated Narratives into a shared *Organization* with multiple permission levels and management features. Additionally, an early version of social feeds informs users of changes happening within their Organizations. Newly added services enable the platform to find and suggest data sets or Narratives that may be of interest to a particular user, based on searching interconnections between the data in KBase. These services will ultimately evolve into knowledge-discovery features, enabling KBase to propose new hypotheses by making connections across the system. KBase is unique in offering these diverse and integrated capabilities to a growing user community that is actively pioneering the use of the platform in their publications.

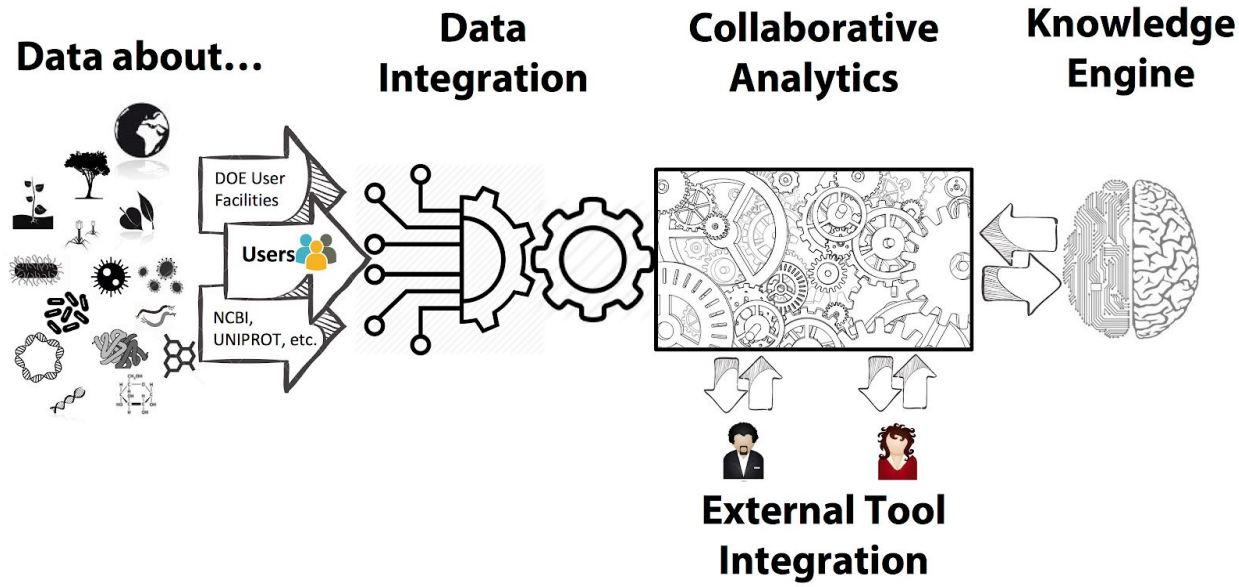


Figure 1. Integration of data and tools into KBase.

KBase is funded by the Genomic Science program within the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award numbers DE-AC02-05CH11231, DE-AC02-06CH11357, DE-AC05-00OR22725, and DE-AC02-98CH10886.

Microbiome Genome Extraction, Phylogenomics, and Metabolic Modeling of Species Interactions in KBase

Dylan Chivian*[1] (DCChivian@lbl.gov), José P. Faria*[2] (JPLFaria@anl.gov), Janaka N. Edirisinghe*[2] (JanakaE@anl.gov), Paramvir S. Dehal*[1] (PSDehal@lbl.gov), Richard S. Canon[1], Elisha Wood-Charlson[1], **Adam P. Arkin[1]**, **Bob Cottingham[3]**, **Chris Henry[2]**, and the KBase Team at the following institutions

[1]Lawrence Berkeley National Laboratory, Berkeley, CA; [2]Argonne National Laboratory, Argonne, IL; [3]Oak Ridge National Laboratory, Oak Ridge, TN; [4] Brookhaven National Laboratory, Upton, NY; [5] Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

http://kbase.us

Project Goals: The DOE Systems Biology Knowledgebase (KBase) is a free, open-source software and data platform that enables researchers to collaboratively generate, test, compare, and share hypotheses about biological functions; analyze their own data along with public and collaborator data; and combine experimental evidence and conclusions to model plant and microbial physiology and community dynamics. KBase's analytical capabilities currently include (meta)genome assembly, annotation, comparative genomics, phylogenomics, transcriptomics, and metabolic modeling. Its web-based user interface supports building, sharing, and publishing reproducible, annotated analysis workflows with integrated data. Additionally, KBase has a software development kit that enables the community to add functionality to the system.

KBase was designed to enable systems biology analysis of communities of microbes and/or plants. KBase is extensible and currently includes powerful tools for metabolic modeling, comparative and phylogenomics of microbial genomes that can be used for developing mechanistic understanding of functional interactions between species in microbial ecosystems. Essential to gaining new insight is obtaining high-quality genomes to annotate, either via cultivation or genome extraction from metagenome assembly. KBase has incorporated and added to a suite of microbiome analysis apps meant to be used in concert, including sequence QA/QC tools such as Trimmomatic and FastQC, taxonomic structure profiling of shotgun metagenome sequence with Kaiju, custom KBase apps for generating sample-specific *in silico* reads for downstream benchmarking, several metagenome assembly algorithms including MEGAHIT, IDBA-UD, and metaSPAdes, custom KBase apps for comparing metagenome assemblies, grouping assembled genome fragments (contigs) into putative genomes (bins) with MaxBin2 and other binners, and genome completeness and contamination assessment with CheckM. Comparison of results and quality assessment of the performance of tools and parameterization against data of various characteristics (e.g. low-complexity, high-complexity) by benchmarking at each stage of this process are offered. Additionally, we've recently released tools and services that allow users to search rapidly (seconds to minutes) all reference genome databases, metagenomes and published metagenome-assembled genomes (MAGs) using their reads, assemblies or MAGs. This is implemented using a MinHash like sketching process that works well for identifying matches above ~90% identity.

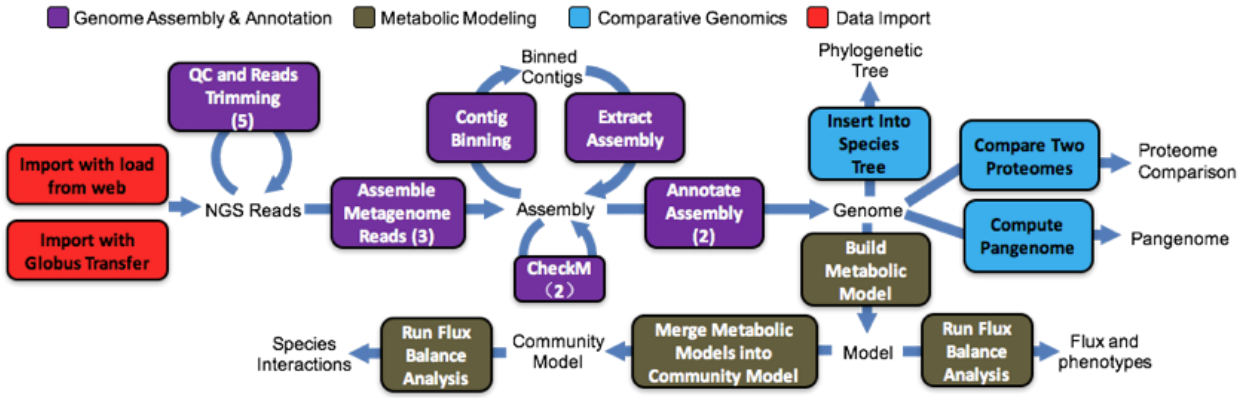


Figure 1. Example approach for use of KBase's microbiome genome extraction tool suite.

We have greatly expanded microbiome analysis in KBase. It is now possible to incorporate and use tools that enable users to get from shotgun reads through to MAGs to phylogenomics and metabolic modeling. As an example from our initial set of tools (Figure 1), a user can upload or find data from collaborators or the public and apply one of the metagenome assembly apps and bin the assembled contigs so that individual genomes can be extracted from the bins. Once individual MAGs are extracted, the highest quality fraction can be piped into a wide range of downstream analysis apps in KBase, including genome annotation, phylogenetic placement and genome content comparison with respect to one another, KBase reference genomes, and other public genome and MAG collections. Additionally, metabolic modeling and RNA-seq alignment can be performed. After generating metabolic models from the genomes assembled from a metagenome, individual metabolic models can be combined into a community metabolic model, which can be applied with the Flux Balance Analysis app to predict trophic interactions between species. Users have applied these tools to study: (i) interactions between plants and microbes in soil; (ii) why some microbes form stable communities; (iii) how a microbial community cooperates to produce a specific product; and (iv) how a community of heterotrophic species can feed on byproducts from an autotroph to grow autotrophically.

In addition to efforts by KBase developers to expand the functionality of our Microbiome tool suite, community developers have been adding tools that they use and have developed, including members of the DOE Joint Genome Institute (MetaBAT2, RQCFileter, JGI Metagenome Assembly Pipeline), the ENIGMA SFA, the LLNL Soils SFA (vConTACT2, VirSorter), and the LANL Bacterial:Fungal Interactions SFA (GOTTCHA2). All Apps in KBase are openly available for users to apply with their own data.

KBase is funded by the Genomic Science program within the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award numbers DE-AC02-05CH11231, DE-AC02-06CH11357, DE-AC05-00OR22725, and DE-AC02-98CH10886.

New pipeline for multi-omics data integration and discovery in KBase to identify the mechanism driving metabolic/OTU dynamics in a microbiome

Adam P. Arkin[1], **Bob Cottingham**[3], **Chris Henry**[2], and the KBase Team at the following institutions

[1]Lawrence Berkeley National Laboratory, Berkeley, CA; [2]Argonne National Laboratory, Argonne, IL; [3]Oak Ridge National Laboratory, Oak Ridge, TN; [4] Brookhaven National Laboratory, Upton, NY; [5] Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

http://kbase.us

Project Goals: The Department of Energy Systems Biology Knowledgebase (KBase) is a knowledge creation and discovery environment designed for both biologists and bioinformaticians. KBase integrates a large variety of data and analysis tools, from DOE and other public services, into an easy-to-use platform that leverages scalable computing infrastructure to perform sophisticated systems biology analyses. KBase is a freely available and developer extensible platform that enables scientists to analyze their own data within the context of public data and share their findings across the system.

Increasingly microbiome systems are being interrogated using a combination of DNA sequencing and metabolome mass spectrometry with the ultimate goal of understanding how microbial communities shape (and are shaped by) the chemistry of their surrounding environments. Yet, the challenge of deciphering the biological mechanisms that give rise to observed dynamics in metabolite and species abundances in a given environment based on this data remains . Here we demonstrate a new workflow in KBase, comprised of many new data types and tools recently added to the KBase platform, that permit users to:
(1) identify metabolites and species that correlate based on cross-comparison of samples;
(2) search for isolates of species of interest that interact with metabolites of interest;
(3) predict biosynthesis pathways for a metabolite of interest in an isolate of interest;
(4) identify gene candidates for gap-filled steps within a predicted pathway;
(5) check phylogenetic neighbor genomes for evidence of conservation of pathway of interest;
(6) find and query available transcriptomes for evidence of expression of pathway of interest;
(7) identify other environments that involve similar species, metabolites, and pathways.

We demonstrate this new pipeline in action by analyzing a dataset from the Ecosystems and Networks Integrated with Genes and Molecular Assemblies (ENIGMA) SFA. In this study, a soil core was retrieved from a contaminated aquifer at the Oak Ridge National Lab Field Research Center (ORNL FRC), and fourteen biological samples were collected from this core at 9 inch vertical intervals. Each sample underwent both amplicon sequencing and metabolomic analysis, identifying a total of 3940 OTUs and 34 metabolites. Within KBase, we were able to correlate the OTUs to the metabolites. As we were particularly interested in applying a multi-omics approach mechanistically to understand these correlations, we applied the *Object Counts by Taxon* tool, which leverage the KBase Relation Engine that connects together related data entities in KBase, to identify which taxons in our dataset had the most multi-omics data available elsewhere in KBase. In this analysis, *Pseudomonas* emerged as the best candidate. Of the metabolites correlated to *Pseudomonas*, betaine had highest positive correlation. A search in

KBase revealed the Web of Microbes dataset uploaded by the Northen lab as part of a JGI collaboration with KBase, which demonstrated that a particular isolate of *Pseudomonas* is known to produce betaine as a byproduct. We applied the *Predict metabolite biosynthesis pathway* tool in KBase to identify 17 reactions and 30 genes involved in betaine biosynthesis in this strain of *Pseudomonas*, with three of the reactions having unknown genes (including the final step in betaine biosynthesis). We applied the *Find Candidate Genes for a Reaction* tool, which is another tool that leverages the KBase relation engine, to identify candidate genes for each of the three gap-filled steps in the betaine pathway. We then used the *Homolog Genome Context* to study how conserved all the betaine pathway genes (including our new candidate genes) are across all close *Pseudomonas* genomes. Close *Pseudomonas* genomes were determined using the *Mash Search* tool in KBase, which was developed in collaboration with JGI. This anlaysis revealed that the betaine pathway is broadly conserved in *Pseudomonas* genomes. We used data discovery tools (driven by the relation engine) in KBase to identify numerous public transcriptome profiles for some of the close *Pseudomonas* genomes identified by MASH. A search of our betaine gene families in these datasets revealed that they are most often expressed in antibiotic and protozoa induced stress conditions. Finally, we applied the *MAG (Metagenomic Assembled Genomes) Mash* tool, which was also developed in collaboration with JGI, to identify metagenome samples that contain similar pseudomonas genomes, revealing other environments where *Pseudomonas* and betaine production are potentially significant.

Overall, using the data discovery and analysis tools in KBase (and less than one hour of user-time), we were able to develop a mechanistic explanation for a metabolite-OTU correlation, identify missing genes in the betaine biosynthesis pathway in pseudomonas, determine that the betaine pathway is highly conserved in pseudomonas, pinpoint some conditions under which the betaine pathway is expressed, and explore some other environments where similar genomes are known to exist. The combination of tools applied in this workflow can be used to perform similar studies on numerous other microbiome and isolate systems, enabling the integration of multi-omics data to translate correlation to mechanistic understanding.

KBase is funded by the Genomic Science program within the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award numbers DE-AC02-05CH11231, DE-AC02-06CH11357, DE-AC05-00OR22725, and DE-AC02-98CH10886.

Transcriptome Profiling and Genome-Scale Metabolic Modeling in Fungal and Plant genomes using KBase.

Sunita Kumari*[1] (kumari@cshl.edu), Vivek Kumar[1], Sam Seaver[2], Janaka Edirisinghe[2], Priya Ranjan[3], Zhenyuan Lu[1], Doreen Ware[1], Paramvir Dehal[4], Ben Allen[3], Jason Baumohl[4], Jay Bolton[4], Shane Canon[4], Stephen Chan[4], John-Marc Chandonia[4], Dylan Chivian[4], Mikayla Clark[3], Meghan Drake[3], Jose Faria[2], Annette Greiner[4], Tian Gu[2], James Jeffryes[2], Marcin Joachimiak[4], Roy Kamimura[4], Keith Keller[4], Miriam Land[3], Sebastian Le Bras[4], Akiyo Marukawa[4], Sean McCorkle[5], Cheyenne Nelson[4], Dan Murphy-Olson[2], Erik Pearson[4], Gavin Price[4], William Riehl[4], Boris Sadkhin[2], Alan Seleman[2], Gwyenth Terry[4], Elisha Wood-Charlson[4], Shinjae Yoo[5], Qizhi Zhang[2], Diane Zheng[4], **Chris Henry[2], Bob Cottingham[3], Adam P. Arkin[4]**

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; [2]Argonne National Laboratory, Argonne, IL; [3]Oak Ridge National Laboratory, Oak Ridge, TN; [4]Lawrence Berkeley National Laboratory, Berkeley, CA; [5]Brookhaven National Laboratory, Upton, NY;

http://kbase.us

Project Goals: The Department of Energy Systems Biology Knowledgebase (KBase) is a knowledge creation and discovery environment designed for both biologists and bioinformaticians. KBase integrates a large variety of data and analysis tools, from DOE and other public services, into an easy-to-use platform that leverages DOE computational infrastructure to perform sophisticated systems biology analyses. KBase is freely available and enables scientists to upload their own data, analyze it alongside collaborator and public data, and share workflows and conclusions. Additionally, KBase has a software development kit that enables the community to add functionality to the system.

KBase is an open, web-accessible computational environment for systems biology research focused on microbes, fungal, plants and their communities. It provides a range of integrated biological data types and associated analysis tools (Apps) that include gene expression analysis, metabolic modeling, comparative genomics and functional genomics. The user-friendly KBase Narrative Interface offers the researchers and bioinformaticians a range of analysis tools and data resources that accelerate the pace of functional genomics research by allowing large-scale sample processing, expression-level quantification and integration of gene expression profiles with downstream functional analysis including clustering of expression profiles based on different algorithms, ontology enrichment, metabolic networks and gene regulation.

KBase currently has 78 plant genomes from the JGI Phytozome database, and 134 fungal genomes from the JGI MycoCosm database. KBase has several data resources that originated from the PlantSEED project which combines plant comparative genomics, functional annotation

of enzymes, and reconstruction of plant primary metabolism for individual species. Plant-specific compounds and reactions, collected from public sources such as KEGG, MetaCyc, and AraCyc, have been integrated into PlantSEED and made available in KBase, where they can be used for plant metabolic modeling.

KBase also integrates reactions, biomass formulations, and gene annotations from 12 published genome-scale metabolic models for diverse fungal genomes. These published models are used by a new app that enables users to automatically construct a new genome-scale metabolic model for any fungal genome. These fungal models are an efficient way of predicting phenotypes across various environmental conditions.

KBase is also actively engaged with the external community to help us improve our tools and workflows for functional genomics especially support for gene expression, regulation, and epigenetics in plant science. These capabilities are directly relevant to important DOE research targets such as optimizing biomass production in biofuel feedstocks.

KBase is funded by the Genomic Science program within the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award numbers DE-AC02-05CH11231, DE-AC02-06CH11357, DE-AC05-00OR22725, and DE-AC02-98CH10886.

KBase User Engagement: Leveraging Outreach and In-Reach Strategies to Build a Collaborative Community

Elisha Wood-Charlson*[1] (elishawc@lbl.gov), Meghan Drake*[3] (drakemm@ornl.gov), Benjamin Allen[3], Mikayla Clark[3], **Adam P. Arkin[1]**, **Bob Cottingham[3]**, **Chris Henry[2]**, and the KBase Team at the following institutions:

[1]Lawrence Berkeley National Laboratory, Berkeley, CA; [2]Argonne National Laboratory, Argonne, IL; [3]Oak Ridge National Laboratory, Oak Ridge, TN; [4]Brookhaven National Laboratory, Upton, NY; [5]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

http://kbase.us

Project Goals: The Department of Energy Systems Biology Knowledgebase (KBase) is a knowledge creation and discovery environment designed for both biologists and bioinformaticians. KBase integrates a large variety of data and analysis tools, from DOE and other public services, into an easy-to-use platform that leverages scalable computing infrastructure to perform sophisticated systems biology analyses. KBase is a freely available and developer extensible platform that enables scientists to analyze their own data within the context of public data and share their findings across the system.

We are building a community at KBase, and we want you to be a part of it! There are many ways to join the KBase family. A KBase user account is free and open to everyone, with full access to public data, ability to upload your own data whether private or public, run data analyses using community-developed tools, share with collaborators, and finally release to the community when you are ready - all via a platform that allows for publication-level reproducibility. KBase also works with community developers to release their tools on the KBase platform through formalized Software Development Kit (SDK) training, one-on-one support via Slack, and regular workshops/webinars.

We are also building User Working Groups (UWGs) in three main research areas: Microbiome, Metabolism, and Functional Genomics. These UWGs represent community members that are ready to get their data, tools, and analysis workflows up and running on KBase. We are also working with several Science Focus Area projects, recently funded by DOE, to get their data and tools on system. Each of the five SFAs are represented in the KBase User Science Presentations session (Monday afternoon), as one of the priorities at KBase is to support and champion our users, developers, and collaborators. Finally, one of the newest KBase features allows users to create a private (non-discoverable) or public landing page for their research group, lab, or organization. This simplifies how you interact with your immediate community (e.g., sharing narratives and inviting new members) via a customizable interface.

More information on the collaborative capabilities inside the KBase platform can be found on the *KBase: The Systems Biology Knowledgebase for Predictive Biological and Environmental Research in an Integrated Data Platform* poster.

KBase also runs regular in-person workshops, has a presence at many key conferences, and continually hosts webinars on a variety of topics to demonstrate functionality on the platform in a venue that allows for community feedback and real-time questions/answers. To receive announcements of upcoming events, please sign up at http://kbase.us/newsletter. If you are interested in knowing more about joining the KBase community or would like to speak to a member of the KBase outreach team outside of the poster session, please email engage@kbase.us.

KBase is funded by the Genomic Science program within the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award numbers DE-AC02-05CH11231, DE-AC02-06CH11357, DE-AC05-00OR22725, and DE-AC02-98CH10886.

Biosensor and Optogenetics for Systems Biology of Yeast Branched-Chain Higher Alcohol Production and Tolerance

Evan Zhao,[1] Sarah Hammer,[1] Yanfei Zhang,[1] and **José L. Avalos** [1,2]* (javalos@princeton.edu)

[1]Department of Chemical and Biological Engineering; [2]The Andlinger Center for Energy and the Environment, Princeton University, Princeton, NJ.

Project Goals:

The goal of this project is to carry out a comprehensive systems biology study of branched-chain higher alcohol (BCHA) production and tolerance in yeast. We will leverage a genetically encoded biosensor of BCHA production to screen various yeast genomic libraries to measure the effects of different genetic perturbations (gene deletion, overexpression, or mutation) on BCHA production or tolerance. In addition, we will carry out transcriptomic studies leveraging again our BCHA biosensor, as well as optogenetic circuits to control BCHA production with light. This combination will allow us to establish closed-loop control systems that we can use to measure transcriptomic changes in well-controlled steady state or dynamic production settings. Ultimately, we will use these genomic and transcriptomic data to discover the key cellular networks involved in BCHA production and tolerance, which will be instrumental in developing better producing strains.

Branched-chain higher alcohols (BCHAs), including isobutanol, isopentanol, and 2-methyl-1-butanol, are some of the most promising advanced biofuels in development. These alcohols have better fuel properties than bioethanol, including higher energy density and better compatibility with current gasoline use and distribution infrastructure. Furthermore, BCHAs can be upgraded to jet fuel, making them excellent renewable fuels for ground as well as air transportation. Existing efforts to commercialize these types of biofuels are challenged by limited productivities, as well as the high toxicity that these alcohols have on strains engineered to produce them. Significant progress has been made in boosting yields and titers, particularly of isobutanol, through extensive metabolic and enzyme engineering based on detailed knowledge of branched chain amino acid metabolism and the structure and function of the enzymes involved. In contrast, virtually nothing is known about the interplay between different cellular networks and BCHA production and tolerance, leaving two basic questions unexplored: i) What are the key cellular networks that influence BCHA production? ii) What cellular networks are most affected by BCHA production, accounting for cellular sensitivity and tolerance to these alcohols?

In this presentation I will describe and demonstrate two new enabling technologies, which we will use to carry out the first systems biology studies on the production, sensitivity, and tolerance to BCHAs in yeast. The first is a genetically encoded biosensor that can monitor the activity of BCHA biosynthesis. We have used this biosensor not only to identify high BCHA producing strains, but also to develop selection screens, which we have used to engineer several enzymes in

the BCHA biosynthetic pathway to enhance their activity. The second enabling technology is a set of optogenetic circuits used to control the expression of different metabolic enzymes with light. Using this platform, we can dynamically control cell ethanol production and growth, as well as BCHA production through light pulses, allowing us to achieve record-breaking yields of these advanced biofuels [1]. Combining this genetically encoded BCHA biosensor with established yeast genomic collections, as well as with our optogenetic control platform to established closed-loop control systems, we will carry out first-in-class systems biology studies, including genomic and transcriptomic measurements to address the two fundamental questions above. These studies will provide new fundamental insights into the production and tolerance of BCHAs in yeast, which will open new avenues for strain development.

References

1. Zhao EM, Zhang Y, Mehl J, Park H, Toettcher JE, Avalos JL. Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature* 29; 555 (7698):683-87 (2018).

This research is supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0019363.

Optimizing Tradeoffs Implicit During Bioenergy Crop Improvement: Understanding the Effect of Altered Cell Wall and Sugar Content on Sorghum-Associated Pathogenic Bacteria

Q. Wang,[1] D. Braun,[2] N. Fahlgren,[1] M. Gehan,[1] W. Vermerris,[3] and **Rebecca Bart**[1]*(rbart@danforthcenter.org)

[1]Donald Danforth Plant Science Center, St. Louis, MO; [2]University of Missouri, Columbia, MO; [3]IUniversity of Florida, Gainesville

Project Goals: The first project goal is to quantitatively model the disease triangle that describes sorghum, pathogenic bacteria, and the environment. Field and laboratory experiments are being combined to determine bacterial susceptibility of genetically diverse sorghum genotypes that differ in cell wall and sugar composition. Standard plant pathology techniques combined with powerful phenomics approaches are providing a holistic view of this pathosystem within variable environments. Further, transcriptomics is being employed to elucidate mechanisms used by bacterial pathogens to induce sorghum susceptibility. Microbial pathogens are known to manipulate the sugar and cell wall characteristics of their hosts. Consequently, these characteristics will be analyzed during pathogen invasion. This research will reveal the mechanisms underlying tolerance to pathogens that must be maintained during biofuel trait optimization.

Plant-derived production of renewable fuels and chemicals has the potential to enhance US farming and agricultural economic opportunities, increase domestic energy security, and reduce fossil fuel dependency and greenhouse gas emissions. Realizing the potential of these alternative energy sources necessitates the development of high-biomass-yielding crops. These specialized crop varieties may harbor modifications to cell walls, which are a major barrier to pathogen entry, and to the tissue distribution of sugars, which are the pathogen's food source; hence they are likely to present previously unseen challenges for disease resistance. Over the last several years, disease from the bacterial pathogen *Xanthomonas*, has caused significant yield losses in many crops where bacterial diseases had historically been rare, including corn and cotton. It is currently unclear why these diseases are emerging. *Xanthomonas* is a known pathogen of sorghum (*Sorghum bicolor* (L.) Moench), though similar to corn and cotton, the incidence and impact of the disease has historically been low. Taken together, these observations highlight a vulnerability in sorghum's resilience to pathogens that is likely to be magnified by alterations in cell wall and sugar content. In this project, we aim to establish the sorghum – *Xanthomonas* pathosystem as a model for deducing how latent microbial pathogens might exploit key biofuel crop traits.

We will report results to date including efforts to classify over one hundred sorghum varieties for resistance/susceptibility to *Xanthomonas* and an accompanying GWAS analysis, dual-RNAseq data that has revealed gene expression patterns in both host and pathogen during compatible and incompatible interactions, and characterization of cell wall and sugar profiles during pathogen attack.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0018072.

Manipulating Carbon Flux between Mesophyll and Bundle Sheath Cells to Optimize Photosynthetic Performance

Jennifer Arp[1]* (jarp@danforthcenter.org), Doug K. Allen,[1,2] Todd Mockler,[1] and **Ivan Baxter**[1]

[1] Donald Danforth Plant Science Center, St. Louis, MO
[2]United States Department of Agriculture, Agricultural Research Service, St. Louis, MO

www.foxmillet.org

Project Goals: This project aims to leverage Setaria viridis as a model system to develop novel technologies and methodologies to redesign the bioenergy feedstock Sorghum bicolor to enhance water use and photosynthetic efficiencies.

C_4 photosynthesis evolved to be more efficient than C_3 photosynthesis in hot and dry environments by utilizing specialized anatomy and biochemistry. C_4 is a highly convergent phenotype, having evolved 22-24 times in the grasses and over 60 times in plants and co-opting a variety of enzymes[1]. Previously, C_4 plants have been classified into three subtypes based on the decarboxylation enzyme used, but evidence now suggests plants may use a mixture of pathways. An untested hypothesis is that the mixed pathway allows smaller pools of transfer metabolites and flexibility in changing environments[2]. Maize uses both the NADP-malic enzyme (NADP-ME) and phosphoenolpyruvate carboxykinase (PEPCK) pathways, with aspartate and malate both serving as transfer metabolites into bundle sheath cells[3]; however, there is no evidence for mixed pathway use in the related grasses setaria and sorghum. The goal of this project is to install a supplemental PEPCK pathway into the NADP-ME type C_4 plants *Setaria viridis* and *Sorghum bicolor* using genetic elements from maize. Constructs are in progress to generate transgenic plants with the maize PEPCK gene driven by the native maize promoter or separately by a bundle sheath-preferred promoter such as RbcS. In addition, guide RNAs targeted to NADP-ME pathway genes have been designed to generate weak alleles and decrease relative flux through that pathway. Last, constructs are underway to modulate asparagine cycling genes to improve availability of aspartate as a transfer molecule in the mesophyll. In parallel, phenotyping tools are being developed to confirm flux through a PEPCK pathway which will be used to characterize the transgenic plants as they become available. The main approach will use isotopic labeling from $^{13}CO_2$ to measure relative flux through aspartate and malate and additionally will enable quantification of flux through the Calvin cycle. Gas exchange measurements will also be performed to measure enhancement of photosynthesis. The transgenic plants will demonstrate increased flux through the photosynthetic carbon concentrating mechanism which can be used to improve the performance of sorghum and related grasses in marginal and variable environments.

References

1. Sage, R. F., Christin, P.-A. & Edwards, E. J. The C(4) plant lineages of planet Earth. *J. Exp. Bot.* **62**, 3155–3169 (2011).
2. Wang, Y., Bräutigam, A., Weber, A. P. M. & Zhu, X.-G. Three distinct biochemical subtypes of C4 photosynthesis? A modelling analysis. *J. Exp. Bot.* **65**, 3567–3578 (2014).
3. Wingler, A., Walker, R. P., Chen, Z. H. & Leegood, R. C. Phosphoenolpyruvate carboxykinase is involved in the decarboxylation of aspartate in the bundle sheath of maize. *Plant Physiol.* **120**, 539–546 (1999).

This project is funded by grant DE-SC0018277 from The DOE Department of Biological and Environmental Research.

Phenotyping for water use efficiency in C_4 grasses *Setaria* and *Sorghum*

Max Feldman[1], Patrick Elsworth, Darshi Banan, Rachel Paul, Hui Jiang[1], Jennifer Barrett[1], Nadia Shakoor[1], Erica Agnew[1], Todd Mockler[1], Asaph B. Cousins, Andrew Leakey and **Ivan Baxter[1]**

[1]Donald Danforth Plant Science Center, St. Louis, MO[1]
[2]Washington State University, Pullman, WA 99163
[3]University of Illinois at Urbana-Champaign, Urbana, IL 61810

Project Goals:

This project aims to leverage *Setaria viridis* as a model system to develop novel technologies and methodologies to redesign the bioenergy feedstock *Sorghum bicolor* to enhance water use and photosynthetic efficiencies.

url: www.foxmillet.org

Abstract: Plant growth and water use are interrelated processes influenced by the genetic control of both plant morphological and biochemical characteristics. Improving plant water use efficiency (WUE) to sustain growth in different environments is an important breeding objective that can improve crop yields and enhance agricultural sustainability. However, genetic improvements of WUE using traditional methods have proven difficult due to low throughput and environmental heterogeneity encountered in field settings. To overcome these limitations we utilizes a high-throughput phenotyping platform to quantify plant size and water use of a populations of the C4 species *Setaria viridis* and *Sorghum Bicolor* under water availability contrasts. Our findings indicate that measurements of plant size and water use in this system are strongly correlated; therefore, a linear modeling approach was used to partition this relationship into predicted values of plant size given water use and deviations from this relationship at the genotype level. The resulting traits describing plant size, water use and WUE were all heritable and responsive to soil water availability, allowing for a genetic dissection of the components of plant WUE under different watering treatments. Linkage mapping identified major loci underlying two different pleiotropic components of WUE. We are in the process of analyzing diversity population data to identify SNPs associated with these traits.

Funding statement: This work was supported by the Office of Biological and Environmental Research in the DOE Office of Science (DE-SC0008769).

Tissue and environment-specific synthetic gene circuits to regulate crown root growth

Jennifer A. N. Brophy[1]* (jbrophy@carnegiescience.edu), Katie J. Magallon[1], and **José R. Dinneny**[1,2]

[1]Carnegie Institution for Science, Stanford, CA; [2]Stanford University, Palo Alto, CA

www.foxmillet.org

Project Goals: This project aims to leverage *Setaria viridis* as a model system to develop novel technologies and methodologies to redesign the bioenergy feedstock Sorghum bicolor to enhance water use and photosynthetic efficiencies.

Grass species develop the majority of their root system through the adventitious initiation of crown roots from the base of the shoot. Crown root growth regulation is important for modulating whole-plant water use and preserving soil water. We are engineering synthetic regulatory circuits that control the development and physiology of crown roots to enhance water use efficiency (WUE) and water acquisition in C4 grass crops. Our approach is grounded in evidence that C4 grasses tend to exhibit overly conservative crown root responses to drought due to their wild origins. By engineering the crown roots of the model C4 grass *Setaria viridis* to respond differently to water, we aim to improve water use efficiency (WUE) and enable the design of plants that are optimized for specific environments.

Three crown root drought response strategies are being engineered and tested for their effects on growth and physiology in the lab and field. First, we are engineering *Setaria* to constitutively produce of crown roots, independent of changes in water availability. This "environment-independent strategy" is based on the observation that domesticated grasses tend to show reduced sensitivity to drought. In an agricultural context, such strategies may have been selected because crops are specifically planted in regions that historically have certain expected patterns of rainfall, whereas a wild species would need to be able to acclimate to a broad range of potential precipitation patterns depending on where it's seed landed. Second, we are engineering *Setaria* to enhance the production of crown roots when water is available. This "enhanced response to water" strategy is based on the idea that plants should maximize the number of crown roots produced during a precipitation event. Finally, we are engineering *Setaria* to enhance the suppression of crown roots during drought. This "enhanced suppression" strategy could be used to shift domesticated crops into more conservative strategies similar to their wild ancestors. This last strategy could be useful for broadening the normal geographic range in which a crop is grown, particularly into regions with sporadic or limited rainfall.

The construction of synthetic regulatory circuits to control crown root growth in these conditions requires fine-scale control of gene expression. We generated an RNAseq data set of crown responses to drought in order to identify promoters active in *Setaria* crowns. These promoters were classified as either constitutive, induced by water deficit, or induced by re-watering of drought stressed plants. Promoters that drive expression specifically in crown roots or surrounding tissues will be used to express genes that alter hormone signaling to control root growth. To generate new spatial patterns of gene expression or combine environmental input signals that control root growth in new ways, we are constructing synthetic genetic circuits using a new library of synthetic transcriptional activators, repressors, and promoters. Through transient

expression assays in *Nicotiana benthamiana* and stably transformed *Arabidopsis thaliana,* we show that each of these parts can be used to regulate transcription in plants and to construct synthetic genetic circuits that implement Boolean logic functions. These model systems enable rapid design-build-test cycles for synthetic genetic regulation.

In order to implement changes in crown root development, we will manipulate signaling pathways associated with root growth under well-watered and water deficit stress: abscisic acid (ABA), auxin, and reactive oxygen species (ROS). In *Arabidopsis,* we show that targeted expression of mutant auxin response regulators (AUX/IAA proteins) can alter lateral root production. A similar approach will be implemented in *Setaria* to inhibit auxin signaling and suppress crown root growth. We used RNAseq datasets to identify auxin response regulators (AUX/IAA-type proteins) that are highly expressed in *Setaria* crowns and likely to be involved in crown root development. Dominant negative mutations are being introduced to these proteins to manipulate auxin signaling and root development in *Setaria.* Genes identified by RNAseq as being induced by re-watering after drought are being explored as potential positive regulators of crown root growth. By expressing these developmental regulators at varying levels in specific conditions, we will attempt to control crown root responses to the environment to generate more robust C4 crop plants.

This project is funded by grant DE-SC0018277 from The DOE Department of Biological and Environmental Research.

Modification of PEPC kinetics to enhance the efficiency of C_4 photosynthesis

Asaph B. Cousins[1]* (acousins@wsu.edu), Robert J. DiMario[1]*(robert.dimario@wsu.edu), Ryan L. Wessendorf[1], Kuenzang Om[1], and **Ivan Baxter**[2]

[1]Washington State University, Pullman, WA 99163, [2]Donald Danforth Plant Science Center, St. Louis, MO 63132

url: http://foxmillet.org

Overall Project Goals: This project aims to leverage *Setaria viridis* as a model system to develop novel technologies and methodologies to redesign the bioenergy feedstock *Sorghum bicolor* to enhance water use and photosynthetic efficiencies. Here we specifically focus on Objective #1*: Engineering photosynthesis to improve performance under water stress.*

Abstract: Due to the predicted increase in food demand, studying the biochemical components of C_4 photosynthesis may provide insight into enhancing photosynthesis in crop plants to increase yield. Currently, photosynthesis can be reduced in C_4 crops by drought conditions which reduce intercellular CO_2 concentrations (C_i) in the plant. The initial carboxylation reaction in C_4 plants is catalyzed by phospho*enol*pyruvate carboxylase (PEPC) and leads to elevated CO_2 around Rubisco. The C_4 isozyme of PEPC originated from a non-photosynthetic PEPC and it has been suggested that specific amino acid substitutions in PEPC confer differences in the affinity of the enzyme for PEP (K_{PEP}). These changes in K_{PEP} may be an unavoidable side effect of selecting for a higher affinity for HCO_3^- (K_{HCO3}) to maintain rates of PEPC when stomatal conductance (g_s) is low. However, experimental evidence for amino acid changes influencing *in planta* kinetic properties of PEPC and rates of C_4 photosynthesis is lacking. Therefore, the objective of this aim is to determine how specific amino acid differences between the C_3 and C_4 isozymes of PEPC influence the efficiency of C_4 photosynthesis when the availability of atmospheric CO_2 is low. To accomplish this objective, we are measuring the kinetic properties of 28 PEPC isozymes from both C_3 and C_4 plants from members of the Poaceae family. These enzymes are being overexpressed and purified from the PEPC-less *PCR1 Escherichia coli* strain. The kinetic measurements will be compared to protein alignments to find specific amino acid residues contributing to the variation in PEPC kinetic properties. To test how these specified amino acids influence PEPC kinetics and C_4 photosynthetic efficiency we will use targeted mutagenesis with base editors and targeted gene replacement to modify specific amino acid residues. PEPC kinetics will be conducted in a temperature-controlled cuvette linked to a mass spectrometer, as previously described. The outcome from this research will determine if changes in specific amino acids confer kinetic differences of PEPC affinity for HCO_3^- and enhance C_4 photosynthesis. Ultimately, the goal is to introduce an enhanced PEPC enzyme into sorghum to increase photosynthesis under drought conditions. The outcome of this research will enhance C_4 photosynthetic efficiency and will lead to an increase in whole plant water use efficiency.

Funding statement: This work was supported by the Office of Biological and Environmental Research in the DOE Office of Science (DE-SC0018277).

Metabolome and transcriptome responses to a water deficit time course in the model C4 grass *Setaria Viridis*

Allen Hubbard[1], Xiaoping Li[1] , Hui Jiang[1], Jennifer Barrett[1], Amber Wolf[1], Ivan Baxter[1]
[1]Donald Danforth Plant Science Center, St. Louis, MO

www.foxmillet.org

Project Goals:
This project aims to leverage *Setaria viridis* as a model system to develop novel technologies and methodologies to redesign the bioenergy feedstock *Sorghum bicolor* to enhance water use and photosynthetic efficiencies.

As a first step in understanding the conserved gene networks underlying the metabolic response to drought in biofuel relevant C4 grass species, we used untargeted metabolomics and transcriptomics to profile *Setaria viridis* at multiple time points during drought stress. We quantified the levels of 7,486 metabolites using hydrophilic interaction chromatography (HILIC) and another 6,157metabolites using reversed phase liquid chromatography (RPLC), along with 35,065 transcripts using 3'end labeled RNA seq. This will allow us to identify key pathways impacted by drought, and describes how their behavior shifts across time. We selected drought responsive metabolites and used principal component analysis (PCA) as well as partial least squares discriminant analysis (PLS-DA) to detect a peak in the drought response at day 6 across the metabolome and transcriptome. We are in the process of conducting the same analysis in Sorghum bicolor. Using *Setaria viridis* and *Sorghum bicolor* diversity panels we will conduct a dual species metabolomic GWAS of drought stress to explore the role of conserved genes regions in regulating the response to water deficit in both species.

Funding statement.
This project is funded by grant DE-SC0018277 from The DOE Department of Biological and Environmental Research

Development of Transgenic Sorghum Lines to Enhance Water Use and Photosynthetic Efficiencies

Albert P. Kausch[1,*](apkausch@uri.edu), Kimberly Nelson Vasilchik[1], Joel Hague[1], Asaph Cousins[2], Dan Voytas[3], Matthew Zinselmeier[3], Colby Starker[3] and **Ivan Baxter**[4].

[1]University of Rhode Island, RI
[2]Washington State University, Pullman WA
[3] University of Minnesota, Saint Paul, MN
[4] Donald Danforth Plant Science Center, Saint Louis, MO
*Presenting Author

www.foxmillet.org

Project Goals: This project aims to leverage *Setaria viridis* as a model system to develop novel technologies and methodologies to redesign the bioenergy feedstock Sorghum bicolor to enhance water use and photosynthetic efficiencies.

Our objectives are towards meeting the needs for stable sorghum transgenics for this project and we are extending these capabilities into new transformation technologies for sorghum. We have established a reliable pipeline for the introduction of stable genetic constructs into sorghum for the projects outlined in the proposal. Our first goal has been accomplished by developing reliable protocols using standard *Agrobacterium*-mediated transformation of BTx430 for routine analyses and the transformation the sorghum reference genome, sorghum (BTx623). The Kausch Lab has generated a pipeline for this project and has many transgenic lines and events in various stages of regeneration for our initial constructs. This pipeline in collaboration with the Voytas and Cousins labs includes the following constructs driven by the maize Ubiquitin promoter and using the BAR gene as a selectable marker: (1) pNG026, BE3 nuclease; (2) pNG028, 1bCPF1 nuclease; (3) pMZ64, Ta Cas 9; (4) pMZ69, dTa Cas9; and , (5) pMZ263, dCas9 transcriptional activator under control of the Zm PEPC promoter targeting all four isoforms of carbonic anhydrase in the sorghum genome. One of our goals is to engineer photosynthesis to improve performance under water stress via (1) a transgenic approach using stably integrated constructs; (2) a transcriptional activation route employing dCas9; and (3) modification of PEPC kinetics to enhance the efficiency of C4 photosynthesis. We are currently in the process of generating these transgenic lines. Numerous events for each are currently in T_0 plants in regeneration for each. Another goal of our project is to develop methods to edit genes and to manipulate gene expression *in vivo* in Sorghum. To accomplish this goal we will use: (1) Targeted mutagenesis with nucleases (2) Targeted mutagenesis with base editors (3)Targeted gene replacement; and, (4) Targeted gene insertion. We are now exploring the use of the transient expression of morphogenic regulators and gene editing functions for rapid target validation and regeneration of edited events. There are currently four publications published on our results, one review paper, and three presentations from our work. Co-PI Kausch and former Co-PI Quemada were active in the discussion and publication regarding the Regulatory and Stewardship implications for using genome editing in sorghum and other crops. Both presented invited presentations at the SIVB meeting on this topic in June 2018.

Publications

1. Kausch A.P., Nelson-Vasilchik K., Hague J., Mookkan M., Quemada H., Dellaporta S., Fragoso C., Zhang Z.J. (2019) Edit at Will: Genotype Independent Plant Transformation in the Era of Advanced Genomics and Genome Editing, Plant Science, in press, available online January 14 2019, https://doi.org/10.1016/j.plantsci.2019.01.006

2. Do, P. T., Lee, H., Nelson-Vasilchik, K., Kausch, A., & Zhang, Z. J. (2018). Rapid and efficient genetic transformation of sorghum via *Agrobacterium*-mediated method. Current Protocols in Plant Biology, 3, e20077. doi: 10.1002/cppb.20077

3. Mookkan, M., Nelson-Vasilchik, K., Hague, J., Zhang, Z.J., and Kausch, A.P. (2018) Morphogenic regulator-mediated transformation of maize inbred B73. Current Protocols in Plant Biology 3 (4), e20075. doi: 10.1002/cppb.20075

4. Nelson-Vasilchik, K., Hague, J., Mookkan, M., Zhang, Z.J., and Kausch, A.P. (2018) Transformation of recalcitrant sorghum varieties facilitated by baby boom and wuschel2. Current Protocols in Plant Biology 3(4), e200076. doi: 10.1002/cppb.20076

5. Mookkan, M., Nelson-Vasilchik, K., Hague, J., Zhang, Z.J., and Kausch, A.P. (2017) Selectable marker independent transformation of recalcitrant maize inbred B73 and sorghum P898012 mediated by morphogenic regulators *BABY BOOM* and *WUSCHEL2*. Plant Cell Reports 36 (9), doi:10.1007/s00299-017-2169-1

Presentations

1. Albert Kausch* Invited Speaker* (2018) The Era of New Agricultural Product Development Via Genomics, Plant Transformation, and Genome Editing. Department of Cell and Molecular Biology, University of Rhode Island, Kingston, RI 02892 Society for In Vitro Biology Meetings, June 2-5 2018 St. Louis MO.

2. Albert Kausch* Invited Speaker* (2018). Public Policy, Societal Impacts, and Considerations for Advanced Biology, Genomics and Gene Editing in Pharmaceutical, Medical and Agricultural Applications. Department of Cell and Molecular Biology, University of Rhode Island, Kingston, RI 02892 Society for In Vitro Biology Meetings, June 2-5 2018 St. Louis MO.

3. Hector Quemada* and Albert Kausch Invited Speaker* (2018) Regulation of Genome Edited Events. Donald Danforth Plant Science Center, Saint Louis, MO 63132 Department of Cell and Molecular Biology, University of Rhode Island, Kingston, RI 02892 Society for In Vitro Biology Meetings, June 2-5 2018 St. Louis MO

This project is funded by grant DE-SC0018277 from The DOE Department of Biological and Environmental Research

Phenomics of stomata and water use efficiency in C_4 species

Andrew D.B. Leakey[1]* (leakey@illinois.edu), John Ferguson[1], Nathan Miller[2], Jiayang Xie[1], Charles Pignon[1], Gorka Erice[1], Timothy Wertin[1], Nicole Choquette[1], Maximilian Feldman[3], Funda Ogut[4], Parthiban Prakash[1], Peter Schmuker[1], Anna Dmitrievna[1], Dylan Allen[1], Elizabeth A. Ainsworth[5,1], **Ivan Baxter[3]**, Thomas P. Brutnell[3], Edward S. Buckler[6,7], Patrick J. Brown[1], Thomas E. Clemente[8], Stephen P. Long[1], Lauren M. McIntyre[4], Edgar Spalding[2]

[1]University of Illinois at Urbana-Champaign, Urbana, IL; [2]University of Wisconsin, Madison, WI; [3]Donald Danforth Plant Science Center, St Louis, MO; [4]University of Florida, Gainsville, FL; [5] USDA Agricultural Research Service, Urbana, IL; [6]USDA Agricultural Research Service, Ithica, NY; [7]Cornell University, Ithica, NY; [8]University of Nebraska, Lincoln, NE

http://foxmillet.org/

Genetically tractable model systems closely related to bioenergy grasses need to be developed to drive the crop improvement required for large scale, ecologically sustainable bioenergy production. *Setaria viridis* is an ideal candidate C_4 panacoid grass. The over-arching objectives of this large, collaborative project are to utilize genomic, computational and engineering tools to begin the genetic dissection of drought response in *S. viridis*. This will be achieved through: 1) Quantitative trait and association genetics; 2) novel controlled environment and field phenotyping combined with molecular and chemical profiling; 3) development of metabolic and gene networks; 4) development of transformation technologies; 5) reverse genetic testing of candidate genes.

Water use efficiency (WUE), which is physiologically distinct from drought tolerance, is a key target for improving crop productivity, resilience and sustainability. This is because water availability is the primary limitation to crop yield globally and irrigation uses the largest fraction of our limited freshwater supply. The exchange of water and CO2 between a leaf and the atmosphere is regulated by the aperture and pattern of stomata. Mechanistic modeling indicates that stomatal conductance could be reduced or stomatal movements accelerated to improve water use efficiency in important C4 crops such sorghum and sugarcane. While molecular genetics has revealed much about the genes regulating stomatal patterning and kinetics in Arabidopsis, knowledge of the genetic and physiological control of WUE by stomatal traits in C4 crops is still poor. Understanding of natural diversity in stomatal traits is limited by the lack of high-throughput phenotyping methods. Two novel phenotyping platforms were developed. First, a rapid method to assess stomatal patterning in three model C4 species grown in the field – maize, sorghum and setaria has been implemented. The leaf surface is scanned in less than two minutes with an optical tomographer, generating a quantitative measurement of a patch of the leaf surface. An algorithm was designed to automatically detect stomata in 10,000s of these images via training of a neural network approach. Second, a thermal imaging strategy, to rapidly screen the kinetics of stomatal closure in response to light has been developed. We identified genotype to phenotype associations for stomatal patterning, leaf gas exchange and canopy water use through quantitative trait loci and genome wide association studies. Transgenically modified expression of stomatal patterning genes has produced sorghum with greater WUE. These plants

were grown in a new field facility for comprehensive evaluation of leaf, root and canopy WUE traits under Midwest growing conditions in summer of 2018.

This research was funded through Subaward No. 23009-UI, CFDA # 81.049 between University of Illinois and Donald Danforth Plant Science Center Under Prime Agreement No. DE-SC0008769 from Department of Energy.

Title: Leveraging *Setaria viridis* and *Sorghum bicolor* pan-genomic variation for gene mining and crop improvement

Scott J. Lee[1]* (slee@danforthcenter.org), Adam Healey[2], Zhenbin Hu[3], Philip Ozersky[1], Jeremy Schmutz[2], Geoffrey P. Morris[3], **Ivan Baxter**[1], and Todd C. Mockler[1]

[1]Donald Danforth Plant Science Center, 975 North Warson Rd., Saint Louis, MO, USA,

[2]HudsonAlpha Institute for Biotechnology, 601 Genome Way, Huntsville, AL 35806

[3]Kansas State University, 1712 Claflin Rd., Manhattan, KS 66506

http://www.foxmillet.org

Project Goals: This project aims to leverage *Setaria viridis* as a model system to develop novel technologies and methodologies to redesign the bioenergy feedstock Sorghum bicolor to enhance water use and photosynthetic efficiencies.

The emergence of *Sorghum bicolor* as premier bioenergy crop is based on its naturally high yield, growth habit, and abiotic stress resistance. A key untapped avenue for sorghum breeders is the characterization of novel and wild alleles. Low coverage sequencing is inadequate to capture all the possible variation within diverse germplasm, especially for genes known to exhibit presence/absence variation. Model grass systems such as *Setaria viridis* can provide additional insights and diversity towards the discovery of novel and useful alleles for the development of bioenergy sorghum. The *de novo* assembly of a diverse collection of *Setaria viridis* and *Sorghum bicolor* lines will provide a novel resource in order to accelerate the development and breeding of sorghum as a bioenergy crop.

We performed high coverage Illumina-based sequencing of a diverse collection of 363 *S. bicolor* lines from the Sorghum Bioenergy Association Panel (BAP), along with de novo assembly of a subset of lines in order to begin to characterize the sorghum pan-genome. These BAP lines are a phenotypically, geographically, and racially diverse collection of sorghum, and were grown in both field and controlled greenhouse conditions. Following phenotyping and tissue collection, the lines were sequenced using the Illumina platform and were then aligned to the reference sorghum genome as well as de novo assembled to capture pangenomic variation. Additional genetic resource development for *Setaria viridis* will aid in these approaches. We performed genome-wide association studies (GWAS) of the phenotypes collected in sorghum, as well as analysis of observed polymorphisms, population structure, genomic selection, and presence/absence variation. We identified several known genes that show signs of selection as well as exhibit PAV within distinct subpopulations of the BAP panel. Additionally, we have identified several potentially novel genes that may be utilized to enhance biomass accumulation and water-use efficiency.

Funding statement: This project is funded by grant DE-SC0018277 from The DOE Department of Biological and Environmental Research

Targeted Mutagenesis and Programmed Transcriptional Regulation in Setaria and Sorghum

Yang Liu[1*] (Liu0177@umn.edu), Arjun Khakhar,[1] Nat Graham,[1] Matt Zinselmeier, [1] Colby Starker, [1] Albert Kausch,[2] Dan Voytas[1] and **Ivan Baxter**[3]

[1]University of Minnesota, St. Paul; [2]University of Rhode Island, West Kingston; [3]The Donald Danforth Plant Science Center, St. Louis, Missouri

Project Goals: This project aims to leverage *Setaria viridis* as a model system to develop novel technologies and methodologies to redesign the bioenergy feedstock *Sorghum bicolor* to enhance water use and photosynthetic efficiencies.

www.foxmillet.org

Improving *Sorghum bicolor* as a biofuel crop requires methods to edit genes and manipulate gene expression *in vivo*. We are optimizing mutagenesis strategies using CRISPR/Cas and CRISPR/Cpf1 nucleases to achieve targeted gene knockouts, gene replacements and transgene insertions. Further, we are implementing base editor technology to achieve precise sequence changes without the need for a DNA double strand break. In addition, we are optimizing the use of programmable transcription factors (activators and repressors) derived from nuclease inactive dCas9 and dCpf1. The programmable transcription factors will be deployed in an innovative strategy for biocontainment of transgenes. To achieve genetic containment, we will identify genes (target genes) that compromise viability of *Sorghum bicolor* when overexpressed by the programmable transcription factors. We plan to introduce mutations into the target gene so that it is no longer recognized by the transcription factors. We will then combine all components of the synthetic circuit needed for genetic containment and test efficacy.

This project is funded by grant DE-SC0018277 from The DOE Department of Biological and Environmental Research

Improving water-use efficiency under fluctuating light

Charles P. Pignon[1] * (pignon2@illinois.edu), Samuel Bonfim-Fernandes[1], Ravi Valluru[2], Nonoy Bandillo[2], Michael Gore[2], Edward Buckler[2], Ivan Baxter[3], Stephen P. Long[1], and **Andrew D. B. Leakey[1].**

[1]University of Illinois at Urbana-Champaign, Urbana, IL;

[2]Cornell University, Ithaca, NY;

[3]Danforth Center, St. Louis, MO

www.foxmillet.org

Project Goals:

This project aims to leverage Setaria viridis as a model system to develop novel technologies and methodologies to redesign the bioenergy feedstock Sorghum bicolor to enhance water use and photosynthetic efficiencies.

Abstract

In leaves, efficient water-use requires coordination of photosynthetic CO_2 assimilation with the activity of stomata, specialized cellular complexes which regulate CO_2 and water fluxes between the leaf and atmosphere. Stomata are slower than photosynthesis to respond to changes in light, and poor coordination between stomatal and photosynthetic processes can diminish productivity and water-use efficiency. Therefore, acceleration of stomatal responses has been hypothesized to improve water-use efficiency in fluctuating light conditions, which predominate in field crop canopies.

We used high-throughput thermal imaging to assess the speed of stomatal closure following a drop in light in a diversity panel of 667 accessions of *Sorghum bicolor* L. Different accessions showed various speeds of initial stomatal closure after the light was dimmed, and varying responses of stomatal re-opening at low light. The initial speed of stomatal closure, and the level of stomatal conductance at low light, were moderately heritable (h_g>0.45). Mapping with SNPs and gene expression data associated stomata kinetic traits to several candidate genes.

Funding statement.

This project is funded by grant DE-SC0018277 from The DOE Department of Biological and Environmental Research

QTG-Finder: A Machine-Learning Algorithm to Prioritize Causal Genes of Quantitative Trait Loci in Plants

Fan Lin, Jue Fan, and **Seung Y. Rhee***

Department of Plant Biology, Carnegie Institution for Science, Stanford, California 94305, USA

Project Goals: This project (www.foxmillet.org) aims to leverage *Setaria viridis* as a model system to develop novel technologies and methodologies to redesign the bioenergy feedstock *Sorghum bicolor* to enhance water use and photosynthetic efficiencies. In this study we developed a computational pipeline is to accelerate the discovery of causal genes in QTLs by reducing the number of candidates to be tested experimentally.

Linkage mapping is one of the most commonly used methods to identify genetic loci that determine a trait. However, the loci identified by linkage mapping may contain hundreds to thousands of candidate genes and require a time-consuming and labor-intensive fine mapping process to find the causal gene controlling the trait. With the availability of a rich assortment of genomic and functional genomic data, it is possible to develop a computational method to facilitate faster identification of causal genes. We developed QTG-Finder, a machine-learning algorithm to prioritize causal genes by ranking genes within a quantitative trait locus (QTL). Two predictive models were trained separately based on known causal genes in Arabidopsis and rice. With an independent validation analysis, we demonstrate the models can correctly prioritize about 80% and 55% of Arabidopsis and rice causal genes when the top 20% ranked genes were considered. The models can prioritize different types of traits though at different efficiency. We also identified several important features of causal genes including non-synonymous SNPs at conserved protein sequences, paralog copy number, and being a transporter. This work lays the foundation for systematically understanding characteristics of causal genes and establishes a pipeline to predict causal genes based on public data. Currently, we are expanding this algorithm to other species such as Setaria and Sorghum by using the othologs of known causal genes as a training set.

Funding statement:

This project is funded by grant DE-SC0018277 from The DOE Department of Biological and Environmental Research.

Computer-aided Engineering of Biomass Production under Drought in Sorghum and Setaria

Cheng Zhao[1], Pascal Schläpfer[1], George D. Cody[2], Jennifer Barrett[3], Ivan Baxter[3], **Seung Y. Rhee**[*,1]

1 Carnegie Institution for Science, Department of Plant Biology, Stanford, California 94305

2 Carnegie Institution for Science, Geophysical Laboratory, Washington, DC 20015

3 Donald Danforth Plant Science Center, St. Louis, MO 63132

Project Goals: This project (www.foxmillet.org) aims to leverage *Setaria viridis* as a model system to develop novel technologies and methodologies to redesign the bioenergy feedstock *Sorghum bicolor* to enhance water use and photosynthetic efficiencies.

In this study, we performed comparative metabolic network analysis under well-watered and water deficient conditions, using Flux Balance Analysis (FBA) to investigate how plants allocate metabolic resources for biomass production in response to drought. First, we collected biomass composition data for *Setaria* and sorghum shoot and root under well-watered and water-limiting conditions at multiple time points, using 13C solid state NMR. Second, we generated genome-scale metabolic models of *Setaria italica*, *Setaria viridis, and Sorghum bicolor* by an automated computational pipeline, converting pathway genome databases to metabolic network reconstructions. The biomass composition will be used to describe the objective function in the *S. bicolor* model. We will constrain the models by the transcriptome data generated from the consortium to further adjust the upper and lower bounds of each reaction. We will perform Flux Variability Analysis for determining the maximum range of flux that every reaction can possibly take on while the network is optimized for biomass production to identify the key reactions that limit biomass production under drought conditions.

This project is funded by grant DE-SC0018277 from The DOE Department of Biological and Environmental Research.

Investigation of Carbonic Anhydrase to Improve Sorghum C_4 Photosynthetic Efficiency

Joshua Rosnow[1]*(jrosnow@danforthcenter.org), Indrajit Kumar[1], Matthew Zinselmeier[2], Dan Voytas[2], Asaph Cousins[3], Todd Mockler[1], **Ivan Baxter**[1]

[1]Donalad Danforth Plant Science Center, St. Louis, MO

[2] College of Biological Sciences, University of Minnesota, St. Paul, MN

[3] The School of Biological Sciences, Washington State University, Pullman, WA

www.foxmillet.org

Project Goal: Leverage *Setaria viridis* as a model system to develop novel technologies and methodologies to redesign the bioenergy feedstock *Sorghum bicolor* to enhance water use and photosynthetic efficiencies.

Abstract:

One objective of this project is to optimize the expression of carbonic anhydrase (CA) to increase the residence time of carbon dioxide in the leaf to improve C_4 photosynthesis and drought resilience. Phylogenetic and gene expression analysis of CA across the grasses demonstrates that there are several α, β, γ - CA genes present in all grass genomes that exhibit a conserved expression in leaf tissue regardless of photosynthetic type. We are working to understand the role of alternative splicing of the most abundant leaf expressed β-CA, with the working hypothesis that two transcript isoforms facilitate CA being localized to different sub-cellular locations. On-going experiments include investigation of CA protein targeting and biochemical analysis of the alternatively spliced protein products. Results suggest that translation of the most abundant C_4 CA transcript does not begin in exon 1 as previously annotated. These results provide insights for designing constructs to optimize CA expression in Sorghum. Constructs that are currently being made for Sorghum include a β-CA overexpression construct and a transcriptional activator, dCas9-TV, for all β-CA isoform promoters with 1-2 sgRNAs per isoform. Further research on CA promoter elements that drive elevated and cell-specific gene expression, will help refine strategies for editing CA promoters to improve photosynthesis in Sorghum.

Funding Statement: This project is funded by grant DE-SC0018277 from The DOE Department of Biological and Environmental Research

Creating Synthetic Lichen Platforms for Sustainable Biosynthesis of Biofuel and Biochemical Precursors

Liqun Jiang,[1] Jackson Jenkins[1]* (jjenki71@jhu.edu), Cristal Zuniga,[2] Karsten Zengler,[2] Pavlo Bohutskyi,[3] Jon Magnuson,[3] Foteini Davrazou,[4] Michael Guarnieri,[4] Jamey Young,[5] and **Michael Betenbaugh**[1]

[1]Johns Hopkins University, Baltimore, MD; [2]University of California, San Diego; [3]Pacific Northwest National Laboratory (PNNL), Richland, WA; [4]National Renewable Energy Laboratory (NREL), Golden, CO; [5]Vanderbilt University, Nashville, TN

The goal of this project is to develop synthetic lichen communities of autotrophic and heterotrophic microbes as a novel sustainable symbiotic platform for the production of biofuel and its precursors. Carbon-fixing autotrophs provide oxygen and organic substrates to their heterotrophic neighbors, which in turn produce carbon dioxide. By optimizing and enhancing these interactions, we can create a robust, sustainable synthetic lichen community. Multi-omics driven genetic engineering will improve metabolite exchange and product generation capabilities with the microbial co-culture.

Lichens are communities of microbes that collect sunlight and carbon dioxide and apply it to power the group's activities, allowing the autotrophic member to optimize photosynthesis and metabolite generation while their heterotrophic fungal partners produce biochemical compounds for the community. Lichens can thrive in the harshest environments on earth, and they represent a robust model for a novel biotechnology platform that can transform CO_2 and sunlight into valuable energy-related biochemicals, eliminating the need for costly substrate feeding.

To create this novel platform, we are investigating multiple co-culture communities to determine optimal synthetic lichen partnerships. Several autotrophs, including engineered sucrose-secreting *S. elongatus+cscB* and other known extracellular polysaccharide (EPS)-secreting *Nostoc 7413* and *Nostoc 6720* were examined for their ability to secrete organic carbon into the culture media. Heterotrophic partners are under evaluation, including filamentous fungi *A. nidulans* and *A. niger*, yeast species *S. cerevisiae*, *R. glutinis*, and *C. curvatus* with the goal of generating energy-related precursors of biofuels or biochemicals of commercial value. We first examined the growth of these microbes in axenic culture and evaluated the sucrose and EPS production of the autotrophic partners, then used the spent media to successfully grow heterotrophs on the carbon produced. We also evaluated combinations of these organisms in co-culture, which showed symbiotic benefits to both partners. Results also show a shift in the cellular fatty acid profile, indicating potential interactions between the partners. Ongoing work with co-cultures will combine insights from community metabolic models, multi-omics profiling, and metabolic flux analysis to promote a fundamental understanding of synthetic lichen communities, to clarify interactions between co-culture partners, and to provide engineering targets for improved metabolic interaction and biofuels and bioproducts generation. The resultant findings will aid in the elucidation of fundamental mechanisms governing microbial symbioses and will in turn generate fundamental design rules and dynamic network properties of microbial systems for rational biosystems engineering.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under Award Number DE-SC0019388.

Synthetic lichens as a sustainable platform for biosynthesis

Cristal Zuniga[1*] (crzuniga@eng.ucsd.edu), Liqun Jiang,[2] Jackson Jenkins,[2] **Michael Betenbaugh,[2]** and **Karsten Zengler[1]**

[1]University of California, San Diego; [2]Johns Hopkins University, Baltimore

The goal of this project is to combine autotrophs and heterotrophs as a novel sustainable symbiotic platform for the production of biofuel and its precursors. The photosynthetic microorganisms in these synthetic lichens are providing substrates and oxygen to the heterotrophs, which in exchange produce CO_2 for carbon fixation via photosynthesis. Synthetic microbial communities of cyanobacterium-bacterium or cyanobacterium-fungus pairs were evaluated through genome-scale metabolic modeling. For this, we manually curated and updated transport capabilities in the model of the cyanobacteria *Synechococcus elongatus* PCC 6803 (*i*JB792) and *Yarrowia lipolytica* (*i*YO844). Three community metabolic models were created by pairing the cyanobacteria model with the models of *Yarrowia lipolytica*, and the two *Escherichia coli* strains W and K-12.

Symbiotic partnership between heterotrophs and phototrophs has promising biotechnology applications. The experimental success of working with these communities relies on the correct selection of the photobiont and the heterotroph partner as well as on our ability to predict the partners' fitness and responses to environmental changes and genetic stability. Here, community metabolic models were reconstructed in order to guide strain selection for a fruitful lichen-like partnership. We paired the photobiont *Synechococcus elongatus* with the heterotrophs *Yarrowia lipolytica, Escherichia coli* W *and* K-12, respectively. Three community metabolic models (CM-models) were reconstructed after updating individual genome-scale metabolic models. CM-models were integrated using the shared metabolite pool (SMP) approach, which includes metabolites that can potentially be exchanged by each synthetic community. Potential exchanges were defined using BIOLOG data. The Constraint-Based Reconstruction and Analysis of Communities (COBRA-C) toolbox has been expanded in the framework of this project. COBRA-C allows obtaining single model statistics and provides tools for the reconstruction of community models. Additionally, COBRA-C contains test functions for several co-culture characterization applications, such as a) prediction and fitting of growth rates and population proportions (constraints-based choice to achieve experimental growth rates); b) determination of possible interactions (theoretical interchange of metabolites, SMP analysis); c) co-culture medium optimization (robustness analysis); d) syntrophic pathway inclusion (metadata contextualization); and e) gene essentiality (knock-out analysis for population)

The reconstructed CM-models helped elucidating syntrophic interactions between community members, while sustaining the heterotroph using CO_2 as only carbon source. We also studied the bioproduction capabilities of organic chemicals and amino acids by the synthetic lichens. Furthermore, we validated our growth and flux distribution predictions using physiological data, untargeted metabolomics, and expression data.

Supported under award DE-SC0019388, Biological and Environmental Research

QPSI: Quantitative Plant Science Initiative

Crysten E. Blaby-Haas*[1] (cblaby@bnl.gov), Agnidipta Ghosh[2], Miriam Pasquini[1], Nam Nguyen[3], Sam Seaver[4], Ping Zhu[1], Daifeng Wang[3], Steven C. Almo[2], Doreen Ware[5,6], Qun Liu[1], Ian K. Blaby[1]

[1]Brookhaven National Laboratory, Upton, NY; [2]Albert Einstein College of Medicine, Bronx, NY; [3]Stony Brook University, Stony Brook, NY; [4]Argonne National Laboratory, Argonne, IL; [5]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; [6]US Department of Agriculture, Agricultural Research Service, Ithaca, NY, USA.

Project Goals: QPSI (Quantitative Plant Science Initiative) is a DOE-mission driven, interdisciplinary, team-based capability that aims to accelerate the acquisition of a core knowledgebase of experimentally validated protein family function. Knowing what and how to engineer and predicting the consequence of that redesign on complex systems, such as a bioenergy crop, requires a systems biology approach grounded in molecular-level understanding. Because of the lack of sequence-based function understanding that affects all BER-relevant plant genomes, our goals are to: (1) reduce uncertainty in plant protein function annotation through integrated computational and experimental approaches, (2) identify genome-based principles underlying highly conserved processes over plant evolution, and (3) define fundamental rules of sequence-function relationships that translate across bioenergy crops. The resulting knowledge will underpin genome-based functional genomics experiments, while accelerating a sequence-based understanding of genotype-to-phenotype in specific plants.

Bioenergy crops that thrive in marginal soils and maintain performance in diverse and fluctuating environments are an essential component of a sustainable bioeconomy. Understanding and predicting biosystem productivity in these environments is hindered by a lack of sequence-to-function information. This knowledge gap impedes accomplishing Biological and Environmental Research (BER) mission to enable the design and reengineering of plants underpinning energy independence. In recognition of the critical role that knowledge of sequence-based protein function plays in the mission of BER and BSSD, Brookhaven National Laboratory (BNL) has established the Quantitative Plant Science Initiative (QPSI). The overarching goal of QPSI is to integrate experimentation with computational approaches to address the knowledge gap that exists between plant genomes and the function encoded within.

Genomics has accelerated our ability to gather systems-wide information regarding plant responses and adaptations to the environment. QPSI aims to capitalize on this cascade of genome-based data by combining scalable functional genomics and computational approaches with protein function characterization. The described activities leverage *in silico* 'omics-derived functional extrapolations and high-throughput *in vitro/in vivo* experimentation. Computational and experimental support is integrated with our *in silico* pipeline for evidence-based propagation of

annotations across bioenergy crops. The resulting gain of foundational knowledge will provide the guiding biological principles for the discovery of niche-specific and lineage-wide processes as targets for the improvement of plant-based bioenergy.

Presently, our priorities are divided into two main objectives:

Objective 1: generate evidence-based function predictions for conserved bioenergy crop protein families and provide a framework for disseminating process-level understanding of plant biology. This objective involves phylogenomic analyses of plant protein families and identification of gene regulatory networks. We are extending a computational platform, already integrated with DOE's Systems Biology Knowledgebase (KBase), for plant subsystem construction and interpretation. Subsystems provide a novel framework for *in silico* hypothesis building and testing in plants, as well as contextualized transfer of knowledge from one genome to another.

Objective 2: experimentally characterize the function of conserved plant protein families. We are generating experimental evidence for the function of conserved protein families using high-throughput (HTP) techniques. QPSI has built a custom HTP automation platform that leverages microbial phototrophs to accelerate the generation of protein family functional information. HTP protein expression/purification approaches are being leveraged to accelerate the generation of *in vitro* experimentally supported protein function evidence. We are also testing the ability of our family-centric computational and experimental functional studies to predict the phenotype of putative loss-of-function alleles in the bioenergy crop *Sorghum bicolor*. Finally, we are performing detailed characterization to provide the missing sequence-to-structure-to-function understanding for specific bioengineering targets.

Here, we present an update on the development of our capability and present progress towards our short-term objectives.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER).

Phylogenomics-guided Approaches for Predicting and Characterizing Functions Across a Multi-functional Protein Family

Miriam Pasquini* (mpasquini@bnl.gov) and **Crysten E. Blaby-Haas**

Brookhaven National Laboratory, Upton, NY

Project Goals: QPSI (Quantitative Plant Science Initiative) is a DOE-mission driven, interdisciplinary, team-based capability that aims to accelerate the acquisition of a core knowledgebase of experimentally validated protein family function. Knowing what and how to engineer and predicting the consequence of that redesign on complex systems, such as a bioenergy crop, requires a systems biology approach grounded in molecular-level understanding. Because of the lack of sequence-based function understanding that affects all BER-relevant plant genomes, our goals are to: (1) reduce uncertainty in plant protein function annotation through integrated computational and experimental approaches, (2) identify genome-based principles underlying highly conserved processes over plant evolution, and (3) define fundamental rules of sequence-function relationships that translate across bioenergy crops. The resulting knowledge will underpin genome-based functional genomics experiments, while accelerating a sequence-based understanding of genotype-to-phenotype in specific plants.

Very few plant-specific proteins are characterized to the level required for inclusion in a redesign strategy. Indeed, most proteins in the plant lineage remain completely uncharacterized. This knowledge gap impedes accomplishing Biological and Environmental Research (BER) mission in enabling design and reengineering of plants for energy independence. However, it is impossible to experimentally characterize every protein: presently 99.8% of functional annotations across plant genomes are predictions. These predictions at the genome-wide scale are typically based on sequence similarity to a database "best hit", for example, by searching against NCBI's RefSeq database using BLAST or using profile hidden Markov models to search domains in Pfam. The annotation from the best-scoring hit is then transferred to the uncharacterized gene product. Unfortunately, in the case of BLAST-based annotations, the "best hit" was also likely annotated by a BLAST search, and the relationship to the original experimentally annotated protein is often lost, and distantly taxonomically related. In the case of model-based annotation, the domain annotation is typically derived from characterization of a single bacterial, yeast, or animal protein and/or mutant phenotype. The result is that over the last 20 years there has been significant error-prone propagation of functional annotations from one genome to another.

The functional annotation challenge is particularly acute for large proteins families that contain multiple paralogs and subfamilies that contain members from taxonomically divergent lineages. With recent advancements in genomics, post-genomic experimentation and high-throughput (HTP) experimental tools, we are able to address protein function in a critical way: integrating genomics, functional genomics, genetics, biochemistry and biophysical characterization. Using

this multi-disciplinary approach, we are addressing the challenge of accurately annotating large, functionally diverse protein families. We are employing large-scale phylogenomic analyses combined with conserved gene neighborhood detection, co-expression networks, co-occurrence profiles, and protein fusion discovery for function prediction. We then take our predictions to the bench for experimental characterization, which further guides evidence-based propagation of annotations across sequenced space. Here, we describe the phylogenomics-guided discovery and biochemical characterization of a protein family that we predict to be required for the maturation of metal-dependent proteins involved in processes that range from bacteriochlorophyll and cobalamin biosynthesis to post-translational modification of chloroplast-localized proteins. In addition, we will present our results toward the experimental verification of these functions, including meta-functional and sub-functional activity assays, and planned experimentation toward understanding the phenotypic manifestation of protein function in *Sorghum bicolor*.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER)

Integration of Physiological Phenotyping and Cell-Type-Specific Omics Approaches to Study Individual and Abiotic Stress Combinations in Poplar

Amir H. Ahkami [1]* (amir.ahkami@pnnl.gov), Maria Del Mar Rubio Wilhelmi [4], Vimal K. Balasubramanian [1], Tanya Winkler [1], Ryan Kelly [1†], Yiran Liang [1†], Samuel Purvine [1], Ying Zhu [1], William Chrisler [1], Yuliya Farris [1], Ziyu Dai [1], Lye Meng Markillie [1], Galya Orr [1], Dusan Velickovic [1], Bryan Stanfill [1], C. Neal Stewart, Jr.[2], Stephen DiFazio [3] and Eduardo Blumwald [4]

[1] Pacific Northwest National Laboratory (PNNL), Richland, WA; **[2]** Dept. of Plant Sciences, University of Tennessee, Knoxville, TN; **[3]** Dept. of Biology, West Virginia University, Morgantown, WV; **[4]** Dept. of Plant Sciences, University of California, Davis, CA

[†] Present address: Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT.

Project Goals: The main goal of SyPro project is the development of transgenic trees with sustained photosynthetic activity and increased biomass production under the simultaneous occurrence of water deficit, increased soil salinity and elevated temperatures. To achieve that, we intend to (1) identify stress-responsive genes and proteins in specific cell-types of poplar leaves and roots; (2) discover novel *cis*-regulatory elements; (3) construct stress-responsive synthetic promoters; and (4) use these promoter-gene fusions to develop abiotic stress-tolerant poplar. The transgenic poplar trees will be evaluated under both controlled and field conditions.

Plant responses to environmental perturbations are dynamic and involve complex cross-talk between different regulatory pathways[1], including metabolic adjustments and gene/protein expression at cellular level for physiological and morphological adaptation at the whole-plant level [2]. Therefore, a single-cell-type analysis approach is needed to effectively reveal the underlying molecular mechanisms regulating developmental processes and plasticity when grown under suboptimal conditions.

In this work, clones of *Populus tremula x alba* (INRA 717 1-B4) were rooted for at least 25 days. Following rooting, plant response(s) to water deficit, salinity, heat, and the combination of all three stresses were monitored. Total biomass (root and shoot) was measured at the end of the experiment, with a reduction of 35%, 70%, 50% and 70% in dry weight of shoots and 40%, 60%, 40% and 70% depletion in dry weight of roots under salinity, water-deficit, heat and combination of three stresses, respectively. While photosynthesis levels and stomatal conductance were reduced under all stresses, the most significant effect was observed under drought and combined stress conditions.

At three sampling-time points (including recovery time), leaf and root tissues were collected, fixed and embedded for cell-type specific transcriptome and proteome analyses. We targeted distinct poplar cell types and tissues including leaf mesophyll, xylem/phloem, root epidermis and cortex cells using cryo-sectioning and laser-capture microdissection (LCM) techniques. For transcriptomics, RNA was extracted from 100-200 cells per cell-type. Full length cDNA and template libraries were generated, and template quantification and preparation for sequencing is in progress. For proteomics, total protein was extracted from 300-800 cells per cell-type, and samples were processed within ultrasmall-volume "nanowells" (nanoPOTS technology)

[3]. Focusing on leaf tissue, a total of 7,515 and 6,023 proteins were identified as palisade mesophyll and vascular-specific proteins, respectively, under all investigated conditions. Among these, 273 (in palisade cells) and 332 (in vascular cells) proteins were identified as candidate cell-type specific abiotic stress related-proteins. For example, several proteins involved in photosynthesis, chloroplast division, carbon allocation to sink tissues, chaperone binding and chromatin organization were identified as drought-, salt-, or combined stress-responsive proteins exclusively in palisade or vascular cells. Our results provide information that is missing in whole tissue-based analyses, including cell population-specific protein profiles that are unique for the distinct cellular layers of poplar leaves and roots. The obtained information (sequences) is being used for motif discovery using bioinformatics approaches and then for promoter engineering for subsequent poplar transformation.

References:

1. L. Shaar-Moshe, E. Blumwald and Z. Peleg. 2017. Unique physiological and transcriptional shifts under combinations of salinity, drought, and heat. *Plant Physiology*, 174 (1) 421-434.
2. Julia Krasensky and Claudia Jonak. 2012. Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. *Journal of Experimental Botany*, 63 (4) 1593–1608.
3. Zhu, Y.; Piehowski, P. D.; Zhao, R.; Chen, J.; Shen, Y.; Moore, R. J.; Shukla, A. K.; Petyuk, V. A.; Campbell-Thompson, M.; Mathews, C. E. 2018. Nanodroplet processing platform for deep and quantitative proteome profiling of 10–100 mammalian cells. en. *Nature Communications*, 9, 882.

The SyPro Poplar project is supported by the U. S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), Genomic Science Program.

Discovery of Novel *Cis*-Regulatory Elements Responsive to Salt-Stress in Hybrid Poplar for Designing Inducible Synthetic Promoters

JunHyung Lee[1]* (jlee220@utk.edu), Yongil Yang[1,2], Yuanhua Shao[1], Magen Poindexter[1], Stephen P. DiFazio[3], Amir H. Ahkami[4], Eduardo Blumwald[5], and **C. Neal Stewart, Jr.**[1,2]

[1]Department of Plant Sciences, University of Tennessee, Knoxville; [2]Center for Agricultural Synthetic Biology, Knoxville, TN; [3]West Virginia University, Morgantown, WV; [4]Pacific Northwest National Laboratory, Richland, WA; [5]University of California, Davis, CA

Project Goals: In SyPro Poplar we intend to (i) study the functions of selected stress-responsive genes; (ii) discover novel motifs and construct stress-responsive synthetic promoters; and (iii) use these promoters to drive the expression of genes shown to confer abiotic stress tolerance in a variety of crops and develop abiotic stress-tolerant poplar seedlings in a coordinated fashion. We will use a combinatorial gene stacking approach with key transgenes driven by stress-responsive synthetic promoters to confer stress resistance. Our plan is to develop a series of abiotic stress-responsive synthetic promoters comprised of the stretch of DNA containing multiple copies of abiotic responsive *cis*-motifs upstream of a core-promoter in which abiotic stress specific transcription factors (TFs) bind to their cognate sequences to drive transcription under multiple abiotic stresses. The aim is the development of transgenic trees with sustained photosynthetic activity and increased biomass production under individual and the simultaneous occurrence of water deficit, increased soil salinity and elevated temperatures.

Advanced -omics data combined with plant synthetic biology technologies are powerful tools to discover novel motifs associated with abiotic stress responses and to construct synthetic promoters to regulate stress-coping genes expression specifically to the stress conditions. In the present study, we designed synthetic promoters responsive to salt-stress using novel *cis*-regulatory elements in hybrid poplar (*Populus tremula* x *Populus alba*). Poplar-transcriptome data[1,2] were used as input for computational-based motif discovery that are being developed as a KBase module. The computational goal is to identify conserved minimal motifs in native plant gene promoters are responsible for upregulated gene expression under salt-stress. Through this process we identified a highly conserved 20-base-long motif, which we refer to as motif 16. Three native promoter variants, which included motif 16, were fused to a green fluorescent protein (GFP) reporter gene. Gene constructs were transfected into poplar mesophyll protoplasts and subjected to salt treatment, which resulted in GFP signal induction after 24 hr. The endogenous gene expression of the three genes in poplar leaves was confirmed by qRT-PCR using greenhouse-grown poplar plants under salt-stress. A synthetic promoter, comprising a 7-repeat (heptamer repeat) of the motif 16 that we placed upstream of the cauliflower mosaic virus (CaMV) -46 35S (minimal) promoter, which was used to drive a GFP reporter. We also tested another synthetic inducible promoter candidate containing heptameric repeats of the ABRE motif (ACGTG, a canonical motif responsive to salt-stress), in the same construct architecture. Each

synthetic promoter strongly induced GFP expression after 24 hr under salt-treatment in poplar mesophyll protoplasts. We are currently in the process of producing transgenic poplar with constitutively-expressed GFP and GFP under the control of three native and two synthetic salt-inducible promoters. Furthermore, we are producing a permutational library using motif 16 to produce 5-base-long fragments for the purpose of discovering the minimal core-motifs responsible for salt-stress induction. The top-performing synthetic promoters will be subsequently deployed to control key abiotic stress resistance genes in poplar with the goal of producing robust and sustainable feedstock.

References

1. Zheng et al. (2015) Transcriptomic analysis reveals importance of ROS and phytohormones in response to short-term salinity stress in *Populus tomentosa. Frontiers Plant Sci.* 6:678.
2. Filichkin et al. (2018) Abiotic stresses modulate landscape of poplar transcriptome via alternative splicing, differential intron retention, and isoform ratio switching. *Frontiers Plant Sci.* 9:5.

The SyPro Poplar project is supported by the Office of Biological and Environmental Research in the Department of Energy Office of Science.

Microbial environmental feedbacks and the evolution of soil organic matter

Nicholas J Bouskill[1]

[1]Climate and Ecosystem Sciences Division, Lawrence Berkerley National Laboratory, Berkeley, CA, 94710.

The vast majority of Earth's organic matter is stored in soil. The products of microbial metabolism as well as dead microbes (necromass), along with residues from plants and other organisms at different stages of decomposition, constitute a large fraction of that soil organic matter (SOM). The ability of microbes to modify and degrade SOM depends on physicochemical characteristics of the soil, affecting SOM stability and persistence. While the contributions of microbes to the decomposition and loss of SOM have been intensively studied, their role in maintaining the terrestrial SOM is poorly understood. Specifically, how fungi, bacteria, and archaea participate in SOM production, its interaction with minerals, and the formation of soil aggregates remains a significant gap in our understanding of the terrestrial nutrient cycle. The chemical composition of SOM is in large measure determined by soil bacterial metabolism, which is impacted by changes in rainfall patterns. This talk will describe progress within the first year of this project. Overall, we employ field and laboratory experiments and computational modeling to understand the role of microbial communities in stabilizing SOM under drought in tropical soils

Enabling Predictive Metabolic Modeling of Diurnal Growth Using a Multi-Scale Multi-Paradigm Approach

Joseph J. Gardner[1] and **Nanette Boyle**[1]

[1]Chemical and Biological Engineering, Colorado School of Mines, Golden, CO 80401, USA

Project Goals: The main goal of this project is to develop more predictive metabolic models of diurnal growth for algal systems. We are using a multi-paradigm multi-scale approach which enables us to include phenomenon not previously integrated into metabolic models, such as diel light, diffusion of metabolites/nutrients, cell-cell interactions, as well as temporal and spatial tracking of cells. This model will further be enhanced with experimental data collected over 24 hour diel growth for transcript abundance and changes in biomass composition. Validation and improvement of the model will be performed by comparing predictions to 13C-MFA of cells grown in the lab as well as in large outdoor ponds.

Photosynthetic microorganisms have the potential to become sustainable sources of fuels; however, we have yet to harness their full power due to a general lack of tools for engineering their metabolism. Metabolic models have been shown to drastically reduce the development time for commercial production strains of heterotrophic bacteria; however, they are less applicable to photosynthetic systems due to the transient nature of diurnal (day/night) growth, which is the most economical option for large-scale production because energy provided to the cells can come strictly from sunlight alone. Current metabolic models are not capable of accurately predicting growth rates in day/night growth cycles, let alone genetic changes which would lead to increased yields. Therefore, it is imperative that the unique physiology of photosynthetic organisms be integrated into the next generation of metabolic models so that they can be used to design, build, and test photocatalytic biofuel production strains. The availability of such models will introduce a new frontier in the ability to use *in silico* tools to investigate the metabolism, growth and phenotype of photosynthetic microorganisms. It will enable us to gain insight into why photosynthetic organisms have drastically different productivities when grown in continuous light compared to diurnal cycles and how to circumvent this. We will present our work to date on the development of an advanced multi-paradigm multi-scale metabolic model of the green alga, *Chromochloris zofingiensis*, using an agent-based modeling framework that both incorporates, and is validated, by experimental data.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0019171

Microbial Responses to Pyrogenic Organic Matter: Cultivating Isolates, Ageing PyOM, Tracing Gas Fluxes, and Meta-analysis

Timothy D. Berry,[1*](tdberry@wisc.edu), Nayela Zeba,[1] Jamie Woolet,[1] **Thomas D Bruns**,[2] Matthew Traxler,[2] Igor Grigoriev,[3] and Thea Whitman[1]

[1]Dept. Soil Science, Univ. Wisconsin, Madison; [2]Dept. Plant & Microbial Biol., Univ. Cal., Berkeley, CA; [3]US DOE Joint Genome Institute, Walnut Creek, CA

http://plantandmicrobiology.berkeley.edu/profile/bruns
https://whitmanlab.soils.wisc.edu/research/

Project Goals: The frequency of large, high severity wild fires is increasing in the western US and in regions around the world due to long-term fire suppression strategies and climate change [1]. These fires have direct, negative effects on soil carbon stocks through combustion, but they have indirect and potentially positive effects on soil carbon stocks through the production of pyrogenic organic matter (PyOM) that has a long residence time and constitutes a major pool of C in fire-prone ecosystems [2-4]. Soil microbes are likely to be involved with the degradation of all these compounds, yet little is currently known about the organisms or metabolic processes involved. **We are dissecting the effects of microbes on post-fire soil carbon dynamics by using a systems biology approach that couples small experimental "pyrocosms", highly controlled production of ^{13}C-labeled pyrolyzed substrates, genomics, transcriptomics, stable isotope techniques, and mass spectrometry to address the following objectives:** (1) Develop improved genomic and other -omic resources for the dominant microbes of fire-affected soils; (2) Determine the temporal response of soil microbes to fire and to PyOM additions; (3) Characterize the temporal patterns of degradation of different sub-fractions of PyOM.

The fungal community in fire-affected soils is increasingly well characterized, and the broader team continues to develop genomic, proteomic, and metabolomic resources for the study of these organisms. In contrast, the bacterial community in fire-affected soils is less well understood and in need of further study. To help meet this need we have isolated >80 bacterial strains from fire-affected soil with the potential to mineralize PyOM. Sequencing of the 16S rRNA genes have identified the majority of theses strains as being members of the *Actinobacteria* phylum, with *Streptomyces* and *Pseudonocardia* being the most prominent.

A custom gas sampling device, the multiplexer, has been constructed to facilitate the analysis of PyOM mineralization by bacterial isolates and whole soils during incubations with ^{13}C enriched PyOM. In conjunction with a Picarro cavity ringdown spectrometer, the multiplexer allows for rapid measurement of the CO_2 concentration and isotopic composition in the headspace of up to 56 incubation vessels. We present data from a small-scale soil incubation using ^{13}C enriched sugar maple PyOM to demonstrate the utility of the instrument. While typical incubations using an isotope ratio mass spectrometer might achieve one measurement per day, we sample the headspace of each soil incubation 11 times over 36 hours; this frequent sampling allows for the detection of fine-scale CO_2 flux dynamics obscured by other methods. Additionally, the large sample number accommodated by the multiplexer allows for the simultaneous inclusion several treatments with replication in a single experiment.

Our understanding of the chemical and biological stability of PyOM has been revised in recent

years with studies observing a change in the physico-chemical properties of PyOM when it is deposited in soil [5]. These changes can occur after PyOM has persisted in soil for a very long time and the process is often referred to as 'ageing' of PyOM [6]. PyOM ageing is likely to influence the degradation of PyOM materials by soil microbes, consequently affecting native soil carbon stocks. We present data from an experiment where we compared pyrogenic carbon mineralization between aged and unaged PyOM by *Streptomyces sp.,* isolated from fire affected soils. We aged eastern white pine wood PyOM produced at 350 and 550 °C using a combination of physical and chemical treatments in lab to accelerate the process of ageing. We incubated aged and unaged pine wood PyOM with *Streptomyces sp.* for a period of 14 days during which the pyrogenic carbon mineralized was measured by sampling the head space of each incubation every 48 hours over a period of 14 days using the Picarro-multiplexer setup. We observed that both chemical and physical ageing reduced pyrogenic carbon mineralization compared to unaged PyOM in the case of 350 °C PyOM. However, in the case of 550 °C PyOM, we observed higher pyrogenic carbon mineralization for the chemically aged PyOM compared to both physically and unaged PyOM incubations.

Finally, we performed a meta-analysis of the effect of PyOM additions on soil bacterial community composition, with the goal of determining: (1) Is there a detectable and consistent "charosphere" community that characterizes PyOM-amended soils – *i.e.*, which is more important for determining bacterial community composition, soil properties or PyOM additions? (2) Are there consistent responses at the phylum level to PyOM amendments? (3) Can we identify individual PyOM-responsive taxa that increase in relative abundance consistently across different soil types? We did not observe consistent community or phylum-level responses to PyOM amendments across studies. However, we did identify genera that tend to be enriched with PyOM additions, including *Sphingomonas sp., Nocardioides sp.*, and *Mesorhizobium sp.*, which may be putative PyOM-degraders.

References

1. Westerling, A.L., H.G. Hidalgo, D.R. Cayan, and T.W. Swetnam, Warming and earlier spring increase western US forest wildfire activity. *Science*, 2006. 313: 940-943.
2. Gonzalez-Perez, J.A., F.J. Gonzalez-Vila, G. Almendros, and H. Knicker, The effect of fire on soil organic matter - a review. *Environment International,* 2004. 30: 855-870.
3. Santin, C., S.H. Doerr, C.M. Preston, and G. Gonzalez-Rodriguez, Pyrogenic organic matter production from wildfires: a missing sink in the global carbon cycle. *Global Change Biology*, 2015. 21: 1621-1633.
4. Lehmann, J. *et al.*, Australian climate-carbon cycle feedback reduced by soil black carbon. *Nature Geoscience*, 2008. 1: 832-835.
5. Cheng, C. *et al*., Oxidation of black carbon by biotic and abiotic processes. *Organic Geochemistry*, 2006. 37: 1477–1488.
6. Heitkötter, J., and Marschner, B., Interactive effects of biochar ageing in soils related to feedstock, pyrolysis temperature, and historic charcoal production. *Geoderma*, 2015. 245–246: 56–64.

Funding Statement: *This work was funded by the DOE, Systems Biology Enabled Research on the Roles of Microbial Communities in Carbon Cycle Processes program, grant DE-SC0016365 to T.D. Bruns, T. Whitman, M. Traxler, and I. Grigoriev. The work conducted by the U.S. DOE Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. DOE under Contract No. DE-AC02-05CH11231.*

Functional genomics of pyrophilous fungi – determining the fate of pyrolyzed carbon in post-fire soils

Steindorff AS[1], Carver A[2], Calhoun S[1]* (secalhoun@lbl.gov), Seong K[2], Mondo SJ[1], Stillman K[2], Liu H[2], LaButti K[1], **Bruns TD**[2], Grigoriev IV[1,2]

[1]US DOE Joint Genome Institute, Walnut Creek, CA; [2]Plant and Microbial Biology, UC Berkeley, Berkeley, CA

https://genome.jgi.doe.gov/Pyrophilous_Fungi

Project Goals: Our overall goal is to understand how the post-fire fungal community affects the fates of pyrolyzed carbon and soil carbon in early post-fire soil environments. We will address this overall goal in the context of the following objectives: 1) Develop improved genomic and other -omic resources for the dominant fungi in this system. 2) Determine the temporal response of soil microbes to fire and to pyrolyzed organic matter additions under controlled conditions. 3) Characterize the extractable fractions of pyrolyzed organic matter, their temporal patterns of degradation, their effects on soil organic carbon mineralization, and the soil microbes driving these processes.

Forest fires are a necessary disturbance in forest ecosystems and historically occur at regular intervals. Although forest fires are a natural occurrence, severe forest fires affect larger areas of land than is expected for their fire regimes, causing damage to wildlife and human communities. Forest fire recovery depends on the ecological succession of soil microbial organisms as well as flora and fauna, but little is known at the microbial level. What is known is that forest soil changes greatly after fire, with an enrichment of dead biomass and partially pyrolyzed or "charred" carbon compounds. Soil chemistry also changes, with melted waxes, lipids and cutins creating hydrophobic conditions that can lead to erosion and landslides. In the direct aftermath of forest fires, post fire soils are dominated by pyrophilous fungi, or filamentous fungi that are known to fruit exclusively on burned soil. As the earliest soil fungal colonizers, pyrophilous fungi may have an important role in post fire soil recovery and may help determine the fate of post fire soil carbon, but little is known about the organisms or their adaptation to their environment. Here, we have sequenced, annotated, and compared the genomes of 10 pyrophilous fungi (7 Ascomycetes and 4 Basidiomycetes) to 24 of their non-pyrophilous relatives and found expansions in gene families related to substrate metabolism, stress response capabilities, and secondary metabolites. Amongst pyrophilous ascomycetes species, we found expansions in many substrate utilization genes including CAZymes such as curdlanases (GH132), cutinases (CE5), carbohydrate binding modules (CBM 50), chitin binding modules (CBM 14), and phospholipases. Expanded protease families include serine aminopeptidases, metalloproteases, Caspases (C14B), and ubiquitin-specific proteases. On the basidiomycetes pyrophilous species, we found several rapidly expanded families such as xylanases (GH10), xylan esterases or cutinases (CE5), multicopper-oxidase (AA1), metallopeptidases, p450, ABC/MFS transporters,

and DNA repair gene XPG1. *Coprinellus angulatus* genome showed the highest rate of expansions compared with genomes of other *Coprinellus* species, totaling 61 families (22 containing signal peptide) and 584 genes including heat-shock protein hsp70, chitin synthase III and signaling proteins G-alpha/Pkinase. This study will help us increase our knowledge of adaptive advantages that have evolved in pyrophilous fungi and the role of these organisms in post fire succession.

This work was funded by the Department of Energy, Systems Biology Enabled Research on the Roles of Microbial Communities in Carbon Cycle Processes program, grant DE-SC0016365.

Rising From the Ashes: Succession and Metabolism of Post-Fire Fungi

Monika S. Fischer*[1] (monikaf@berkeley.edu), Neem J. Patel[1], Sydney I. Glassman[1,2], Thomas D. Bruns[1], **Matthew F. Traxler**[1]
[1]University of California, Berkeley, [2]University of California, Riverside

Project Goals:

- To define the pattern of post-fire fungal succession across a gradient of fire severities in the Sierra Nevada mountains.
- To elucidate the nutrient niche of post-fire fungi.
- To identify how post-fire fungal communities affect the fate of pyrolyzed carbon stocks and soil organic matter.

Wildfires increasingly pose a threat to ecosystems and communities across the western United States. Given the current trajectory of climate change, wildfires are predicted to dramatically increase in size and severity. We are already witnessing this effect. Each year, wildfires have set new records for the amount of land burned and destruction caused. One critically important aspect of mitigating the long-term impacts of wildfires is understanding and promoting ecosystem recovery after wildfires. Fire transforms soil organic carbon into pyrolyzed carbon, which alters the accessibility of carbon for soil dwelling organisms that depend on soil carbon stocks for survival. Soil bacteria and fungi consume soil carbon and are also are vitally important for the growth and survival of plants. While decades of research has defined the dynamics of post-fire plant succession, we have only recently acquired the tools necessary (i.e. environmental high-throughput sequencing) to begin to define post-fire fungal succession and how this process supports the reestablishment of plants. Serendipitously, in 2013 the Rim Fire burned two long-term sampling plots near Yosemite National Park (YNP). Fungal community sequencing of soil samples in addition to observing the development of fungal fruiting bodies for one year after the fire demonstrated a clear pattern of succession. In an ongoing study at the Blodgett Forest Research Station (~70miles north of YNP), we are working to confirm whether or not the fungal succession pattern observed after the Rim Fire is predictable and reproducible, similar to how post-fire plant succession is generally predictable. Within two months after the Rim Fire, the quick-growing Ascomycete *Pyronema omphalodes* dominated. The Ascomycete *Morchella spp*. dominated roughly four months post-fire, and then the Basidiomycete *Pholiota molesta* dominated roughly ten months post-fire. Representatives of these three species are easily cultured in the lab, and we are using these cultures to directly investigate the metabolism of these fungi. Our preliminary data indicates that the nutrient niche of these fungi is broad, and that these fungi respond uniquely to various chemical aspects of pyrolyzed soil. Ongoing work will determine how the chemistry of post-fire soils changes over time, connecting specific fungi with the transformation of specific compounds in these soils. Together, these data will determine the fate of pyrolyzed carbon, and demonstrate how fire-adapted fungi survive, thrive, and promote the re-establishment of forest ecosystems after wild fires.

Microbial Community Succession and Metabolism in Post-Fire Environments

Neem J. Patel[1]* (neempatel@berkeley.edu), Monika S. Fischer[1], Thomas D. Bruns[1], **Matthew F. Traxler**[1]
University of California, Berkeley[1]

Project Goals

- Determine the temporal responses of soil microbes in fire-effected soils and identify patterns of succession of bacterial communities.
- Characterize the extractable chemical fractions of pyrolyzed organic matter, their temporal patterns of degradation and their effects on soil organic carbon mineralization.
- Identify how the post-fire bacterial communities affect the fates of pyrolyzed carbon stocks and soil organic matter and drive degradation.

Over the past few decades, we have observed an increased frequency of high-severity wild-fires, a consequence of long-term fire suppression strategies and further exacerbated by climate change. These fires have a significant impact on soil carbon stocks: a direct, negative effect due to combustion and an indirect positive effect, via the production of pyrolyzed organic matter (PyOM). PyOM is chemically heterogeneous and while a small fraction of it can be mineralized by microbes, much of it is highly aromatic with slow decomposition rates, creating a long-term carbon sink. Reports suggest that PyOM may constitute up to 80% of total soil organic carbon (SOC) in fire-effected ecosystems. Thus, it is important to chemically characterize and study the constituents of PyOM and total SOC in these environments and elucidate the role microbial communities play in their fates.

Fire-fueled changes also dramatically alter soil microbial communities. There is a considerable decrease in total soil biomass. Rehabilitation of these fire-affected soil is acutely delayed - requiring decades to fully re-establish healthy and diverse communities. Studies suggest that this delay in recovery is due to changes in functional capabilities (i.e. metabolic potentials and rates of organic matter decomposition) of the organisms utilizing this transformed soil. While we know that fires have dramatic effects on soil communities, the fundamental mechanisms and key taxa involved in driving the recovery of these communities remain enigmatic.

Recent work by Dr. Thomas Bruns' lab, has identified a group of pyrophilous fungi as the first colonizers of these pyrolyzed soils. One of these fungi, *Pyronema omphalodes*, was found at 26 sample locations in Stanislaus National Forest within the first month following the Rim Fire in Northern California. The Bruns' lab continued sampling these sites over the course of a year and found a clear, temporal, pattern of succession which was dominated by known pyrophilous fungi. Much less is known about the bacterial communities in these environments. Preliminary experiments in Dr. Thea Whitmans' lab have shown that bacterial communities can be selectively enriched by the addition of PyOM, some of these pyrophilous bacteria have been identified. Further the Whitman lab has shown that the dominant responders are members of the phyla Actinobacteria, Proteobacteria and Bacteroidetes. However, environmental studies have yet to be done to complement this work.

In order to close the gap and gain a clearer picture of the affects of fire on soil and microbial community recovery, we are conducting a year-long field study at Blodgett Forest Research Station in Georgetown, CA. We will use 16S rDNA bacterial community profiling and hope to observe a successional pattern, similar to the pyrophilous fungi observed by the Bruns' lab. Addi-

tionally, we will us LC/MS analysis to understand the pre- to post-fire soil chemistry and further understand how this changes over time. Together these data have the power to unveil the temporal patterns of PyOM degradation and the microbial powerhouses associated with these actions. Finally, we aim to recapitulate these observations in pyrocosms, which are experimentally tractable systems we use to mimic wild-fire conditions. We hope this work will further scientists' knowledge, insight and expertise of these blistered ecosystems and the community of organisms associated with them. Understanding microbial community re-establishment is essential for the secondary succession of plant communities, thus elucidating the initial stages of succession will greatly aide in the recovery of our forest ecosystems.

References

1. Westerling, A.L., H.G. Hidalgo, D.R. Cayan, and T.W. Swetnam, Warming and earlier spring increase western US forest wildfire activity. Science, 2006.313(5789): p. 940-943.
2. Gonzalez-Perez, J.A., F.J. Gonzalez-Vila, G. Almendros, and H. Knicker, The effect of fire on soil organic matter - a review. Environment International, 2004.30(6): p. 855-870.
3. Lehmann, J., J. Skjemstad, S. Sohi, J. Carter, M. Barson, P. Falloon, K. Coleman, P. Woodbury, and E. Krull, Australian climate-carbon cycle feedback reduced by soil black carbon. Nature Geoscience, 2008. 1(12): p. 832-835.
4. Köster, K., F. Berninger, A. Lindén, E. Köster, and J. Pumpanen, Recovery in fungal biomass is related to decrease in soil organic matter turnover time in a boreal fire chronosequence, in Geoderma2014, Elsevier B.V. p. 74-82.
5. Fontúrbel, M.T., A. Barreiro, J.A. Vega, A. Martín, E. Jiménez, T. Carballas, C. Fernández, and M. Díaz-Raviña, Effects of an experimental fire and post-fire stabilization treatments on soil microbial communities, in Geoderma2012, Elsevier B.V. p. 51-60.
6. Seaver, F.T., Studies in Pyrophilous Fungi: I. The Occurrence and Cultivation of Pyronema. Mycologia, 1909. 1(4): p. 131-139.
7. Whitman, T., Z. Zhu, and J. Lehmann, Carbon Mineralizability Determines Interactive Effects on Mineralization of Pyrogenic Organic Matter and Soil Organic Carbon, in Environmental Science & Technology2014. p. 13727-13734.

The Ecophysiological Determinants of Mineral-Associated Carbon in Agricultural Soil

Roland C. Wilhelm* (rwilhelm@cornell.edu) and **Daniel H. Buckley**

School of Integrative Plant Science, Cornell University, Ithaca, NY

Project Goals

Our project aims to develop understanding of the ecological determinants of microbially-processed soil carbon (C) fated to become mineral-associated C (MAC). We aim to identify ecological traits and genomic signatures of bacteria that are predictive of the quantity and/or quality of MAC. We study the trade-offs between growth rate and metabolite production / carbon utilization efficiency in bacteria common to agricultural soils, where management favors fast-growing, ruderal organisms and the gradual loss of MAC. By linking ecology, metabolite production and MAC, our research tests the ecophysiological controls on MAC formation and identifies where such information can be used to improve C-cycle modeling and management practices.

Abstract

Microbiota in cropland soils experience punctuated changes in physical, chemical and biological properties at varying frequency and intensity depending on management practices. Disturbances favor organisms that respond rapidly, and often transiently, to new conditions (i.e. 'zymogenous' species), while negatively affecting those with competitive strategies favored by periods of stability (i.e. late-stage or 'climax' species). While members of these groups broadly differ in their metabolic capabilities, we hypothesize that the general trade-offs necessary for rapid growth may exert an overarching influence on the quality and quantity of MAC produced, stemming from limits to the richness of an organism's metabolome and/or carbon utilization efficiency (CUE). We selected cosmopolitan soil bacteria isolates representing each ecological group based upon physiological, genomic and environmental characterizations with pairings obtained from seven major phylogenetic groups found in soil (belonging to *Actinobacteria, Alphaproteobacteria* and *Firmicutes*). We then characterized CUE and exometabolite profiles (i.e. low molecular weight extra-cellular metabolites) of each during exponential growth (replete conditions) and during maintenance growth (limiting conditions) in a soil-defined media. Exometabolite samples were then filtered through a set of common minerals, representing end-members for hydrophilicity (Montmorillonite, Kaolinite, Illite and Hematite) and hydrophobicity (Pyrophyllite and Talc), to characterize MAC-forming potential. While this work is on-going, we will present an overview of isolate selection and highlight genomic signatures that define the ecological groups. These signatures offer potential parameters for microbe-explicit C-cycling models.

Funding Statement

This material is based upon work supported by the Department of Energy, Office of Biological & Environmental Research Genomic Science Program under Award Numbers DE-SC0010558 and DE-SC0004486.

Simulation and Experiment Uncover Complementary Regulation in the Circadian System *Neurospora crassa*

William R. Cannon[1]*, Jennifer M. Hurley[4], Meaghan S. Jankowski[4], Jeremy D. Zucker[1], Douglas J. Baxter[2], Neeraj Kumar[1], Scott E. Baker[3], Tina Kelliher[5], Hannah De Los Santos[4], **Jay C. Dunlap**[5]

[1] Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, 99352;

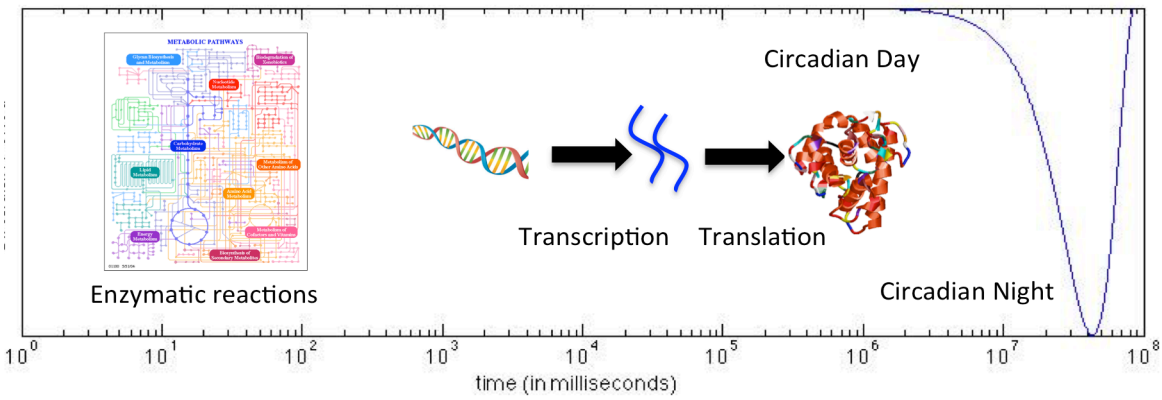
[2] Research Computing Group, Pacific Northwest National Laboratory, Richland, WA, 99352

[3] Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, 99352

[4] Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY, 12180

[5] Department of Molecular and Systems Biology, Geisel School of Medicine at Dartmouth, Hanover, NH 03755

Project Goals: The goal of this research is to develop and implement a new computational and theoretical method for modeling biological systems that fills a gap in modeling mass action dynamics. Based on statistical thermodynamics, the method bridges data-poor scales (parameters for mass action kinetics) and data-rich scales (chemical potentials of metabolites, and metabolite, protein & transcript data) to enable predictive modeling from enzymatic reactions (10^{-3} to 10^{0} s^{-1})



Timescales that the simulations using statistical thermodynamics will cover. Enzymatic reactions occur on the millisecond to second timescale while gene and protein expression occur on the minute to ~30-minute scale and the circadian rhythm occurs over a period of 24 hours.

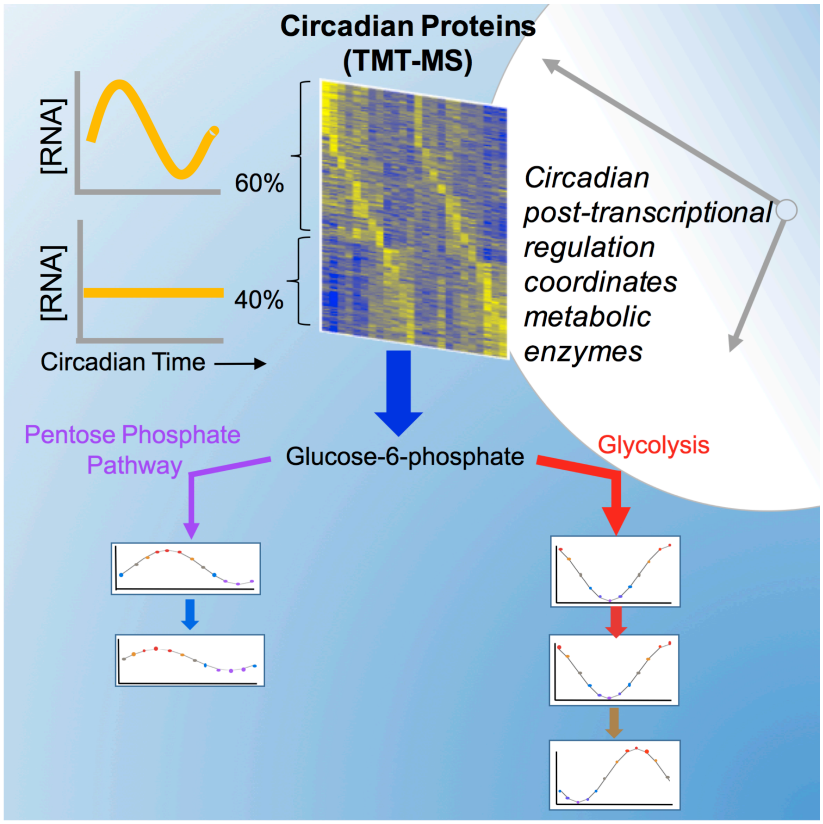
to gene and protein regulation (~20 minutes) to circadian rhythms (24 hours).

Progress: Nature selects those organisms that can reproduce the fastest while maintaining fitness and extracting the least amount of energy from their environment. This challenge makes regulation of natural systems critically important. We report the computational prediction of regulation, metabolite levels and rate constants using a maximum entropy method [1], and the experimental detection of circadian regulation of proteins and transcripts [2].

The computational method is applied in four steps: (1) a new constrained optimization approach based on Marcelin's 1910 mass action equation is used to obtain the maximum entropy distribution, (2) the predicted metabolite concentrations are compared to those generally expected from experiment using a loss function from which post-translational regulation of enzymes is inferred, (3) the system is re-optimized with the inferred regulation from which rate constants are determined from the metabolite concentrations and reaction fluxes, and finally (4) a

full ODE-based, mass action simulation with rate parameters and allosteric regulation is obtained. The method is applied to central metabolism and the flow of material through the three competing pathways of upper glycolysis, the non-oxidative pentose phosphate pathway, and the oxidative pentose phosphate pathway are evaluated as a function of the NADP/NADPH ratio.

The simulations complement experimental transcriptional and translational experiments performed over the circadian cycle of *Neurospora*. Transcriptional/translational feedback loops in fungi and animals drive circadian rhythms in transcript levels that provide output from the clock, but post-transcriptional mechanisms also contribute. To determine the extent and underlying source of this regulation, we applied novel analytical tools to a long-duration, deeply-sampled, circadian proteomics time course comprising half of the proteome. We found a quarter of expressed proteins are clock-regulated, but >40% of these do not arise from clock- regulated transcripts. Contrary to predictions, rhythmic protein degradation plays little role in post-transcriptional regulation but instead rhythms arise from oscillations in translation. Our data highlighted the impact of the clock on metabolic regulation, with central carbon metabolism reflecting both transcriptional and post-transcriptional control and opposing metabolic pathways showing peak activities at different times of day. CSP-1, a transcription factor with a role in metabolic regulation of the clock, contributes significantly to determining the rhythmicity and phase of clock-regulated proteins. The experimental data demonstrate that the rhythmic proteins within the Pentose-Phosphate pathway peak in the circadian morning, while conversely, in glycolysis and the TCA cycle, the rhythmic proteins peak in the circadian evening. That is, the rhythmic proteins of glycolysis are in anti-phase to the rhythmic proteins of the Pentose-Phosphate pathway.



References

1. Cannon, W.R., et al., *Prediction of metabolite concentrations, rate constants and post-translational regulation using maximum entropy-based simulations with application to central metabolism of Neurospora crassa.* Processes, 2018. **6**(6)https://doi.org/10.3390/pr6060063
2. Hurley, J.M., et al., *Circadian Proteomic Analysis Uncovers Mechanisms of Post-Transcriptional Regulation in Metabolic Pathways.* Cell Systems, 2018https://doi.org/10.1016/j.cels.2018.10.014

This project is supported by the U.S. Department of Energy's Office of Biological and Environmental Research and the National Institute of Biomedical Imaging and Bioengineering.

Screening Fungal Genome Sequencing Data and Culture Collections to Better Understand Bacterial:Fungal Interactions

Geoffrey L. House[1]* (ghouse@lanl.gov), Andrea Lohberger[2], Fabio Palmieri[2], La Verne Gallegos-Graves[1], Julia M. Kelliher[1], Demosthenes P. Morales[1], Armand E. K. Dichosa[1], Debora F. Rodrigues[3], Hang N. Nguyen[3], Saskia Bindschedler[2], Jean F. Challacombe[4], Jamey D. Young[5], Pilar Junier[2], and **Patrick S. G. Chain[1]**

[1]Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico; [2] Institute of Biology, University of Neuchâtel, Neuchâtel Switzerland; [3]Civil and Environmental Engineering, University of Houston, Houston, Texas; [4] College of Agricultural Sciences, Colorado State University, Fort Collins, Colorado; [5]School of Engineering Vanderbilt University, Nashville, Tennessee

Project Goals: Interactions between bacteria and fungi are critical to the above- and below-ground functioning of terrestrial ecosystems, yet little is known about these interactions or how they function. It is increasingly clear that these interactions also underpin multiple DOE research priorities, including understanding the possible effects of Earth system change, overcoming current and future energy and environmental challenges, and developing renewable energy sources. This Science Focus Area (SFA) project seeks to better understand of these bacterial:fungal interactions, including the diversity of bacteria and fungi involved, the functions of these interactions in a variety of conditions, and the mechanistic basis of these interactions. Here we outline computational methods we are using as part of this project to better understand the diversity of these bacterial:fungal interactions represented in publicly available DNA sequence data as well as in fungal culture collections.

https://genomicscience.energy.gov/research/sfas/lanlbfi.shtml

The amount of publicly available fungal genome sequencing data is increasing quickly, due in large part to the 1000 Fungal Genomes Project through the Joint Genome Institute (JGI), with data from over 1200 fungal isolates now represented in JGI's Mycocosm database. However, because many fungi have bacteria and viruses associated with them, these DNA sequence datasets from fungi may also begin to provide information about the associated microbiome of these fungal isolates. Furthermore, because the 1000 Fungal Genomes Initiative deliberately seeks to span the full range of known fungal diversity, this presents a unique opportunity to use these DNA sequences to start understanding the diversity of bacteria that may form associations with a wide range of fungi.

To this end, we have developed a bioinformatics pipeline that consists of commonly used tools and custom scripts to identify signals of bacteria that co-occur with hundreds of different fungal isolates. We begin the analysis with raw DNA sequencing reads from fungal genome projects and then remove all identifiable fungal DNA sequencing reads in order to reduce spurious similarities to bacteria. Next, we assemble the remaining reads into longer contigs that contain more information that can aid their taxonomic classification. For this analysis, we discard reads that do not assemble into contigs. We then use multiple taxonomy classifiers to identify contigs that signal the presence of specific bacteria.

Using this bioinformatics pipeline, we have been able to identify specific fungal isolates that have strong signals of likely associated bacteria (e.g. *Rhizobium* sp. and others), while other fungal isolates have strong signals of bacterial contamination (e.g. *Escherichia coli*). However, for other, ambiguous cases, determining whether the identified bacterial signals likely represent contaminants or whether they represent bacteria that may form associations with the fungi remains a challenge.

To help address the problem of differentiating signals of true bacterial associates from likely contaminants, we have also independently used 16S amplicon screens of more than 200 fungal isolates from multiple fungal culture collections for associated bacteria to look for concordance between the results of these screens and the results of the bioinformatics analysis. In addition, we have used the results from screening the fungal cultures in order to guide our bioinformatics screening, finding a clear signal of *Massilia timonae* in the sequencing reads from an isolate of *Aspergillus glaucus*.

By combining DNA sequence data from both fungal collections and publicly available fungal sequencing data, we have identified strong signals of associated bacteria in some fungal isolates. We will continue tuning this analysis pipeline and will also apply it to a wider variety of metagenome samples, including complex soil metagenomes, to identify potential bacterial associates of fungi from a range of environments. This information can then be used to target specific fungi for isolation from environmental samples and to determine how a range of these bacterial:fungal interactions affect both fungal and bacterial phenotypes.

This SFA is supported under the Computational Biosciences Program of the Office of Biological and Environmental Research in the DOE Office of Science through grant DE-AC52-06NA25396

LA-UR-19-20509

Democratization of fungal highway columns as a tool to investigate bacteria associated with soil fungi

Pilar Junier[1]* (pilar.junier@unine.ch), Olivia C. Trautschold[2], Nicholas G. Parra-Vasquez[2], Saskia Bindschedler[1], Ilona Palmieri[1], Fabio Palmieri[1], Aislinn Estoppey[1], Danae Bregnard[1], Wafa M. Kooli[1], Andrea Lohberger[1], Geoffrey L. House[3], La Verne Gallegos-Graves[3], Julia M. Kelliher[3], Armand E. K. Dichosa[3], Hang N. Nguyen[4], Simone Lupini[4], Jean F. Challacombe[5], Jamey D. Young[6], Debora F. Rodrigues[4], and **Patrick S. G. Chain[3]**

[1]Institute of Biology, University of Neuchâtel, Neuchâtel Switzerland; [2]Materials Science and Technology, Los Alamos National Laboratory, Los Alamos New Mexico; Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico; [4]Civil and Environmental Engineering, University of Houston, Houston, Texas; [5]College of Agricultural Sciences, Colorado State University, Fort Collins, Colorado; [6]School of Engineering Vanderbilt University, Nashville, Tennessee.

https://genomicscience.energy.gov/research/sfas/lanlbfi.shtml

Project Goals: Bacteria-fungi interactions (BFI) are recognized as essential for numerous environmental processes. These interactions are modulated not only by the nutritional conditions of the surrounding environment, but also by its physicochemical constrains. In soils, one of these physical constrains relates to the unsaturated nature of the substrate and the consequence of this for the dispersal and activity of bacteria. Developing new tools to investigate how bacteria overcome dispersal limitation at a cellular scale is important to unravel the mechanisms by which BFI affect processes at an ecosystem scale. This is not only relevant for soils, but also for other unsaturated matrices in which microbial activity is an important driver of function. In many unsaturated systems, bacteria and fungi engage in a cooperative interaction in which bacteria use fungal hyphae to disperse. This interaction is referred to as fungal highways. Here we present the design and validation of a sampling system obtained by additive printing to enrich interacting bacterial-fungal couples that form these fungal highways.

Abstract

Fungi and bacteria form different types of associations that are central to numerous environmental processes. Although many bacteria are able to move by different mechanisms, occasionally requiring the production of specific appendages, most bacterial motility types are only efficient in the presence of a liquid film (1). Thus, in soil and other water-unsaturated matrices, bacterial dispersal is highly limited (2). In contrast, the filamentous growth of fungi is an ideal morphology for colonization of unsaturated matrices. Some bacteria can use the liquid film surrounding hyphae for their dispersal in a mechanism known as fungal highways (3). In order to investigate this dispersal mechanism, a sampling device was designed to select and isolate both bacteria and their fungal highway partner directly from an environmental sample by including an attracting and a target medium in a unsaturated column. These devices, called fungal highways columns, have been used to investigate the diversity and function of fungal highways in soils (4). However, one major limitation for the routine and standardized use of such devices as a scientific instrument for investigating fungal highways in soils has been the variability and time-consuming manual production of these devices by experienced scientists. In this study we present the results of the design and testing of a fungal highway column generated by additive 3D printing. This new model

considers the main features of the manually built device, namely the possibility to include any type of targeting and attracting media, an hourglass shape to avoid the formation of a continuous water film, and a built-in modular lattice that simulates the presence of soil particles and air-filled gaps.

References

(1) Harshey RM. Bacterial motility on a surface: many ways to a common goal. Annu Rev Microbiol 2003 57 249 73.
(2) Or D, Smets BF, Wraith JM, et al. Physical constraints affecting bacterial habitats and activity in unsaturated porous media – a review. Adv Water Res 2007 30 1505 27.
(3) Kohlmeier S Smits TH Ford R et al. Taking the fungal highway: mobilization of pollutant-degrading bacteria by fungi. Environ Sci Technol 2005 39 4640 6
(4) Simon A, Binschedler S, Job D et al. Exploiting the fungal highway: development of a novel tool for the in situ isolation of bacteria migrating along fungal mycelium. FEMS Microbio Ecol 2015 fiv116.

This study was supported by a U.S. Department of Energy Biological and Environmental Research Science Focus Area grant (grant no. DE-AC52-06NA25396).

Using Gel Microdroplets to Investigate Bacterial Influence on Fungal Spore Germination

Julia M. Kelliher[1]* (jkelliher@lanl.gov), Geoffrey L. House[1], Laverne Gallegos-Graves[1], Demosthenes P. Morales[1], Savannah A. Martinez[1], Armand E. K. Dichosa[1], Debora F. Rodrigues[3], Hang N. Nguyen[3], Saskia Bindschedler[2], Jean F. Challacombe[4], Jamey D. Young[5], Pilar Junier[2], and **Patrick S. G. Chain[1]**

[1]Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico; [2]Institute of Biology, University of Neuchâtel, Neuchâtel Switzerland; [3]Civil and Environmental Engineering, University of Houston, Houston, Texas; [4]College of Agricultural Sciences, Colorado State University, Fort Collins, Colorado; [5]School of Engineering Vanderbilt University, Nashville, Tennessee

https://genomicscience.energy.gov/research/sfas/lanlbfi.shtml

Project Goals: The vast taxonomic diversity and the complexity of interactions within the soil microbiome present a unique challenge. Many of the interactions between soil-dwelling bacteria and fungi are not yet well understood, and a more comprehensive understanding would lead to substantial agricultural, environmental, and energy-focused advancements. These potential developments align with the focus areas of the DOE, and would influence multiple scientific fields. The aim of this Science Focus Area (SFA) is to better understand the diverse and abundant interactions within the soil rhizosphere, specifically between fungi and bacteria, and to decipher the mechanisms behind their communication. Herein, we discuss a pipeline for the effective interrogation and monitoring of bacterial:fungal partners that will provide insight into the environmental conditions and biotic associations that affect fungal spore germination.

An understanding of the interactions between microbes inhabiting the soil rhizosphere is becoming increasingly pertinent. Bacterial:fungal interactions are at the forefront of this undertaking, specifically regarding how bacteria can promote the growth of fungal species that are used for biocontrol, and how bacteria can inhibit the growth of pathogenic fungi. Elucidating which bacteria promote or inhibit fungal spore germination would have substantial implications in understanding impacts on plant fitness, agricultural practices, and more broadly, ecosystem-level nutrient flux. In order to address this challenge, our lab has developed a pipeline to rapidly assess bacterial:fungal partners of interest during the early stages of fungal growth.

Gel microdroplet (GMD) technology was previously optimized by Los Alamos National Laboratory for use in the screening of algae, bacteria, and various microbiomes. These microsphere agarose droplets range from 20-70μm in diameter and are used to encapsulate viable cells for cultivation and downstream molecular assays. For this project, we have employed GMDs to interrogate bacterial:fungal interactions by co-capturing bacterial cells with fungal spores. These GMDs allow us to test how fungal spore germination is affected by various ratios of bacteria to fungal spores while maintaining close proximity between the captured cells and allowing for nutrient exchange. In addition, the GMDs are able to be efficiently interrogated and sorted via flow cytometry and visualized using microscopy; and finally, the GMDs have demonstrated improved genomic recovery for downstream analyses compared to traditional culturing techniques.

We have successfully co-captured several genera of fungal spores and bacteria within GMDs, and monitored their growth by flow cytometry and microscopy. We optimized the

targeted flow cytometry-based enrichment of the GMDs of interest in order to achieve an 80% success rate in isolating only the GMDs containing the desired bacterial:fungal ratios. The sorted GMDs can now be monitored either in bulk or individually on novel static culture microscopy slides. Initial experiments established that the GMDs are best suited for monitoring the early stages of spore germination, as the fungal hyphae can easily extend out of the droplets when they become too long. The ease, utility, and speed of the GMD screening process makes this pipeline an ideal candidate for observing altered fungal germination rates in the presence of various bacterial partners. This technology will be applied to fungal and bacterial partners of interest, and will be coupled with time-lapse transcriptomics to better understand the mechanisms through which bacteria can affect fungal spore germination.

This SFA is supported under the Computational Biosciences Program of the Office of Biological and Environmental Research in the DOE Office of Science through grant DE-AC52-06NA25396

LA-UR 19-20592

Taking a Closer Look at Bacterial & Fungal Interactions in Soil Using Imaging

Demosthenes Morales, III,[1,2*] (dmorales@lanl.gov), Geoffrey House,[2] Julia Kelliher,[2] Jim Werner,[1] and **Patrick Chain**[2]

[1]Center of Integrated Nanotechnologies, Los Alamos National Laboratory; Los Alamos, NM
[2]Biosecurity and Public Health, Los Alamos National Laboratory, Los Alamos, NM

https://genomicscience.energy.gov/research/sfas/lanlbfi.shtml

Project Goals: This project intends to inform the use of soil microbiomes to address DOE priorities in overcoming energy and environmental challenges. We are focusing on understanding the role of bacterial:fungal interactions in ecosystem development by connecting microbial diversity to actionable phenotypic responses. To do so, genomic and metagenomic sequencing of soil microbes will be combined with advanced imaging techniques and metabolomics to determine a mechanistic route in which these organisms associate to augment soil fertility and plant growth.

There is a great complexity of conditions affecting soil that have no absolute value, making it difficult to generate models predicting states for optimal plant growth. One such factor is the soil microbiome which has long been recognized to be important for nutrient solubilization and transfer required for plant development, but difficulties in characterizing microbial composition and function has made studying their influence quite elusive. Recent ease in accessibility to whole genome sequencing and advances in multi-omics analyses has improved the ability to quantitate plant and soil microbe relationships. Here we will take a quantitative look at how bacteria directly alter the genetic profile of a single fungal cell in a spatial dependent manner as a result of bacterial-fungal interactions using single molecule fluorescence in situ hybridization (smFISH). By determining how the intimate relationships between bacteria and fungi influence soil conditions we might better predict how microbial structure ultimately affects the quality of plant life.

This SFA is supported under the Computational Biosciences Program of the Office of Biological and Environmental Research in the DOE Office of Science.

Bacterial Oxalotrophy as a Biocontrol Mechanism against Fungal Pathogens

Aislinn Estoppey[1], Fabio Palmieri[1]* (fabio.palmieri@unine.ch), Geoffrey L. House[2], Debora F. Rodrigues[3], Jamey D. Young[4], Saskia Bindschedler[1], Pilar Junier[1], **Patrick S. G. Chain**[2].

[1]Laboratory of Microbiology, Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland; [2]Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico; [3]Civil and Environmental Engineering, University of Houston, Houston, Texas; [4]School of Engineering Vanderbilt University, Nashville, Tennessee

https://genomicscience.energy.gov/research/sfas/lanlbfi.shtml

Project Goals: The worrisome rise of antifungal drugs resistance leads to the need of finding alternative therapeutic strategies to fight against fungal pathogens. Here we propose to apply bacterial nutritional interference as an alternative biocontrol strategy in order to control the growth of these pathogenic fungi.

The worldwide emergence of drug-resistant pathogenic fungi in the last decades is a major problem not only in human health, but also in agriculture and food security. Nowadays, only a very limited number of therapeutic alternatives are available for the treatment of fungal infections. Generally, the same compounds are used in both agriculture and medicine, creating an interconnected system favoring the spread of resistance. This is particularly worrisome for pathogens with a broad range of hosts. Therefore, there is an urgent need to find more sustainable therapeutic approaches to mitigate the rise of antifungal drug resistances.

Some phytopathogenic fungi such as *Sclerotinia sclerotiorum* or *Botrytis cinerea* are known to use oxalic acid as a pathogenicity factor. Indeed, oxalic acid chelates calcium ions in the middle lamella of the plant cell wall in order to access and degrade pectin. Moreover, oxalic acid often precipitates as calcium oxalate (CaOx) crystals. These crystals mechanically break the plant cell wall, facilitating penetration and infection of the plant tissues. Oxalic acid is one of the most common low molecular weight organic acids (LMWOA) produced by fungi and it plays multiple roles aside pathogenicity in processes such as interspecies competition, mineral weathering, and lignocellulose degradation. In soil, it is primarily present in the form of CaOx crystals. This pool of CaOx can be consumed by oxalotrophic bacteria, leading to a local pH increase. The effect of bacteria on CaOx and pH are the stepping stones of this biocontrol project. In the case of normal infection, the secretion of oxalic acid by the fungal pathogen leads to the acidification of the host tissues and makes the environment favorable for infection. However, when oxalotrophic bacteria are added to the system, the bacterial consumption of oxalic acid leads to a pH increase and makes the environment less favorable for infection.

Confrontation experiments in Petri dishes were conducted with different phytopathogenic fungi (*Botrytis cinerea* and *Sclerotinia sclerotiorum*) and oxalotrophic bacteria (*Cupriavidus necator, Cupriavidus oxalaticus*, and *Burkholderia phytofirmans*) in different nutritional conditions. In addition, a non-oxalotrophic bacterium (*Pseudomonas putida*) and a Δ*oxc*-mutant of *B. phytofirmans* were considered as controls. *C. necator* and *C. oxalaticus* inhibited the growth of both *B. cinerea* and *S. sclerotiorum*. This was also the case for *B. phytofirmans*, with a pronounced growth inhibition in confrontation with *S.*

sclerotiorum. The Δ*oxc*-mutant of *B. phytofirmans* still inhibited the growth of *S. sclerotiorum* but to a lesser extent. Moreover, oxalic acid concentration was quantified in liquid mono- and co-cultures. Plant-bacterium-fungus interactions are proposed to be investigated in the future with the plant model *Lactuca sativa* (lettuce).

This study was supported by a U.S. Department of Energy Biological and Environmental Research Science Focus Area grant (grant no. DE-AC52-06NA25396).

Diversity of endo- and exo-bacteria associated with soil fungi

Debora F. Rodrigues[1]* (dfrigirodrigues@uh.edu), Hang N. Nguyen[1], Simone Lupini[1], Geoffrey L. House[2], La Verne Gallegos-Graves[2], Julia M. Kelliher[2], Armand E. K. Dichosa[2], Jean F. Challacombe[2], Jamey D. Young[4], Fabio Palmieri[3], Saskia Bindschedler[3], Andrea Lohberger[3], Pilar Junier[3], and **Patrick S. G. Chain[2]**

[1]Civil and Environmental Engineering, University of Houston, Houston, Texas; [2]Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico; [3]Institute of Biology, University of Neuchâtel, Neuchâtel Switzerland; [4]School of Engineering Vanderbilt University, Nashville, Tennessee

https://genomicscience.energy.gov/research/sfas/lanlbfi.shtml

Project Goals: Interactions between bacteria and fungi are important determinants of ecosystem function, yet little is known about these interactions or how they operate. This is a critical knowledge gap as these interactions are important in addressing multiple DOE priorities including developing renewable energy sources, understanding the possible effects of Earth system change, and understanding how these interactions may help overcome energy and environmental challenges. Here we present preliminary results of the diversity and function of bacterial:fungal interactions in soil ecosystem. Using a combination of single cell isolation and cultivation techniques, as well as bioinformatics-based data mining of 16S rRNA and ITS sequences, we are beginning to understand the diversity of bacteria that form associations with fungi, and how these associations affect both fungal and bacterial growth.

Abstract

Fungi are cosmopolitan microorganisms with complex genetic make-up and metabolism.[1, 2] Furthermore, this group of microorganisms possesses important roles in ecology, agriculture, forestry and human health. In soil, fungi are one of the most abundant group of microorganisms and are known to interact with different microorganisms, including bacteria. Bacterial:fungal interactions in soils can be positive or negative (synergistic or antagonistic). In the present study, we obtained 45 fungal isolates and 53 exobacteria associated with the fungal isolates from soil microcosms from six different locations. These fungi were obtained with different growth media of plant origin, namely cornmeal, oatmeal, sorghum grain, and potato carrot. The fungal isolates were also investigated via real-time PCR for endobacteria. Out of the 45 fungal isolates, 41 were associated with endobacteria. From this culture collection, four fungal isolates (*Didymella* spp., *Neopestalotiopsis* spp., *Staphylotrichum coccosporum*, *Aspergillus* spp.) and four bacterial isolates (*Exiguobacterium sp.,* two *Paenibacillus spp.*, and *Pseudomonas sp.*) associated with these fungi with synergistic and antagonistic interactions from the same microcosm were selected for further investigation. The selection was based on confrontational assays. These confrontational assays allowed us to distinguish bacteria that would inhibit or enhance fungal growth. In addition to the exobacteria interaction with fungi, another five fungi were selected to further investigate for the endo-bacteria association based on the real time PCR results. The fungi selected were *Didymella glomerata*, *Cladosporium sp.*, *Aspergillus fumigatus*, *Aspergillus sp.* and *Byssochlamys spectabilis*. Hydrophobicity analyses via contact angle and surface composition via Attenuated total reflection Fourier transform infrared (ATR-FTIR) were also employed to better understand the physical interactions of the bacteria with these fungi. All nine fungal isolates presented cell

surfaces with moderate to high hydrophobicity. The main functional groups observed on the surface of these fungi belonged to polysaccharides as well as phosphate compounds and proteins. Further investigations of the metabolites involved in the communication between the fungi and the bacteria triggering negative and positive interactions are underway.

References

[1] L. de Aguirre, S.F. Hurst, J.S. Choi, J.H. Shin, H.P. Hinrikson, C.J. Morrison, Rapid Differentiation of Aspergillus Species from Other Medically Important Opportunistic Molds and Yeasts by PCR-Enzyme Immunoassay, Journal of Clinical Microbiology, 42 (2004) 3495-3504.
[2] R.E. McKinney, Environmental pollution control microbiology: a fifty-year perspective, CRC Press, 2004.

This study was supported by a U.S. Department of Energy Biological and Environmental Research Science Focus Area grant (grant no. DE-AC52-06NA25396).

Parallelized *in vivo* Construction of a Synthetic 57-Codon *E. coli* Genome

Nili Ostrov,[1] Maximilien Baas-Thomas[2]* (maximilien@g.harvard.edu), Alexandra Rudolph,[2] Matthieu Landon,[3] and **George Church**[1,4]

[1]Department of Genetics, Harvard Medical School, Boston, MA; [2]Program in Biological and Biomedical Sciences, Harvard University, Cambridge, MA; [3]Program in Systems Biology, Harvard University, Cambridge, MA; and [4]Wyss Institute for Biologically Inspired Engineering, Boston, MA

http://arep.med.harvard.edu

Project Goals: We are building a fully recoded, 3.97 Mb *Escherichia coli* genome, in which seven codons are replaced with synonymous alternatives in all protein coding genes. For this aim, the recoded genome was *de-novo* synthesized and assembled in vivo into 87 segments. In the final steps of genome construction, we combine all segments to a single fully recoded genome.

We present progress towards a fully recoded, 57-codon *Escherichia coli* genome, in which seven codons are replaced with synonymous alternatives in all protein coding genes. Synthetic recoded DNA was assembled into 50 kb episomal segments, which were individually tested and combined *in vivo* into larger recoded "clusters" via Cas9-assisted recombineering. Each segment replaces its corresponding wild-type sequence. This integration efficiency for 50 kb segments was enhanced from 30% to 80% by increasing homology size and optimization of gRNA sequences. Cycle time was further decreased by streamlining of screening and sequence analysis. Currently, we are simultaneously constructing eight recoded clusters, aiming for 500 kb of recoded DNA in each. Thus far, no significant decrease in growth rate has been observed in intermediate strains carrying up to 200 kb of genomically recoded DNA (~200 recoded genes). In parallel, we are actively pursuing de-risking of the final assembly protocol by conjugative assembly (CAGE) of recoded clusters. As we approach the final assembly steps of a virus-resistant *E.coli* genome, intermediate clusters are also used to test expression of genes which depend on non-standard amino acids, to enable strict biocontainment of the final strain. Our work expands the toolkit available for large scale engineering in living cells and opens new avenues for rewriting genomes.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under Award Number DE-FG02-02ER63445 and Defense Advanced Research Projects Agency grant BAA-12-64. MML and GMC are inventors on patent application #62350468 submitted by President and Fellows of Harvard College that covers Methods for Rule-based Genome Design. Dr. Church is a founder of the following companies in which he has related financial interests: ReadCoor; EnEvolv; and 64-x. For a complete list of Dr. Church's financial interests, see also http://arep.med.harvard.edu/gmc/tech.html.

Establishing and optimizing RecT-mediated homologous recombination in bacteria beyond *E. coli*

Gabriel T. Filsinger[1,9]* (filsinger@g.harvard.edu), Timothy M. Wannier[2,9,†], Felix B. Pedersen[3,†], Isaac D. Lutz[4,†], Julie Zhang[5,†], Devon A. Stork[6,9], Kevin Gozzi[7], Helene Kuchwara[2], Verena Volf[8,9], Stan Wang[2,9], Michael T. Laub[7], **George M. Church[2, 9]**.

[1]Department of Systems Biology, Harvard Medical School, Boston, Massachusetts, USA.
[2]Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA.
[3]Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense M, Denmark.
[4]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
[5]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
[6]Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts, USA.
[7]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
[8]Harvard University John A. Paulson School of Engineering and Applied Sciences, Cambridge, Massachusetts, USA.
[9]Wyss Institute for Biologically Inspired Engineering, Harvard University, Cambridge, Massachusetts, USA.
† These authors contributed equally

http://arep.med.harvard.edu

Project Goals: The Lambda phage encoded Red recombination system (λ-Red) is among the most widely used genome editing tools in *E. coli* due to its ability to perform scarless mutagenesis using ssDNA oligonucleotides (oligos), gene knockouts and integrations, and multiplexed editing. In many bacteria, alternative methods of genome editing have significant limitations. Site-specific nucleases alone are lethal, site-specific recombinases leave genomic scars and cannot be flexibly programmed, base editors are imprecise, and native homologous recombination proceeds through inefficient random crossover. λ-Red is species-specific however, and does not function in most bacteria beyond *E. coli*, severely limiting the applicability of this technology. Here we aim to provide a framework for establishing and improving oligo-mediated editing in bacteria beyond *E. coli*.

The Lambda phage RecT-family protein Beta improves homologous recombination and facilitates multiplexed genome editing in *E. coli*. However this technology is host-restricted and has low activity outside *E. coli*. We find that host bacterial single-stranded binding proteins (SSB) affect the species-restriction of RecT proteins, and through co-expression we improve activity in foreign hosts by up to 3 orders of magnitude. We identify the SSB C-terminal 7 amino acids as a major recognition domain, and reprogram RecT-SSB compatibility by swapping this domain. We demonstrate the utility of this approach by establishing oligo editing for the first time in *C. crescentus*. Then in *L. lactis*, we explore other variables limiting editing efficiency

including mismatch repair and protein expression, improving rates of single-nucleotide mutations 500-fold to 23.8% at one site, and performing gene integrations. We optimized editing to generate millions of combinatorial variants within a ribosomal antibiotic target in *L. lactis*, finding that the majority of resistant mutants cannot be predicted by resistant single mutants alone.

This work was supported through National Institute of General Medical Sciences under Grant No. 1U01GM110714-01 and the Department of Energy DE-FG02-02ER63445 to G.M.C.. This work was also supported by NIH grant R01GM082899 to M.T.L. who is an Investigator of the Howard Hughes Medical Institute. G.T.F. was supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE1745303. D.A.S was supported by a Landry Cancer Biology Research Fellowship. K.G. was supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE1745302. Dr. Church is a founder of the following companies in which he has related financial interests: ReadCoor; EnEvolv; and 64-x. For a complete list of Dr. Church's financial interests, see also http://arep.med.harvard.edu/gmc/tech.html.

Deciphering Microbial Structures *in situ* Using Subcellular RNA Sequencing.

Andrew C. Pawlowski[1,2]*(Andrew_Pawlowski@hms.harvard.edu), Richie E. Kohman[2], Jenny M. Tam[1,2], John Aach[1], **George M. Church**[1,2]

[1]Harvard Medical School, Boston; [2]Wyss Institute for Biologically Inspired Engineering, Boston, Massachusetts

http://arep.med.harvard.edu/

Project Goals:

This study aims to develop methods for characterizing microbial communities with single-cell and spatial resolution. We are employing fluorescent *in situ* sequencing (FISSEQ) technology for acquiring genomic and transcriptomic information to better understand the spatial arrangement of microbes and their gene expression in natural and synthetic microbial communities.

Most investigations reduce microbial physiology to monoculture conditions, which does not consider their abundant interactions in natural environments. Indeed, much of our understanding of microbes stems from gene deletions, heterologous expression, and *in vitro* enzyme characterization. Microbes need to be studied *in situ* where their spatial organization holds biological importance for responding to the environment (*e.g.*, polymicrobial metabolism, biofilms, horizontal gene transfer). However, *in situ* characterizations of these complex phenotypes in natural environments are generally not feasible and remain largely unknown.

Microbial cell walls and membranes are notorious for their low permeability to all molecules, such as oligonucleotides and enzymes needed for subcellular sequencing. We envision that methods for removing the cell wall and cell membrane will enable spatial characterization of a wide-range of microbes that are normally recalcitrant to single-cell techniques (e.g., fluorescent *in situ* hybridization (FISH), subcellular RNA sequencing). The Church Lab in collaboration with Ed Boyden's lab at MIT have developed Expansion FISH and subcellular Expansion Sequencing (ExSeq), based on the Expansion Microscopy (ExM) technique. ExM is a super-resolution imaging technique whereby biological samples are embedded in an acrylamide/acrylate hydrogel and physically expanded (in ExSeq; ~3.5x linear expansion, ~40x volumetric expansion).

Here, we describe our progress towards subcellular sequencing of microbes using FISSEQ paired with ExM. We have found conditions for performing ExM on *E. coli* and *Saccharomyces cerevisiae* and find that cell wall degradation is required for the successful expansion of cellular components. Further, we have developed a series of chemical treatments that covalently attach total cellular nucleic acids to the ExM hydrogel, which is a pre-requisite for generating sequencing libraries while maintaining their original spatial location. Notably, we establish protocols for generating *in situ* sequencing libraries that enable multiplexed assays for identifying microbial taxonomy and gene expression, and demonstrate preliminary results for whole-transcriptome sequencing.

This project has been funded by DOE grant DE-FG02-02ER63445. Dr. Church is a founder of companies in which he has related financial interests: ReadCoor; EnEvolv; and 64-x. For a complete list of Dr. Church's financial interests, see also arep.med.harvard.edu/gmc/tech.html.

Discovery and characterization of disease resistance loci using a unique gene copy number variant population

Héloïse Bastiaanse[1,2*], Maria Marlin[4], Isabelle Henry[3], George Newcombe[4], Randall Rousseau[5], **Luca Comai**[2,3], Andrew Groover[1,2]

[1] US Forest Service, Pacific Southwest Research Station, Davis CA; [2] UC Davis, Department of Plant Biology; [3] UC Davis, Genome Center; [4] University of Idaho, Department of Forest, Rangeland and Fire Science; [5] Mississippi State University, Department of Forestry

bastiaanse.h@gmail.com

Abstract

Pathogenic fungi that colonize poplar leaves and stems reduce yield and can cause failure of industrial bioenergy plantations. Despite extensive study of poplar pathosystems, the genetic control of poplar resistance to pathogens is still poorly understood, underscoring the need for new approaches. We developed a unique functional genomics resource based on gene dosage variation in an elite biomass poplar hybrid. We pollinated *Populus deltoides* with gamma irradiated *P. nigra* pollen to produce ~ 800 F1 seedlings. These contain large-scale deletions and insertions that together tile each chromosome multiple times. Under natural infection in the field, as well as under controlled inoculations in the greenhouse, we observed a wide variation in disease resistance within our population and were able to identify dosage QTLs influencing resistance of poplar to some of its most important fungal diseases: leaf rust and leaf spot (*Melampsora* sp., *Alternaria* sp.). Next, time-course analysis of gene expression during progression of disease symptoms will be performed for selected genotypes, and used to develop predictive models of transcriptional networks underlying disease susceptibility. A final set of experiments will aim to identify candidate genes for functional analysis by manipulation using CRISPR-Cas9. Such dosage-sensitive candidate genes with significant effects on disease resistance phenotypes could be exploited in breeding programs through the selection of germplasm with naturally-occurring allelic variation or indels/copy number variation covering resistance loci.

***EvoNet*: A Phylogenomic and Systems Biology approach to identify genes underlying plant survival in marginal, low-Nitrogen soils**

Gil Eshel, [1*] (ge30@nyu.edu), Kranthi Varala,[2] Rodrigo Gutiérrez,[3] Dennis Stevenson,[4] Robert DeSalle,[5] W. Richard McCombie,[6] Jean-Michel Ané,[7] Heidi Kaeppler[7] and **Gloria Coruzzi**[1]

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

http://evonet.org

Project Goals: This DOE BER sustainability project aims to identify the key genes and gene regulatory networks that enable "extreme survivor" plants to adapt and grow in marginal, extremely nitrogen (N) -poor soils in the hyperarid Atacama Desert in Chile. These "extreme survivor" species cover the main branches in flowering plants, and include 7 grass species of particular interest for biofuels. We focus on 28 "extreme survivor" Atacama species and compare their genomes to Californian "sister" species that live in a N-replete conditions in arid (27 species) or mesic (27 species) environments. Deep RNA-sequencing of these "triplet species" was used to fuel a novel phylogenomic analysis that helps identify individual genes that support the evolutionary divergence of the extreme survivors in Atacama Desert from their sister species in California. The genes thus identified will help to discover the mechanisms underlying physiological and developmental processes that allow plant survival in nitrogen-poor, dry soils. The genes and network modules so uncovered can potentially be translated to biofuel crops to greatly increase biomass and nitrogen use efficiency in marginal, low-fertility soils.

This collaborative project exploits the genomes of "extreme survivor" plants adapted to thrive in marginal, extremely Nitrogen (N) poor soils in the hyperarid Atacama Desert in the Chilean Andes. It uses a previously validated phylogenomic pipeline we developed called PhyloGeneious [1], which can identify genes that provide positive support to species divergence. By applying this phylogenomic pipeline to the gene sequences of these "triplet species", we can identify the genes that distinguish these "extreme survivors" in Atacama from their related "sister" species adapted to similarly dry regions in California (CA) not constrained by N and/or water availability and to mesic "sister" species growing in N and water (W) replete conditions. These "extreme survivor" species from the Atacama desert broadly cover the main branches in flowering plants, and therefore offer a wide range of genomic backgrounds within which the survival traits repeatedly arose i.e., multiple independent origins of trait.

Key to our phylogenomic approach is the "triplet species" sampling strategy. To maximize our ability to separate the trait-relevant signature from overall speciation events, our "triplet species" sampling will cover multiple independent origins of the low-N adaptive trait. In published studies, we showed that our phylogenomic pipeline could; i) identify genes that underlie convergent evolution of antioxidant synthesis in Rosids in a study of 150 plant genomes [1]; and ii) identify 100+ genes associated with the loss of Arbuscular Mycorrhizal (AM) symbiosis in the *Brassicaceae* [2]. We now extend this phylogenomic approach to the study of "extreme survivor strategies" as follows:

Aim 1. **Species collection and deep transcriptome sequencing**: (NYU, NYBG, Chile). ***Progress:*** We sequenced all 28 species collected in the Atacama Desert and 39 of their CA relative "sister" species (Table 1), with resulting average gene coverage of 87%, based on the BUSCO conserved single-copy orthologs existence assessment.

Aim 2. **Phylogenomic Analysis**: Perform phylogenomic analysis of 82 "triplet species" to identify genes that repeatedly support nodes that distinguish the extreme survivors in the Atacama Desert from their sister species in CA (AMNH, NYU). ***Progress:*** We performed phylogenomic analysis that includes each of the major plant lineages (*Poaceae* - 13 taxa; *Caryophyllales* – 9 taxa; *Lamiids* – 12 taxa; *Campanulids* – 25 taxa; *Fabaceae* – 11 taxa). This analysis identified hundreds of genes that provide phylogenetic support to the splits between the "extreme survivors" species that thrive in the Atacama Desert and their closest relatives from CA that grow in drought-adapted or mesic environments. We are further analyzing these gene sets to narrow down the best candidates for validation studies in a model grass.

Aim 3. **Network Analysis**: To combine phylogenomics (encoded protein sequence) and gene expression data to identify genes and network modules associated with adaptations to marginal, low-N soils (NYU, Chile). ***Progress:*** To exploit a comparative analysis of gene regulatory networks, we are currently developing a new module called **PhyloExpress** that extends the PhyloGeneious pipeline to include gene expression data. As a test data set, we are reconstructing, in parallel from sequence as well as gene expression data, the phylogeny of five diverse plant species (2 *Poaceae* - Maize, Rice, 2 *Rosids* - Soybean, Arabidopsis and 1 *Asterid* – Tomato).

Aim 4. **Functional Validation**: To functionally validate top-ranked candidate genes for low-N adaptation in Arabidopsis and Brachypodium (NYU, Chile, U Wisconsin). ***Progress:*** We have begun to transform Brachypodium with the most promising candidates from our preliminary analysis using our Atacama set and their closest sequence available sister species.

Table 1. Extreme survivor species in Atacama Desert, Chile, and their "sister" species in CA (Drought or Mesic). *Our project studies 28 triplet of species from Marginal (Dry +low-N, Atacama), Dry (Dry, California), and moist (Mesic, California) soils. All the Atacama species and 39/54 of the California species have already been sequenced (**RED**) while 15/54 California species are being collected (**BLACK**).*

Atacama Desert, Chile Extreme Survivor Species (Drought + low-N)	California "Sister" Species (Drought)	California "Sister" Species (Mesic)
Atriplex imbricata	*Atriplex lentiformis*	*Atriplex watsonii*
Mulinum crassifolium	*Sanicula crassicaulis*	*Conium maculatum*
Ambrosia artemisioides	*Ambrosia chamissonis*	*Ambrosia psilostachya*
Baccharis boliviensis and Baccharis tola	*Baccharis glutinosa*	*Baccharis salicifolia*
Trichocline caulescens	*Cirsium occidentale*	*Cirsium fontinale*
Chuquiraga atacamensis	*Coreopsis douglasii*	*Rudbeckia californica*
Parastrephia quadrangularis	*Grindelia hirsutula*	*Eriophyllum confertiflorum*
Senecio puchii	*Senecio californicus*	*Senecio mikanioides*
Tagetes multiflora	*Pectis papposa*	*Pluchea sericea*
Phacelia pinnatifida	*Eriodictyon tomentosum*	*Phacelia nemoralis*
Pycnophyllum bryoides	*Cerastium viride*	*Cerastium beeringianum*
Adesmia spinosissima	*Amorpha californica*	*Amorpha fruticosa*
Lupinus oreophilus	*Lupinus nana*	*Lupinus arboreus*
Lupinus subinflatus	*Lupinus hirsutissimus*	*Lupinus latifolius*
Aristida adscensionis	*Danthonia unispicata*	*Danthonia californica*
Bouteloua simplex	*Bouteloua curtipendula*	*Muhlenbergia filiformis*
Calamagrostis crispa	*Calamagrostis rubescens*	*Calamagrostis breweri*
Calamagrostis cabrerae	*Festuca californica*	*Festuca subuliflora*
Munroa decumbens	*Munroa squarrosa*	*Munroa utilis*
Nassella nardoides	*Nassella cernua*	*Nassella manicata*
Jarava frigida	*Stipa coronata*	*Stipa kingii*
Chorizanthe conmisuralis	*Chorizanthe palmeri*	*Rumex crispus*
Exodeconus integrifolius	*Lycium cooperi*	*Physalis lancifolia*
Fabiana denudata	*Nicotiana glauca*	*Petunia parviflora*
Solanum chilense	*Datura wrightii*	*Solanum douglasii*
Acantholippia deserticola	*Aloysia wrightii*	*Phyla nodiflora*
Fagonia chilensis	*Fagonia laevis*	*Tribulus terrestris*

References

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

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

Anthracnose resistant response of the USDA-NPGS sweet sorghum [*Sorghum bicolor* (L.)] germplasm collection

Hugo E. Cuevas [1, *](hugo.cuevas@ars.usda.gov), Louis K. Prom[2], Joseph E. Knoll[3], Wilfred Vermerris[4]

[1]USDA-ARS Tropical Agriculture Research Station, Mayaguez, PR; [2]USDA-ARS Crop Germplasm Research Unit, College Station, TX; [3]USDA-ARS Crop Genetics and Breeding Research Unit, Tifton, GA; [4]University of Florida, Department of Microbiology & Cell Science, Gainesville, FL

Project Goals:

The sweet sorghum collection of the USDA-ARS National Plant Germplasm Systems (NPGS) is the primary source of genetic diversity for the development of new sweet sorghum varieties for biofuel production. Although it has been demonstrated that the collection contains valuable genes/alleles that can aid the commercial production of biofuels, few sweet sorghum varieties have been phenotyped for biofuel applications**.** Anthracnose is a disease caused by the fungal pathogen *Colletotrichum sublineolum* that currently limits large-scale commercial production of sorghum in the southeastern U.S. and Caribbean. **The identification of novel anthracnose resistance sources present in sweet sorghum germplasm will expedite the development of new resistant sweet sorghum cultivars and hybrids by avoiding time-consuming introgression breeding approaches with non-sweet sorghums serving as donor of the resistance alleles**.

We evaluated 1,801 sweet sorghum accessions from the NPGS in single replicated trial in Puerto Rico during 2015 and 2016. The analysis identified 654 accessions with Brix value >10.0º of which 171 accessions had values >15.0º. Based on Brix values and geographic origin a subset of 233 sweet sorghum accessions were selected for further genotypic and phenotypic characterization. This diversity panel is composed of accessions from 19 countries, including advanced breeding materials and it represents ~10-15% of the NPGS sweet sorghum collection.

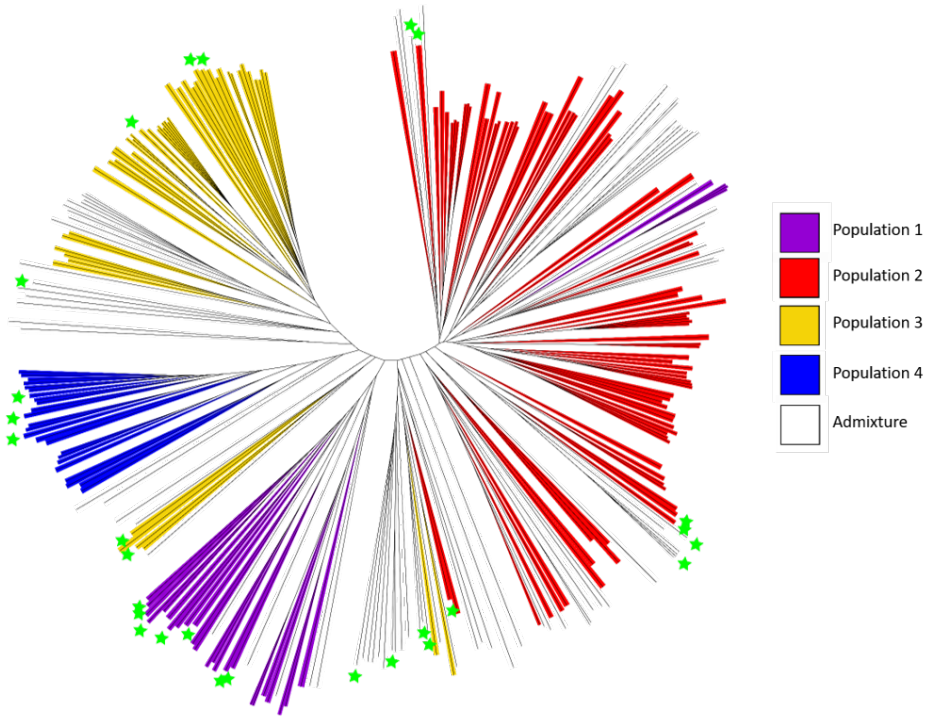
Genotype-by-sequencing analysis of the diversity panel identified 157,843 single-nucleotide polymorphism (SNPs). Population structure analysis based on 2,345 unlinked SNPs separated most of the accessions (~70%) of the panel into four populations and an admixture group (~30%) (**Figure 1**). The population structure could be related with breeding programs, since most of the advanced germplasm is highly genetically related and is distributed among three populations. The anthracnose resistance response of the diversity panel was determined based on replicated trial evaluations in Texas, Georgia, Florida and Puerto Rico during 2017 and 2018. The average of anthracnose resistance response in the diversity panel was less than 3.0 (on a 1 to 5 scale), indicating that most of the accessions were infected at the bottom leaves. A total 145 accessions showed different anthracnose resistance responses among different locations, while 29 and 59 were resistant and susceptible, respectively, at all four locations. The 29 resistant accessions were found at higher frequency in one population (**Table 1**).

An adequate knowledge of the genetic diversity in anthracnose-resistant sweet sorghum germplasm is necessary to effectively use the diversity in sorghum breeding programs. An unrooted neighbor-joining tree was generated to understand the genetic relationships among the 29 resistant accessions across locations (**Figure 1**). Anthracnose resistant accessions were present in each of the four populations. Nevertheless, the high genetic relationship among some resistant accessions suggests that the genes in these accessions are identical by descent (i.e., the accessions share the same resistance genes). Certainly, a limited number of anthracnose resistance genes are controlling the resistance response observed in the NPGS sweet sorghum collection.

Table 1. Frequency of the anthracnose resistant response in a subset of sweet sorghum germplasm from NPGS germplasm collection

Population	Resistant	Susceptible	X^2	Means ± S.D.
Population 1	8	19	0.00	2.2 ± 0.2
Population 2	4	69	0.08	3.0 ± 0.7
Population 3	5	30	0.60	2.4 ± 0.5
Population 4	3	15	0.45	2.8 ± 0.6
Admixture	8	63	0.72	2.7 ± 0.7

Figure 1. Unrooted neighbor-joining tree of a subset of sweet sorghum accessions from NPGS germplasm collection. Anthracnose resistant accessions are labeled with green starts.



Funding statement.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0014171.

Characterizing the transcriptomes of *Rhodoccocus opacus* strains adapted to model lignin breakdown products

Winston Anthony[1]* (winstoneanthony@wustl.edu), Rhiannon Carr[2], Tayte Campbell[1], William R. Henson[2], **Tae Seok Moon**[2], and **Gautam Dantas**[1,3,4,5]

[1] The Edison Family Center for Genome Sciences and Systems Biology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63110, USA; [2] Department of Energy, Environmental & Chemical Engineering, Washington University in St. Louis, St. Louis, MO, 63130, USA; [3] Department of Pathology and Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63108, USA; [4] Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO, 63130, USA; [5] Department of Molecular Microbiology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63108, USA

Project Goals: Using *R. opacus* strains adapted to model lignin breakdown products, we are identifying the genetic components and RNA expression changes which impart aromatic tolerance.

The oleaginous microbe *R. opacus* is naturally tolerant to aromatic compounds found in lignin-derived compound mixtures. *R. opacus* has been selected as a potential biological chassis for lignin conversion to lipids, and its mutants with improved aromatic tolerance have been generated by adaptive evolution. We have previously characterized the genetic alterations to *R. opacus* after aromatic adaptation. However, the differential expression profiles of mutant clones have only been described for individual compounds or one mixture of these compounds, limiting our understanding of aromatic tolerance and utilization in *R. opacus* [1, 2].

In this work, we used a multipartite study design to comprehensively characterize these mutants. We completed growth experiments, comparing wild-type *R. opacus* PD630 (WT) and the mutant strains in minimal media at a total aromatic concentration permissive to WT growth. For each compound and each compound mixture, we also measured the growth rate of adapted mutant strains grown at a high concentration not permissive to WT growth. Additionally, ^{13}C metabolite profiling and lipidomics have been performed along with transcriptomics. Importantly, we are now contextualizing the increased growth and compound utilization by comparing the transcriptomic profiles of the adapted strains grown on aromatic compounds at the permissive or nonpermissive concentrations. This study will deepen our understanding of aromatic tolerance and utilization mechanisms in diverse *R. opacus* mutants by expanding the list of aromatic compound mixtures in which these strains are fully characterized.

Publications

1. WR Henson, T Campbell, D DeLorenzo, Y Gao, B Berla, SJ Kim, M Foston, TS Moon and G Dantas. Multi-omic elucidation of aromatic catabolism in adaptively evolved *Rhodococcus opacus*. *Metab. Eng.* 49, 69–83 (2018)

2. A Yoneda, WR Henson, NK Goldner, KJ Park, KJ Forsberg, SJ Kim, MW Pesesky, M Foston, G Dantas and TS Moon. Comparative transcriptomics elucidates adaptive phenol tolerance and utilization in lipid-accumulating *Rhodococcus opacus* PD630. *Nucleic Acids Res.* 44, 2240–2254 (2016)

This work is funded by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research, Biological System Sciences Division, award # DE-SC0018324.

Multi-omic analysis of adapted *Rhodococcus opacus* strains for the characterization of aromatic metabolism

Tayte P. Campbell[1] (tayte.campbell@wustl.edu), Drew DeLorenzo[2], William R. Henson[2], Yu Gao[2], Bertram Berla[1], Soo Ji Kim[2], **Marcus Foston[2]**, **Tae Seok Moon[2]**, and **Gautam Dantas[1,3,4,5]**

[1]The Edison Family Center for Genome Sciences and Systems Biology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63110, USA; [2]Department of Energy, Environmental & Chemical Engineering, Washington University in St. Louis, St. Louis, MO, 63130, USA; [3] Department of Pathology and Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63108, USA; [4] Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO, 63130, USA; [5] Department of Molecular Microbiology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63108, USA

Project Goals: We aim to combine adaptive evolution and multi-omic approaches to identify aromatic tolerance mechanisms in *Rhodococcus opacus*. We aim to utilize identified genes and pathways to engineer an optimized *R. opacus* strain for the conversion of lignocellulose into valuable products.

Depolymerization of lignin results in a mixture of aromatic compounds that inhibit microbial conversion of lignocellulose into value-added fuels and chemicals. *R. opacus* is a promising strain for biofuel production due to its inherent tolerance for aromatic compounds. However, mechanisms for its aromatic tolerance and utilization remain undercharacterized. *R. opacus* has also been underutilized due to a lack of reliable engineering tools. To identify degradation pathways and tolerance mechanisms for aromatic mixtures in *R. opacus*, we performed adaptive evolution on 32 combinations of lignin model compounds, and used multi-omic approaches, including genomics, transcriptomics, targeted metabolomics, and lipidomics. Additionally, we developed a toolbox for reliable gene expression control and genome modification in *R. opacus*, and applied the developed tools to confirm our findings from these multi-omic analyses.

The key findings are as follows. First, adapted strains showed improved growth compared to wild-type on model lignin compound mixtures (up to 1900% growth improvement). Second, whole genome sequencing identified genes involved in redox reactions that were mutated in four distinct phenolic adaptation experiments with no mutations in two control adaption experiments. One of the mutated enzymes, superoxide dismutase, showed decreased activity in PVHG6, the adapted strain with the most improved growth, suggesting that redox reactions are important for aromatic tolerance and utilization. Third, using transcriptomics, we identified multiple aromatic transporters, five aromatic funneling pathways, and two degradation routes of the β-ketoadipate pathway for five aromatic compounds. Fourth, synthetic biology tools, including a reliable system for genome modification and the first reported CRISPR interference (CRISPRi) system for *Rhodococcus*, have been developed and applied to confirm the identified degradation pathways for lignin model compounds. Fifth, we found that the alteration of the mycolic acid and phospholipid membrane composition is likely to be a strategy of *R. opacus* for phenol tolerance. Together, our results provide insights into mechanisms for aromatic tolerance and utilization, facilitating the further development of *R. opacus* as an improved chassis for renewable bioproduction.

Publications

1. WR Henson, T Campbell, D DeLorenzo, Y Gao, B Berla, SJ Kim, M Foston, TS Moon and G Dantas. Multi-omic elucidation of aromatic catabolism in adaptively evolved *Rhodococcus opacus*. *Metab. Eng.* 49, 69–83 (2018)

2. WR Henson, F Hsu, G Dantas, TS Moon and M Foston. Lipid metabolism of phenol tolerant *Rhodococcus opacus* strains for lignin bioconversion. *Biotechnol Biofuels*. 11:339 (2018)

3. DM DeLorenzo and TS Moon. Selection of stable reference genes for RT-qPCR in *Rhodococcus opacus* PD630. *Sci. Rep.* 8:6019 (2018)

4. DM DeLorenzo, AG Rottinghaus, WR Henson and TS Moon. Molecular toolkit for gene expression control and genome modification in *Rhodococcus opacus* PD630. *ACS Synth. Biol.* 7, 727–738 (2018)

This work is funded by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research, Biological System Sciences Division, award # DE-SC0018324.

A Concerted Systems Biology Analysis of Aromatic Metabolism in *Rhodococcus opacus* PD630

Garrett W. Roell[1]*, Rhiannon R. Carr[1], Tayte Campbell[2], Zeyu Shang[1], William R. Henson[1], Jeffrey Czajka[1], **Hector Garcia Martin**[3,4,5,6], **Fuzhong Zhang**[1], **Marcus Foston**[1], **Gautam Dantas**[2,7,8,9], **Tae Seok Moon**[1], and **Yinjie J. Tang**[1]

[1]Department of Energy, Environmental and Chemical Engineering, Washington University in St. Louis, St. Louis, MO, 63130, USA; [2] The Edison Family Center for Genome Sciences and Systems Biology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63110, USA; [3] DOE, Joint BioEnergy Institute, Emeryville, CA, 94608, USA; [4] DOE, Agile BioFoundry, Emeryville, CA, 94608, USA; [5] Biological Systems and Engineering Division, Lawrence Berkeley National Lab, Berkeley, CA, 94720, USA; [6] BCAM, Basque Center for Applied Mathematics, Bilbao, Spain; [7] Department of Pathology and Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63108, USA; [8] Department of Biomedical Engineering, Washington University in St. Louis, St Louis, MO, 63130, USA; [9] Department of Molecular Microbiology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63108, USA

Project Goals:

1. **Use ^{13}C-MFA to reveal R. opacus' phenol metabolism**
2. **Connect flux data with transcription profiling and metabolite analysis to show phenol metabolism regulations.**
3. **Determine whether R. opacus phenol utilization is hindered by other aromatic and non-aromatic substrates.**
4. **Test adaptively evolved strains to determine how their central metabolic network has changed.**

Rhodococcus opacus PD630 metabolizes aromatic substrates and naturally produces branched-chain lipids, which are advantageous traits for lignin valorization. To provide insights into its lignocellulose hydrolysate utilization, we performed ^{13}C pathway tracing, transcriptional profiling, biomass composition analysis, and metabolite profiling in conjunction with ^{13}C-metabolic flux analysis (MFA) of phenol metabolism. We found that 1) phenol is metabolized through the ortho branch of the β–ketoadipate pathway; 2) phenol-fed cultures have high TCA cycle fluxes with overflow succinate secretion; 3) NADPH is generated mainly via NADPH-dependent isocitrate dehydrogenase; 4) Active cataplerotic fluxes increase plasticity in the TCA cycle; and 5) gluconeogenesis occurs partially through the reversed Entner–Doudoroff pathway (EDP). We also found that phenol-fed *R. opacus* PD630 generally has lower sugar phosphate concentrations (e.g., fructose 1,6-bisphosphatase < 5%) compared to metabolite pools in glucose-fed *Escherichia coli* (set as 100%), while pool sizes of its TCA metabolites (malate, succinate, and α-ketoglutarate) are higher than those in *E. coli*. In addition, glucose catabolite repression is absent in *R. opacus*, but phenol utilization can be hindered by the presence of other aromatic substrates (e.g., benzoate). Three adaptively-evolved strains display different growth rates when fed with phenol as a sole carbon source, but they demonstrate a conserved central flux network**.**

This project is supported by the DoE BER grant (DESC0018324).

Exploring the Aging of Lignin Breakdown Products for Biological Conversion

Zeyu Shang[1] (zeyu.shang@wustl.edu), Rhiannon Carr[1], **Tae Seok Moon[1]**, **Gautam Dantas[2,3,4,5]** , and **Marcus Foston[2]**

[1] Department of Energy, Environmental & Chemical Engineering, Washington University in St. Louis, St. Louis, MO, 63130, USA; [2]The Edison Family Center for Genome Sciences and Systems Biology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63110, USA; [3] Department of Pathology and Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63108, USA; [4] Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO, 63130, USA; [5] Department of Molecular Microbiology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63108, USA

Project Goals: We aim to combine adaptive evolution and multi-omic approaches to identify aromatic tolerance mechanisms in *Rhodococcus opacus*. We aim to utilize identified genes and pathways to engineer an optimized *R. opacus* strain for the conversion of lignocellulose into valuable products.

Future commercial viability of second generation biofuel and biochemical production depends on lignin valorization. To this end, we have developed a hybrid (i.e., thermo-catalytic and biological) conversion platform to generate value-added products from lignin. This hybrid conversion platform consists of two sequential processes: a thermo-catalytic process that depolymerizes lignin into an aqueous-soluble lignin breakdown products (LBPs) and a biological process that converts and funnels compounds in LBPs into a single bio-product. To overcome well-known challenges to lignin depolymerization, LBPs are generated via a "lignin-first" depolymerization process using an activated carbon supported pallium catalyst and an alkaline solvent from milled poplar. In a practical hybrid conversion process, LBP may have to be stored prior to biological conversion. Moreover, complete microbial utilization of the LBP can take up to 72 hours. During our early efforts, we noted that over the course of several days, our LBP underwent a significant color change while a solid precipitant formed. As a result we studied, the stability of LBPs stored at different temperatures (room temperature and 4 °C) over the course of 32 days. We characterized the LBPs for compositional changes and changes in the growth and utilization of the LBP by the bacterium, *Rhodococcus opacus*, as a function of aging time.

This work is funded by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research, Biological System Sciences Division, award # DE-SC0018324.

Solvent-induced Membrane Disruption Revealed by Neutron Scattering and Molecular Simulation

Micholas Dean Smith,[1,2] Sai Venkatesh Pingali,[1] James G. Elkins,[1] John Katsaras,[1] Loukas Petridis,[1] Hugh O'Neill,[1,2] Robert Frank Standaert,[3] Dmitry Bolmatov,[1] Frederick Heberle,[4] Jeremy Smith,[1,2] and **Brian H. Davison[1,2]* (davisonbh@ornl.gov)**

[1]Oak Ridge National Laboratory, Oak Ridge, TN; [2]University of Tennessee Knoxville, Knoxville, TN, [3]East Tennessee State University, Johnson City, TN, [4]University of Texas Health Science Center at Houston, Houston, TX

http://cmb.ornl.gov/index.php/research/bioenergy/dynamic-visualization-of-lignocellulose

Project Goals: To realize the potential of lignocellulosic biomass to play a major role in generation of renewable biofuels, key limitations in biomass pretreatment and microbial fermentation need to be addressed. For both processing steps, presence of nonaqueous co-solvents can disrupt key biological structures. In the case of biomass, this disruption is desirable, as it facilitates the solubilization and fractionation of lignocellulosic polymers for subsequent cellulosic conversion and lignin valorization. In contrast, the solvents can inhibit fermentation by disruption of microbial cell membranes. The Scientific Focus Area in Biofuels is developing "Visualization of Solvent Disruption of Biomass and Biomembrane Structures in the Production of Advanced Biofuels and Bioproducts" for multiple-length scale, real-time imaging during processing with non-aqueous co-solvents to provide the fundamental information that is needed to improve conversion of renewable lignocellulosic biomass to biofuels.

A challenge to the efficient conversion of plant biomass to biofuels is obtaining high titers of biofuels and biofuel precursors from microbial fermentation. Key to overcoming this challenge is developing robust fermentative microbes with membranes that can sustain the solvent stressors introduced during the fermentation of pretreated biomass. Here, using a combination of atomistic Molecular Dynamics (MD) simulations and Small-Angle Scattering Neutron Scattering (SANS), we determine how tetrahydrofuran (THF), a residual pretreatment solvent, and butanol, a fermentation product, disrupt a model bacterial membrane composed of 70:30 POPE (1-palmitoyl-2-oleoyl phosphatidylethanolamine) and POPG (1-palmitoyl-2-oleoyl phosphatidylglycerol). Both SANS and MD show that both solvents decrease the thickness of the model membranes, and that the disruption caused by butanol is greater than that by THF. Further, the simulations show that organic solvents increase membrane "fluidity" and that butanol is localized mainly at the lipid head groups, whereas THF is found both around the head groups and the tails. The fundamental understanding obtained here of how nonaqueous solvents can be detrimental to fermentative microbes by disrupting cell membranes will provide a rational basis to develop optimized industrial biomass conversions by limiting inhibition of microbial fermentation.

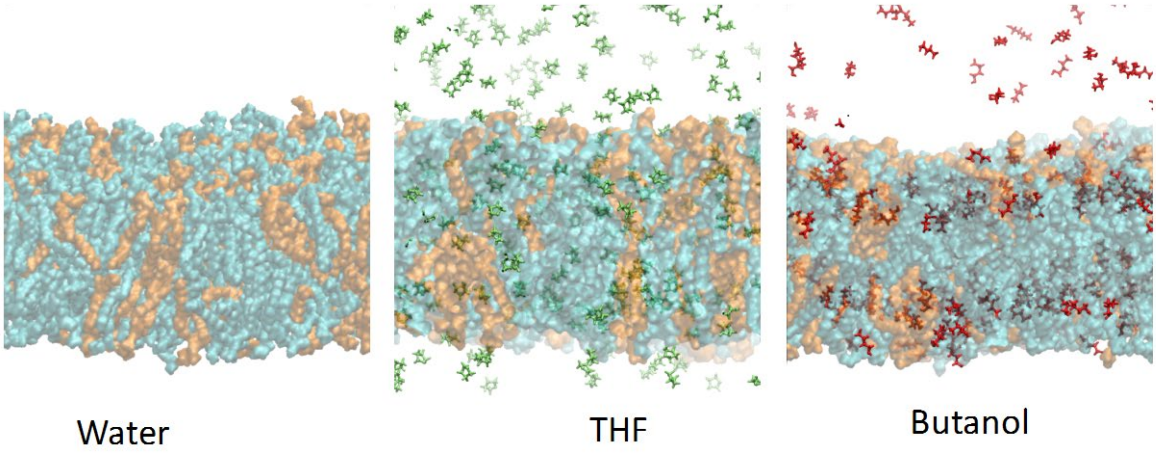


Figure: MD snapshots. POPE is cyan, POPG orange, THF green and butanol red. Water molecules not shown.

Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under contract no. DE-AC05-00OR22725. This program is supported by the U. S. Department of Energy, Office of Science, through the Genomic Science Program, Office of Biological and Environmental Research, under FWP ERKP752.

Structural Characterization of Poplar Variants Provides New Insights into Plant Cell Wall Recalcitrance

Riddhi Shah[1,2], Samarthya Bhagia[1], Sai Venkatesh Pingali[1], Arthur Ragauskas[1,2], Jaroslav Ďurkovič[3], Rastislav Lagaňa[4], Monika Kardošová[3], Wellington Muchero[1], Jay Chen[1], Hugh O'Neill*[1,2] (oneillhm@ornl.gov), **Brian H. Davison**[1]

[1]Oak Ridge National Laboratory, Oak Ridge, TN; [2]University of Tennessee Knoxville, Knoxville, TN, [3]Department of Phytology, Technical University in Zvolen, 960 53 Zvolen, Slovak Republic, [4]Department of Wood Science, Technical University in Zvolen, 960 53 Zvolen, Slovak Republic

http://cmb.ornl.gov/index.php/research/bioenergy/dynamic-visualization-of-lignocellulose

Project Goals: To realize the potential of lignocellulosic biomass to play a major role in generation of renewable biofuels, key limitations in biomass pretreatment and microbial fermentation need to be addressed. For both processing steps, presence of nonaqueous co-solvents can disrupt key biological structures. In the case of biomass, this disruption is desirable, as it facilitates the solubilization and fractionation of lignocellulosic polymers for subsequent cellulosic conversion and lignin valorization. In contrast, the solvents can inhibit fermentation by disruption of microbial cell membranes. The Scientific Focus Area in Biofuels is developing "Visualization of Solvent Disruption of Biomass and Biomembrane Structures in the Production of Advanced Biofuels and Bioproducts" for multiple-length scale, real-time imaging during processing with non-aqueous co-solvents to provide the fundamental information that is needed to improve conversion of renewable lignocellulosic biomass to biofuels.

Poplar is a fast-growing hardwood but field-grown poplars vary significantly in cell wall properties that affects efficiency of biomass conversion technologies[1]. Understanding the structural differences that impact biomass recalcitrance is crucial to valorize this important bioenergy source. A genome wide association study identified two naturally occurring poplars, BESC-316 and GW-11012, that had 22% and 18% reduced klason lignin content and syringyl/guaiacyl ratio of 2.58 and 3.13, respectively. The lower lignin content genotype showed greater sugar release before and after hot water pretreatment compared to the higher lignin counterpart[2]. Atomic force microscopy using PeakForce quantitative nanomechanical property mapping found that modulus of elasticity was significantly higher in secondary cell walls than the compound middle lamella in both variants. Lignin content in these poplars did not affect modulus of elasticity, dissipation and deformation forces of cell walls. Transmission electron microscopy revealed that for the same type of anatomical cells, lignin distribution varies significantly in poplar variants. Moreover, lower syringyl/guaiacyl ratio of high lignin variant agreed with the assertion that lower S/G ratio in poplar can negatively affect sugar yield. The structural properties of BESC-316 and GW-11012 were studied using small-angle neutron scattering (SANS) and wide-angle X-ray scattering (WAXS) before and after hot water

pretreatment. Cellulose microfibril arrangement in GW-11012 is consistent with aggregated microfibrils and differed significantly from the well-ordered cellulose microfibrils in BESC-316 before pretreatment. Post-pretreatment, little change was seen in cellulose arrangement for GW-11012 whereas BESC-316 showed aggregation of microfibrils. After pretreatment, both genotypes have very similar scattering patterns indicating that the structural changes that occurred in the cell walls were similar. Cellulose accessibility measured using the modified Simons' stain before and after pretreatment was similar for GW-11012 and BESC-316. Our data suggest that lignin distribution and cellulose organization both play an important role in lowering lignocellulose recalcitrance. Through the application of recently developed AFM PeakForce quantitative nanomechanical property mapping and first use of SANS for comparing field-grown poplar variants, this study presents new advancements in understanding cell wall recalcitrance of poplar.

The poplar variants were provided by the Center for Bioenergy Innovation.

References

1. Muchero et al, High-resolution genetic mapping of allelic variants associated with cell wall chemistry in Populus, BMC Genomics, (2015);16 (1),24.

2. Bhagia et al. Natural genetic variability reduces recalcitrance in poplar, Biotechnology for Biofuels (2016); 9:106

Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under contract no. DE-AC05-00OR22725. This program is supported by the U. S. Department of Energy, Office of Science, through the Genomic Science Program, Office of Biological and Environmental Research, under FWP ERKP752.

Dynamics of the Lignin Glass Transition

Loukas Petridis[1]* (petridisl@ornl.gov), Derya Vural,[1] Hugh M. O'Neill,[1] Yunqiao Pu,[1] Sai Venkatesh Pingali,[1] Arthur J. Ragauskas,[1] Jeremy C. Smith,[1] and **Brian H. Davison**[1]

[1]Oak Ridge National Laboratory, Oak Ridge, Tennessee

http://cmb.ornl.gov/index.php/research/bioenergy/dynamic-visualization-of-lignocellulose

Project Goals: To realize the potential of lignocellulosic biomass to play a major role in generation of renewable biofuels, key limitations in biomass pretreatment and microbial fermentation need to be addressed. For both processing steps, presence of nonaqueous co-solvents can disrupt key biological structures. In the case of biomass, this disruption is desirable, as it facilitates the solubilization and fractionation of lignocellulosic polymers for subsequent cellulosic conversion and lignin valorization. In contrast, the solvents can inhibit fermentation by disruption of microbial cell membranes. The Scientific Focus Area in Biofuels is developing "Visualization of Solvent Disruption of Biomass and Biomembrane Structures in the Production of Advanced Biofuels and Bioproducts" for multiple-length scale, real-time imaging during processing with non-aqueous co-solvents to provide the fundamental information that is needed to improve conversion of renewable lignocellulosic biomass to biofuels.

At room temperature, lignin is mechanically rigid, which impedes industrial processing. High temperature is employed to soften it. At the molecular scale, this is achieved by enhancing the underlying lignin atomic dynamics. We combined MD simulations with neutron scattering experiments to understand the lignin glass transition at the atomic level. We characterized the atomic motions giving rise to the technologically important lignin glass transition and how they differ above and below the glass transition temperature T_g [1]. Below T_g, lignin exhibits mainly internal and localized motions. Above T_g, the mobility of segments, consisting of 3–5 lignin monomeric units, is enhanced. The temperature dependence of the lignin relaxation time was found to switch from Arrhenius to non-Arrhenius as the temperature increased above T_g. Despite the heterogeneous and complex structure of lignin, its glass transition dynamics can be described by concepts developed for chemically homogeneous polymers. We further probed the dependence of lignin dynamics on its chemical composition. We found syringyl (S) units to be more dynamic than guaiacyl (G) and the three-carbon chains to be more dynamic than the phenol rings[2]. The data are consistent with previous studies showing lignin with a high S/G ratio has lower T_g [3]. Heating biomass above the lignin T_g is expensive. Our results show that using feedstocks whose lignin has a higher syringyl content may offer ways to lower processing costs.

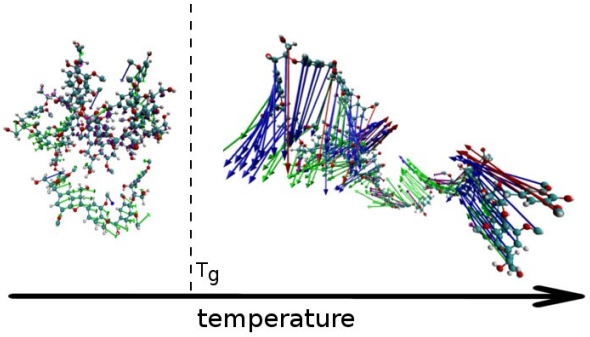


Figure: The motions of atoms in lignin below and above Tg. Red arrows indicate the largest atomic motions, green arrows the smallest.

References

1. Vural, D.; Smith, J. C.; Petridis, L., Dynamics of the lignin glass transition. *Phys. Chem. Chem. Phys.* **2018,** *20* (31), 20504-20512.
2. Vural, D.; Gainaru, C.; O'Neill, H.; Pu, Y.; Smith, M. D.; Parks, J. M.; Pingali, S. V.; Mamontov, E.; Davison, B. H.; Sokolov, A. P.; Ragauskas, A. J.; Smith, J. C.; Petridis, L., Impact of hydration and temperature history on the structure and dynamics of lignin. *Green Chem.* **2018,** *20* (7), 1602-1611
3. Sun, Q. N.; Khunsupat, R.; Akato, K.; Tao, J. M.; Labbe, N.; Gallego, N. C.; Bozell, J. J.; Rials, T. G.; Tuskan, G. A.; Tschaplinski, T. J.; Naskar, A. K.; Pu, Y. Q.; Ragauskas, A. J., A study of poplar organosolv lignin after melt rheology treatment as carbon fiber precursors. *Green Chem.* **2016,** *18* (18), 5015-5024.

Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under contract no. DE-AC05-00OR22725. This program is supported by the U. S. Department of Energy, Office of Science, through the Genomic Science Program, Office of Biological and Environmental Research, under FWP ERKP752.

Influence of microbial diversity on carbon use efficiency in soils

L.A. Domeignoz-Horta[1], G. Pold[2], X.J.A. Liu[1], J. M. Melillo[3],
S.D. Frey[4] and **K.M. DeAngelis[1]**

[1]Department of Microbiology, University of Massachusetts, Amherst, MA 01003, USA. [2]Graduate Program in Organismic and Evolutionary Biology, University of Massachusetts, Amherst, MA 01003, USA. [3]Ecosystems Center, Marine Biological Laboratory, Woods Hole, MA 02543, USA. [4]Department of Natural Resources and the Environment, University of New Hampshire, Durham, NH 03824, USA.

Goals:

1. To develop a soil model system in which microbial communities and abiotic conditions can be manipulated
2. Quantify the relative importance of biotic and abiotic drivers of CUE

Abstract:

Soils retain the largest organic carbon pool in the terrestrial biosphere and represent an important source of carbon dioxide (CO_2) to the atmosphere. Microbial carbon use efficiency (CUE) – or the fraction of C taken up by a cell that is converted to biomass – is a central determinant of how much of this carbon is able to be retained in soil. CUE has been proposed to vary with both biotic and abiotic conditions. Therefore, understanding how this pivotal variable in the C cycle will respond to the dually changing climate and biodiversity is of utmost importance. The present study addresses how microbial diversity, community composition, rising temperatures and different moisture impact CUE and its consequences for soil CO_2 fluxes.

Here, we extracted communities exposed to chronic warming and ambient temperature controls from a long-term experimental site at the Harvard Forest. The diversity of these communities was manipulated prior to inoculation in an artificial soil environment. Microcosms were incubated for 120 days under two temperatures and moistures. Subsequently ^{18}O-H_2O was added and CUE was evaluated after microbes were growing for 24h. The diversity of bacteria and fungi was determined by Illumina MiSeq. Because there was no observed difference in CUE that could be attributed to long-term warming treatment, this factor was dropped from the analysis, though it was maintained in the experiments.

Overall, more diverse treatments showed higher CUE compared to less diverse treatments. The greatest difference in CUE was observed between microcosms containing bacteria and those containing both bacteria and fungi, where microcosms with fungi had higher CUE compared to those without. Moisture and incubation temperature influenced the bacterial and fungal microbial communities. Moreover, moisture was a stronger predictor of communities than temperature, explaining 7.9% and 2.5% of bacterial community structure, respectively. CUE decreased in response to increasing temperature from 15 to 25ºC, ranging from -0.03 to +0.008 $^{o}C^{-1}$. Decreasing moisture from 60 to 30% water holding capacity mostly decreased CUE to up to 56%. Bacterial community structure explained the highest fraction of variance in CUE (25%), while moisture and temperature explain both less than 1%. Fungal alpha and beta diversity showed a minor influence on CUE. This could be due to a slower growth of fungi compared to

bacteria which has not been captured by the relatively short incubation time for the CUE ^{18}O-H_2O method. We showed that moisture, a factor that has been overlooked in many CUE studies, was important to understand microbial CUE. Altogether, these results lay the foundation for a better understanding on how microbial community composition and abiotic factors may interact to drive changes in soil carbon cycling.

Funding statement:

This work is supported by the DOE Genomic Sciences Program under Contract No. DE-SC0016590.

Heavy, Wet, and Stuck in the Mud: Quantifying the Effect of Violating Common Assumptions of the $^{18}OH_2$ Method of Measuring Carbon Use Efficiency

Grace Pold[1,*] (apold@umass.edu), Luiz D. Horta,[2] **Kristen M. DeAngelis**[2]

[1]Graduate Program in Organismic and Evolutionary Biology, University of Massachusetts, Amherst, MA;
[2] Department of Microbiology, University of Massachusetts, Amherst, MA

Project Goals:

Our goals are 1. to determine how violating assumptions of the $^{18}OH_2$ method affect estimates of microbial growth during soil carbon use efficiency measurements; and 2. to quantify the impact of violating these assumptions on conclusions drawn regarding the sensitivity of carbon use efficiency to biotic and abiotic manipulations.

Abstract:

Measuring carbon use efficiency (CUE) – or the fraction of carbon taken up by a cell and incorporated into biomass – is essential to understanding the ability of soil microbial communities to retain carbon on short to intermediate time frames. Historically, CUE was measured by tracing the heavy carbon from single ^{13}C-labeled compounds into biomass and respiration. However, in response to dual critiques of failure to distinguish between the uptake and anabolism of substrates, and failure to assess growth efficiency on realistic substrates, the $^{18}OH_2$ incorporation into DNA method was developed. Accurate use of this heavy water method requires making a number of assumptions, but the consequences of violating these assumptions is unknown.

Here we used a combination of a literature search and numerical simulations to evaluate the effect of violating these assumptions on the degree to which CUE is biased. These include heterogeneous microbial growth rates and DNA contents; DNA and microbial biomass carbon extraction efficiencies; and the use of sources other than exogenous water for oxygen in DNA.

We found that estimates of CUE using the $^{18}OH_2$method are sensitive to both DNA and MBC extraction efficiency biases, particularly when the bulk MBC:DNA ratio used for converting DNA growth to carbon growth is not representative of the actively growing community. We will apply our results to our own $^{18}OH_2$-CUE data in order to validate and interpret results in light of the method's assumptions - not in spite of them - and encourage other researchers to do the same. Once complete, R scripts will be made publicly available for this purpose.

Funding statement.

This work is supported by the DOE Genomic Sciences Program under Contract No. DE-SC0016590, and the American Association for University Women American Dissertation Fellowship Program.

Using Marginal Lands for Bioenergy Crop Production in the Rainfed Regions of Continental U.S.: Economic Potential and Environmental Benefits

Sarah Acquah[1*](sa31@illinois.edu), Chongya Jiang[1], Kaiyu Guan[1], and **Madhu Khanna[1]**

[1]University of Illinois at Urbana-Champaign, Urbana, IL.

https://cabbi.bio/research/sustainability-theme/

Project Goals: This project seeks to: (A) assess the availability of low-quality land that is currently not under active and continuous crop cover or continuously under grass cover in the Continental United States; and (B) determine the supply of biomass from energy crops (Miscanthus and Switchgrass) and crop residues (corn stover and wheat straw) at various biomass prices considering the availability of different land types and the yields of energy crops on these land types.

The Renewable Fuel Standard (RFS), established in 2007 by the Energy Independence and Security Act (EISA), sought to increase biofuel production through conventional and advanced biofuels based on the type of feedstocks. While the mandate for 15 billion gallons of corn ethanol has been met, a minimum level of 16 billion gallons of advanced biofuels from cellulosic biomass is yet to be established. To achieve the cellulosic target, emphasis is placed on cellulosic feedstock production from perennial energy crops such as miscanthus and switchgrass. Both are particularly attractive because of their high yields per unit land, low input requirement, and the ability to grow on low-quality lands that are not productive for row crop agriculture. A major concern with large-scale biofuel production is the competition it poses for land under food/feed production versus fuel, its implications for food/feed prices, and indirect land use change that can offset some of the greenhouse gas benefits of biofuels.

Several studies have proposed that idle, degraded, and abandoned lands could be used to produce perennial energy crops that would avoid competition with food production, since the yields of these crops are less susceptible to soil quality than those of annual row crops (Tilman et al., 2006; Campbell et al., 2008; Fargione et al., 2008; Christopher et al., 2007; Cai et al, 2010; Nijsen et al., 2011; Clifton-Brown et al., 2011). These studies have attempted to measure land availability using different data sources and different definitions of available land categories. Yet, the non-consensus characterization of land availability has led to results that are not only difficult to compare, but also provide only a partial assessment of land available for bioenergy crop production. For example, Campbell et al. (2008) estimate the amount of abandoned land, Cai et al. (2010) estimate marginal agricultural land, while Nijsen et al. (2011) estimate the amount of degraded land available for biomass production. Campbell et al. (2008) defined abandoned land as land no longer in crop or pasture due to the relocation of agriculture and due to degradation from intensive use. Cai et al. (2010) consider several possible definitions of

marginal land based on its biophysical productivity and do not consider whether that land is currently under crop production or idle. Nijsen et al. (2011) define degraded land based on expert consultation with some soil scientists and qualitative soil characteristics that indicate the reduced suitability of that land for crop production, and Gelfand et al. (2013) examine the availability of marginal lands, based on soil quality and slope. In addition, prior assessment of land availability for biofuels based on a biophysical definition of marginal land disregards the fact that land use is determined by economic factors, such as crop prices, which vary over time and can affect the extent to which even low-quality land might be used for conventional crop production.

We propose a novel approach for assessing the amount of marginal land in the rainfed region of the U.S. Marginal land is defined as land that was classified as cropland at a point in the past but has been classified as a non-cropland cover type for at least two consecutive years. Potential marginal lands are identified by removing permanent croplands, permanent natural vegetation, permanent non-vegetated areas, and low-productive soils. This is implemented through a comprehensive spatiotemporal analysis using the 10 years 30-m resolution satellite data and soil survey data across the continental U.S. Available marginal land will be coupled with simulated data on yields of miscanthus and switchgrass for different types of land using the DayCent Model. Data on land availability and the productivity of that land for energy crops will be used with the Biofuel and Environmental Policy Analysis Model (BEPAM) model to examine the supply of biomass from energy crops at various prices at a county scale. BEPAM will be used to analyze the economic incentives for cellulosic feedstocks production on marginal land relative to cropland. We will also quantify the soil carbon sequestration benefits and the greenhouse gas mitigation that can be achieved by growing energy crops for biofuels on marginal lands.

The results are expected to show the amount of marginal land available in rainfed U.S. that is economically and biophysically suitable for energy crop production and its potential to produce biomass from miscanthus and switchgrass. The analysis will also indicate the total biomass likely to be produced and the location and amount of land allocated to it at various biomass prices using BEPAM. This research is based on a dynamic view of marginal land as reflected in the variations in land use (and not just biophysical attributes) and its availability for producing bioenergy crops which varies with economic drivers. The prevailing impression implicit in most analyses to date is that land availability is fixed and static. Our approach incorporates economic determinants of land availability and will provide a realistic view of the production capacity on low-quality land. The results will stimulate discussion on the potential for using marginal land for bioenergy crops and the extent to which these crops can avoid competition for land with food crop production.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Oilcane: Metabolic Engineering and Genome Editing to Improve Energy Density, Agronomic and Conversion Performance of Sugarcane

Fredy Altpeter[1,2*] (altpeter@ufl.edu), Baskaran Kannan[1,2], Guangbin Luo[1,2], Tufan Oz[1,2], Duoduo Wang[1,2], Saroj Parajuli[2], Ratna Karan[2], Hui Liu[1,3], Yunjun Zhao[1,3], Chang-Jun Liu[1,3], and John Shanklin[1,3]

[1]DOE Center for Advanced Bioenergy and Bioproducts Innovation; [2]University of Florida, Gainesville, Fla.; [3]Brookhaven National Laboratory, Upton, N.Y.

https://cabbi.bio/research/feedstocks-theme/

Project Goals: Boosting the accumulation of extractable triacylglycerol (TAG) in sugarcane stems for production of biodiesel; developing and deploying enabling biotechnologies for precision breeding of oilcane; and elevating the agronomic and conversion performance of oilcane postharvest residues.

Metabolic engineering to divert carbon flux from sucrose to oil in a high biomass crop like sugarcane has been proposed as a strategy to boost both energy density of high biomass crops and lipid yields per acre for biodiesel production. Recently we succeeded with metabolic engineering to drastically increase TAG content in vegetative sugarcane tissue (Zale et al. 2016) by upregulating fatty acid synthesis, TAG synthesis and optimization of TAG storage. Current efforts focus on the identification of optimal regulatory elements and transgene variants. A library of stem specific and constitutive promoters, variants of transcription factors involved in fatty acid biosynthesis, as well as transcription terminators was generated and assembled into multi-gene expression constructs using Golden Gate cloning to evaluate the contribution of different regulatory elements and gene variants to lipid accumulation in vegetative sugarcane tissues.

We are also deploying genome editing with CRISPR/Cas9 for genetic improvement using both targeted mutagenesis and/or homology directed repair of targeted double strand breaks to generate “loss” and/or “gain of function” mutants with improved agronomic or feedstock performance. We have successfully developed an efficient HDR mediated genome editing approach conferring herbicide resistance by introduction of mutations in the endogenous acetolactate synthase (ALS) gene in highly polyploid sugarcane. Transgenic sugarcane plants were vegetatively propagated and grown in the greenhouse. Molecular characterization of vegetative progeny indicated the CRISPR/Cas9 mediated multi-allelic mutations was faithfully transmitted to the vegetative progeny. We are also generating a rapid readout system to explore alternative delivery options for the editing reagents, which provide opportunities for increased efficiency and specificity of genome modifications in sugarcane while producing transgene-free events. For this approach, target genes for “loss of function mutations” include those that are well known to create albino or chlorotic plantlets following loss of function mutations like phytoene desaturase (PDS). PDS, an essential plant carotenoid biosynthetic enzyme, is involved in the biosynthesis pathway of β-carotene which provides photoprotection of chlorophyll.

Flowering affects both sugar and biomass yields in sugarcane, since vegetative growth ceases upon flower induction and sucrose that has accumulated in the stalks is re-mobilized for use in reproductive development. Often flowering also leads to dehydration of the stalk tissues, which negatively affects stalk density and plant weight, and may also compromise sugar extraction in

conventional sugarcane or lipid extraction in metabolically engineered lipid cane. Therefore, we will focus on suppression of flowering in sugarcane by CRISPR-Cas9 mediated genome editing. We have identified and sequence confirmed multiple alleles from several candidate flowering inducing genes from sugarcane cultivar CP 96-1252. Several guide RNA's were designed in silico for multi-allelic cleavage of the target genes and we selected superior gRNA's with the help of an in-vitro assay. A vector for delivery of selected gRNA's along with the Cas9 nuclease and selectable marker is currently being constructed and will be transferred to sugarcane.

Lignin is a recalcitrance factor for the conversion of sugarcane processing residues to advanced biofuels like butanol. We have recently demonstrated that reducing lignin content and/or monolignol ratio in sugarcane by genome editing or RNAi translates into improved saccharification efficiency (Jung et al. 2013; Kannan et al. 2018). Currently we are exploring alternative approaches to lignin modification including the expression of enzymes that interfere with the normal process for cell wall lignification. We are currently evaluating the overexpression of an artificial monolignol 4-O-methyltransferase to compromise lignin polymerization. Vector containing different versions of the lignin modification genes were introduced into sugarcane callus by biolistic gene transfer and will be regenerated on the media containing selection agents. Regenerated plants will be characterized for the transgene expression and lignin modification at maturity.

References

1. Zale, J., J.H. Jung, J.Y. Kim, B. Pathak, R. Karan, H. Liu, X. Chen, H. Wu, J. Candreva, Z. Zhai, J. Shanklin, F. Altpeter. 2016. Metabolic engineering of sugarcane to accumulate energy-dense triacylglycerols in vegetative biomass. Plant Biotech. J. 14: 661-669.
2. Jung, J. H., W. E. Vermerris, M. Gallo, J. R. Fedenko, J. E. Erickson, and F. Altpeter. 2013. RNAi suppression of lignin biosynthesis increases fermentable sugar yields for biofuel production from field-grown sugarcane. Plant Biotech. J. 11: 709-716
3. Kannan B., Jung J.H., Moxley G.W., Lee S.-M., Altpeter F. 2018. TALEN mediated targeted mutagenesis of more than 100 COMT copies/alleles in highly polyploid sugarcane improves saccharification efficiency without compromising biomass yield. Plant Biotech. J. 16: 856-866

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Miscanthus sinensis Genome: Updates on the Genome Assembly and Analysis of the Transcriptome and Small RNA

Mohammad B. Belaffif[1*] (mbelaffif@hudsonalpha.org), Brandon T. James,[1] Adam Session,[2] Therese Mitros,[3] Stephen P. Moose,[4] Daniel S. Rokhsar,[3] and **Kankshita Swaminathan**[1]

[1]HudsonAlpha Institute for Biotechnology, Huntsville, AL; [2]Joint Genome Institute, Walnut Creek, CA; [3]University of California, Berkeley; and [4]University of Illinois at Urbana-Champaign, Urbana

Project Goals:

Generated a reference genome for *Miscanthus sinensis* to enable comparative genomics in the *Andropogoneae* grasses.

- Chromosome-scale assembly of the *M. sinensis* genome
- Characterization of subgenome-specific signatures that reflect recent allotetraploidy
- Annotation of both gene and small RNA producing regions to lay the foundation for functional genomics and biotechnology

Understanding key genes and networks regulating stem carbon partitioning, nutrient remobilization, and pereniality

- Identify genes with expression patterns that can be used to engineer pathways in a tissue-specific manner
- Understanding the spatiotemporal dynamics of gene regulatory networks governing carbon and nitrogen partitioning

Miscanthus is a perennial C4 grass which exhibits a number of desirable traits for sustainable biomass production, especially in temperate regions. The availability of a high quality, well annotated, reference genome for *Miscanthus* will provide a foothold for understanding the genetic basis of traits of interest and facilitate genomics enabled breeding. Here we present an update of our progress in providing a chromosome scale genome assembly for *Miscanthus sinensis* (available on Phytozome https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Msinensis_er).

The draft genome presents an almost complete coverage of the 19 chromosomes of the *M. sinensis* genome, with 91% of assembled sequence placed on chromosomes. We have identified LTR retrotransposons that allow us to identify an allotetraploidy within the *Miscanthus* clade and mark the two subgenomes. Our current annotation predicts a total of 67,789 gene models, representing gene coding regions, 82% of which have transcriptional evidence. To better understand the developmental biology *Miscanthus* and mine genes of interest, we have generated both RNA-seq and small RNA-seq data from *Miscanthus x giganteus*, a natural derived triploid of *M. sinensis* and *M. sacchariflorus*. We used the transcriptomic data to annotate *M. sinensis*

genome. By leveraging the high-quality genome of *M. sinensis*, we were able to identify genes preferentially expressed in leaves, stems or rhizomes and are in the process of obtaining a more detailed spatiotemporal expression for these genes using promoter-reporter fusions.

Small RNAs play an important role in plant growth and development and are implicated in important phenomena like fine regulation of gene function and heterosis. We used small RNA sequencing to characterize the types of small RNA and regions producing them. We have categorized the small RNA producing regions into six types: microRNA, putative microRNA, natsiRNA, putative natsiRNA, putative cncRNA, and phasiRNA. We have identified 430 high confidence microRNA or putative microRNA producing regions. Most of these are conserved between the two subgenomes of *M. sinensis*. We are in the process of combining the expression data from both the transcriptomic and small RNA to build a composite gene regulatory network to better understand gene regulation and their role in development in *Miscanthus*.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Developing a Novel Plant-Microbial Interactions Model to Predict the Impacts of Bioenergy Crops on Soil Carbon and Nitrogen Cycling

Edward R. Brzostek[1*] (erbrzostek@mail.wvu.edu), Kara E. Allen[1], Joanna R. Ridgeway[1],

[1]West Virginia University, Morgantown, WV;

https://cabbi.bio

Project Goals: The Center for Advanced Bioenergy and Bioproducts Innovation's (CABBI) mission is to develop efficient ways to grow bioenergy crops, transform biomass into valuable chemicals, and market the resulting biofuels and other bioproducts. A key part of this mission is to ensure the sustainability of a bioenergy economy in the rainfed US by maintaining or enhancing existing ecosystem services (e.g., soil carbon sequestration, nitrogen retention). However, our ability to predict the extent to which different management strategies and bioenergy feedstocks impact ecosystem services is limited by simplified predictive model formulations that do not represent the diversity of microbial traits as well as the ability of plant-microbial interactions to feedback on soil biogeochemical cycling. Thus, the goal of this project within CABBI is to develop and validate plant-microbial interactions model that predicts interactions between plant and microbial traits and the resulting impacts on ecosystem services.

Accurate projections of the impacts of bioenergy crop production on soil carbon (C) and nitrogen (N) cycling by ecosystem models depend on how they represent the interactions between roots, free-living microbes, and symbiotic microbes. While empirical research has highlighted the rhizosphere as a hotspot for the trading of photosynthate C for soil N between roots and microbes; this plant C allocation pathway and the resulting impacts on soil microbial activity are missing from most models. To bridge this knowledge gap, we have adapted a state-of-the-art model that dynamically predicts plant-microbial interactions, Fixation and Uptake of Nitrogen – Carbon, Organisms, and Rhizosphere Processes in the Soil Environment (FUN-CORPSE), to bioenergy crop systems by incorporating representations of tillage, fertilization, and harvest fluxes and timing. In addition, we have begun work to integrate quantitative stable isotope probing data that allows us to develop distinct fungal and bacterial guilds in the model that vary in key traits such as carbon use efficiency and turnover.

First, we confronted and validated the newly developed bioenergy crop model with datasets from two bioenergy crop systems—switchgrass and corn. We ran model experiments to examine the extent to which C-N dynamics vary as a function of agricultural management practices and N availability in soils. Results from our model experiments showed that soil C in corn was more sensitive than in switchgrass to management practices. However, incorporating a no-till method in the model on the corn systems was able to reduce the impact on soil C losses. Altering harvesting methods in our model to allow more standing biomass to remain as crop residues resulted in a significant increase in soil C stocks. Second, we performed a preliminary model

experiment to test whether parameterizing distinct fungal and bacterial guilds that vary in carbon use efficiency, turnover, and their ability to degrade soil organic matter substrates were able to capture soil respiration data from a lab incubation experiment. As we iteratively improved parameterizations of whole microbial community traits to distinct fungal and bacterial traits, the ability of the model to predict soil respiration improved. Collectively, our newly developed bioenergy crop model provides a novel framework that will enhance our ability to predict how interactions between plant and microbial traits impact the sustainability of new bioenergy crops and management techniques.

Funding statement: This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

The Nitrogen Biogeochemical Impacts of Energy Sorghum, a Potential Bioenergy Crop

Mark B. Burnham[1]* (mburnham@illinois.edu), Taylor N. Bozman,[1] Avarna Jain,[1] D.K. Lee,[1] **Evan H. DeLucia**,[1] **Wendy H. Yang**[1]

[1]University of Illinois at Urbana-Champaign, Urbana, IL

https://cabbi.bio/research/sustainability-theme/

Project Goals: We determined if high-yielding energy sorghum inhibits nitrification and denitrification in its rhizosphere compared to bulk soil and if these effects were changed under different fertilization rates. We also measured if sorghum rhizosphere effects resulted in a change in nitrous oxide flux from sorghum field soils compared to maize.

The net environmental benefit of bioenergy crops depends on how they affect agroecosystem nutrient cycling. Nitrogen (N) fertilizer addition affects ground and stream water by increasing leaching of nitrate (NO_3^-), and increases emissions of nitrous oxide (N_2O), a potent greenhouse gas. Perennial biofuel crops produce high yields with low N fertilizer requirements and small NO_3^- and N_2O losses, but their establishment in the Midwestern United States has been limited. Annual sorghum is more easily incorporated into crop rotations, but high-yielding energy sorghum still requires significant N input. However, sorghum can release compounds into the rhizosphere that cause biological nitrification inhibition (BNI), potentially reducing ecosystem losses of NO_3^- and N_2O compared to maize. We measured potential nitrification and denitrification in rhizosphere and bulk soil across two N fertilization rates (0 and 168 kg N ha^{-1}) in a sorghum agronomy trial at the University of Illinois Energy Farm, and measured soil N_2O emissions in separate sorghum and maize plots. Across all dates, sorghum inhibited potential nitrification by 8.6% in the rhizosphere relative to bulk soil ($P = 0.028$). Mid-growing season, when plants were growing fastest, sorghum inhibited potential nitrification by an average of 16% (± 4.8%). This inhibition was stronger in unfertilized plots (26.8%) compared to fertilized plots (11.6%) ($P = 0.025$). Across all dates, potential denitrification was stimulated by 36.6% in the sorghum rhizosphere compared to bulk soil ($P = 0.01$), and N_2O flux from sorghum fields was higher than from maize. In sorghum fields, carbon-rich root exudates may have stimulated denitrification, a heterotrophic process, causing higher N_2O emissions. The decline of BNI with fertilizer addition indicates that BNI is likely facultatively expressed to reduce N loss as NO_3^- and increase N retention in the soil. As fertilizer rates increase, we expect an exponential increase in NO_3^- leaching as BNI declines.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Economic Incentives for Achieving Conservation and Renewable Energy Goals with the Conservation Reserve Program

Luoye Chen*[1] (luoyec2@illinois.edu), Elena Blanc-Betes[1], Tara W. Hudiburg[2], Evan H. DeLucia[1], and **Madhu Khanna**[1]

[1] University of Illinois at Urbana-Champaign, Urbana, IL; [2] University of Idaho, Moscow, Idaho

https://cabbi.bio/research/sustainability-theme/

Project Goals: The goal of this project is to examine the economic and ecological benefits from producing harvestable dedicated energy crops on land enrolled in the Conservation Reserve Program (CRP). Specifically, we examine the economic incentives for growing miscanthus and switchgrass for cellulosic biofuels on land enrolled in the CRP at various biomass prices and the extent to which this can displace oil while also lowering the costs of maintaining the CRP for the government. We also assess the potential for additional carbon mitigation by growing harvestable miscanthus and switchgrass produced on CRP acres and the change in soil nitrate levels as compared to leaving CRP with unharvested permanent cover.

The objectives of reducing dependence on fossil fuels, increasing energy independence and mitigating greenhouse gas (GHG) emissions has led to ambitious policies in the US, such as the Renewable Fuel Standard (RFS), to encourage the use of renewable fuels in the US. At least 16 billion gallons of cellulosic ethanol out of a total 36 billion gallons of biofuels was mandated by the RFS to be blended annually with gasoline by 2022. However, the production of biofuels has raised concerns about the competition for land they pose for food and feed production, resulting in higher global crop prices that lead to indirect land use change by creating incentives for the conversion of non-cropland to crop production and releasing carbon stored in soils and vegetation.

These concerns are mitigated in the case of high-yielding dedicated energy crops, such as, miscanthus and switchgrass, that can be grown productively on low-quality land. This has led to an interest in the potential to use land already enrolled in the CRP in the United States for producing harvestable bioenergy crops like miscanthus and switchgrass. This is appealing for both economic and ecological reasons. It can provide a source of revenue for landowners that could reduce the rising annual rental payments for enrolling land in CRP. Using CRP acres to produce harvestable biomass can enhance soil carbon sequestration, produce low carbon biofuels and reduce soil nitrate levels as compared to leaving CRP with unharvested permanent cover.

In this study, we assess the extent to which the conversion of CRP land to produce harvestable bioenergy crops can contribute to biomass production and displace foreign oil while maintaining or enhancing the ecosystem services and lowering maintaining cost of the government using an

integrated model, which combines the a biogeochemical model (Daycent) with an economic model (BEPAM).

We simulate a Business-as-Usual Scenario (BAU) in which no CRP acres are converted to energy crops, and we examine the optimal land allocation for energy crops, biomass prices, GHG and nitrate effects while meeting the 16 billion-gallon goal. Then we consider four other alternative policy scenarios in which CRP can be used for energy crops with varying levels of reduction in land rental payment by the government to assess the incentives to convert CRP acres in the rainfed region to energy crops and its effect on biomass prices, GHG and nitrate effects. We also consider four additional policy scenarios in which the government provides a cost-share subsidy that covers 50 percent of the cost of establishing harvestable energy crops. We conduct this research by compiling a comprehensive county-specific dataset on acres enrolled in the CRP and their soil attributes. We also compile information on the land rental payment and dominant cover practices adopted as well as expiring CRP acres in each county over the next 15 years, and we overlay various soil attributes data layers on CRP parcels, including, soil texture, bulk density, and PH, that are available from the SSURGO database. Our analysis shows that very limited CRP land will be converted to produce harvestable biomass if farmers are expected to forgo the full rental payments they currently receive for keeping land in CRP. In our results, given the government payment for enrolling in CRP program remains at the full level, 1.6 million ha (66.4%) of CRP acres in the rainfed region will convert to energy crops and produce 60 Million megatons of biomass. If the government payment is reduced by 25 percent, little CRP acres will convert to energy crops, and an additional 0.4 million croplands and 1.2 million marginal lands are required to harvest energy crops to meet the RFS mandate. We also find the cost-share assistance program contributes to the conversion of CRP acres to harvestable energy crops, but the incentive is very limited.

However, harvesting miscanthus and switchgrass in CRP lands would be accompanied by a significant increase in soil carbon stocks and a reduction in nitrate runoff relative to the status quo use of CRP acres. The GHG analysis shows the conversion of CRP acres to harvestable energy crops can increase accumulative soil carbon sequestration of the U.S. by 63.2% over 2017-2030, compared to the BAU scenario. In our simulation of 2030, we found that nitrate leaching can also be reduced by 17% compared to the BAU scenario.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Economics of Plant Oil Recovery: A Review

Ming-Hsun Cheng,[1,2]* (minghsun@illinois.edu), **Bruce S Dien**,[2,3] **Vijay Singh,**[1,2]

[1]University of Illinois at Urbana-Champaign, Urbana, IL; [2]Center for Advanced Bioenergy and Bioproducts Innovation, Urbana, IL; [3]Bioenergy Research USDA ARS, Peoria, IL

https://cabbi.bio/

Project Goals: In this study, plant oil production from solvent extraction and mechanical pressing, including on-farm oil production, industrial and commercial operations are investigated. The goal is to provide an overview of plant oil production cost to form the basis for evaluating oil applications in the oleochemical industry.

Plant oil is a major agricultural commodity used in food, feed, and chemicals. Presently, plant oil is produced from oil seeds either using mechanical pressing or solvent extraction. These technologies have steadily improved, resulting in increased oil recovery; however, production cost is especially important for a commodity. Herein, the costs associated with on-farm pressing, industrial mechanical pressing, and solvent extraction are reviewed. Solvent extraction is the dominant technology because it offers high oil recovery and low production cost. In contrast, industrial mechanical pressing has the highest production cost because of its low oil yield; nevertheless, the simple process results in the lowest fixed capital investment. For on-farm pressing, lower material cost results in lower production cost than industrial mechanical pressing. Additionally, credits from co-products play an important role in determining total revenues, especially for mechanical pressing. Therefore, broadening the applications and values of the co-product is also critical for profitability for the plant oil industry.

Keywords: Plant oil production, On-farm pressing, Industrial mechanical pressing, Solvent extraction, Production cost.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or

recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Bioconversion of Pelletized Big Bluestem, Switchgrass, and Low-Diversity Grass Mixtures into Sugars and Bioethanol

Dien, Bruce S[1] (Bruce.Dien@ars.usda.gov); Mitchell, Robert B [2]; Bowman, Michael J [3]; Jin, Virginia L [4]; Quarterman, Joshu [3]; Schmer, Marty R [4]; Singh, Vija [5]; Slininger, Patricia J. [1]

[1]USDA-ARS, Peoria, IL (United States); Univ. of Illinois, Urbana-Champaign, Urbana, IL (United States)

[2]Univ. of Nebraska, Lincoln, NE (United States)

[3]USDA-ARS, Peoria, IL (United States)

[4]Univ. of Nebraska-East, Lincoln, NE (United States)

[5]Univ. of Illinois, Urbana-Champaign, Urbana, IL (United States)

Project Goals:

To investigate preprocessing, pretreatment, and bioconversion of three bioenergy crops adapted for production on northern marginal farmland.

Three crops of warm-season grasses are being developed for biomass production on northern rain-fed marginal farmland: big bluestem (BBS), switchgrass (SG), and a low diversity mixture of grasses (LDM). In this study, biomass harvested from established fields were compared for pelletization and subsequent conversion to sugars and ethanol. Each biomass was successfully pelletized to similar bulk densities without adding a binder at a commercial feed operation. Pelletizing increased the bulk density by 407% on average and was equally effective on all three biomass samples (528–554 kg/m 3). Chemical analysis of the samples indicated that glucan and xylan contents were slightly reduced during pelletizing (by 23 and 16 g/kg, respectively), as well as theoretical ethanol yields, which are based upon total carbohydrate contents. Pellets and milled straws were pre-treated with either liquid hot-water or low-moisture ammonium hydroxide (LMA) and subsequently hydrolyzed with cellulases. Glucose and total sugar yields were similar for non-pellets and pellets using either pre-treatment; carbohydrates present in pellets were more efficiently recovered compared to non-pellets. LMA pretreated samples were separately hydrolyzed and fermented to ethanol using *Scheffersomyces stipitis* yeast. Hydrolysis recovered 69.7–76.8% of the glucose and 66.5–73.3% of the xylose across all samples. Glucose yields were 251–279 g/kg, db and were significantly lower for SG as compared to the other biomass samples. Recovered sugars were fermented to ethanol at 77.7–86.7% of theoretical yield. As a result, final ethanol yields (245.9–275.5 L/Mg, db) were similar for all of the grasses and estimated to equate to production levels for BBS, LDM, and SG of 1,952, 2,586, and 2,636 l of ethanol per ha, respectively.

References

Dien BS, Mitchell RB, Bowman MJ, Jin VL, Quarterman J, Schmer MR, Singh V, Slininger PJ. Bioconversion of pelletized big bluestem, switchgrass, and low-diversity grass mixtures into sugars and bioethanol. Frontiers in Energy Research. 2018.

Funding statement.

This work was partially funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

Evaluating the effects of the Renewable Fuel Standard on Water Quality: An Integrated Ecosystem-Economic Assessment

Kelsie Ferin[1*] (kmferin@iastate.edu), Sarah Acquah,[2] Jia Zhong,[2] Madhu Khanna,[2] and **Andy VanLoocke[1]**.

[1]Iowa State University, Ames, IA; [2]University of Illinois, Urbana, Ill.

https://cabbi.bio/research/sustainability-theme/

Project Goals: This research aims to assess the impact of Renewable Fuel Standard (RFS) on water quality improvement in the Mississippi River Basin and Gulf of Mexico Hypoxic Zone, above-ground and soil carbon sequestration and GHG emission benefits, and spatial differences in the observed co-benefits.

Increased demand for biofuel production motivated by the Renewable Fuel Standard (RFS) has increased the area of land under corn production to meet the 15 BG corn ethanol mandate. A major consequence of expanded corn-based ethanol production has been intensification of nitrogen application on cropland and worsening of the Gulf of Mexico hypoxia (GMH) problem. Diversifying cropland to produce perennial crops such as Miscanthus and Switchgrass has the potential to reduce nutrient run-off from cropland. However, the extent of these benefits will depend on where energy crops are produced, the type of land allocated to energy crop production and the mix of crops and bioenergy feedstocks across the rainfed region of the US. We used a combination of economic (BEPAM) and ecosystem models (Agro-IBIS and THMB) to quantify the benefits of these mixes under three scenarios based on the RFS policy that represents the next two decades: No-Policy, RFS – corn ethanol only (RFSco), and RFS – corn and cellulosic ethanol (RFScc). BEPAM simulations were used to determine the mix of crop production and management that would meet biofuel mandate production targets at minimal cost and maximal total economic benefit. The water quality simulations were evaluated using observations from the USGS stream monitoring database. The amount of land allocated to annual row crops and fertilizer application increases in the RFSco and RFScc relative to the No-Policy scenario. Total nitrogen delivered to the Gulf of Mexico increased by 6.4% under the RFSco and was reduced by 3.6% under the RFScc, compared to the No-Policy. Our findings indicate that implementing cellulosic bioenergy production under current economic policy scenarios could reduce GMH, however simulated nitrogen export surpassed GMH targets even in the RFScc. This suggests that for future renewable fuel policies to meet GMH targets, they will likely require more targeted changes to land use and crop management that are linked to water quality improvements and costs of nutrient reduction.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

TOWARDS PARAMETERIZING THE CENTRAL DOGMA OF MOLECULAR BIOLOGY ON A GENOME-WIDE SCALE

Meera Gupta*,[1,2] Hsin-Jung (Sophia) Li,[1] Ned Wingreen,[1] Zemer Gitai,[1] **Martin Wühr[1,2 #]**

1 Department of Molecular Biology & the Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA
2 DOE Center for Advanced Bioenergy and Bioproducts Innovation
*presenting author: meerag@princeton.edu
correspondence: wuhr@princeton.edu, 617-2307625

https://cabbi.bio/

Project Goals: Flux through any biochemical network depends on enzyme abundance. We aim to develop experimental techniques and statistical framework [1] to quantify all parameters of the central dogma of molecular biology that determine protein expression levels on a genome wide scale. To this end we combine RNA seq, ribosomal profiling, and mass-spectrometry based quantitative proteomics. We aim to determine how the expression level of proteins is set and how cells adapt their proteome to different environments. This will help us understand the economies of enzyme production. This insight will likely help engineer cell's enzyme levels to produce desired metabolic products. We have started the development of these techniques in *E. coli* since its simplicity speeds up the technology development. We anticipate the generated insights and technology to be easily transferrable to yeast and other microbes relevant for biotechnology.

Abstract

Proteins are the direct determinants of cellular functions. Being able to understand how each protein is formed is critical to tuning their expression levels to obtain the desired product. Modulation of protein levels can be achieved by controlling transcription, translation, or protein degradation. While, we understand this regulation for few of the individual genes, a quantitative and comprehensive understanding of regulatory processes is still lacking. To address this shortcoming, we need to parametrize the central dogma of molecular biology for the entire proteome. Recent advances in high throughput genomics have enabled us to globally quantify transcription and translation rates. However, till date, the measurement of degradation contribution at a global scale remains a challenge. This is mainly due to the lack of accurate and precise measurements of absolute and relative protein expression levels.

As a preliminary result, we have global measurements of differential degradation rates in Nitrogen-limited vs the Phosphorous-limited growth conditions for *E. coli*. These measurements were obtained by comparing translation rates, measured through ribosomal profiling, with the relative protein expression levels in the steady state growth conditions. Remarkably, the results indicated that using the recently published TMTc+ method [2] for protein measurements removes the systematic biases in relative proteomics measurements. However, this approach provides only insight into proteins degradation rates that differ between conditions. To confirm our measurements, and to obtain absolute degradation rates we measured the decline of protein

abundances upon ribosomal translation inhibition. This approach has been previously used on a proteome-wide scale with TAP-tagged libraries [3] . However, we find strong evidence that this approach is prone to artifacts and the half-life of many proteins is drastically altered upon drug addition. To overcome this limitation, we have started to measure protein turn-over rates by a pulse-chase experiment with heavy isotope labeled amino-acids. Our preliminary shallow experiments indicate that only a small subset of proteins in *E. coli* (<10%) shows protein half-lives which are significantly shorter than doubling time. Once these techniques have been established, we intent to apply them to yeast species that are relevant for metabolic engineering and determine the parameters that set metabolic enzyme levels under various growth conditions.

References

1. Peshkin, L., et al., *Bayesian Confidence Intervals for Multiplexed Proteomics Integrate Ion-Statistics with Peptide Quantification Concordance.* bioRxiv, 2019: p. 210476.
2. Sonnett, M.Y., E. ; Wühr, M, *Accurate, Sensitive, and Precise Multiplexed Proteomics using the Complement Reporter Ion Cluster.* 2017, Princeton University.
3. Belle, A., et al., *Quantification of protein half-lives in the budding yeast proteome.* Proc Natl Acad Sci U S A, 2006. **103**(35): p. 13004-9.

This work was supported by the U.S. Department of Energy (DOE) under Grant # DE-SC0018420.

Identification of Sorghum stem-preferred genes and their cis-regulatory elements through de novo promoter analysis

Madara Hetti Arachchilage[1,2]* (mhettiar@illinois.edu) and **Amy Marshall-Colon**[1,2,]

[1]University of Illinois at Urbana-Champaign, Urbana, Ill.; [2]DOE Center for Advanced Bioenergy and Bioproducts Innovation, Urbana, Ill.

https://cabbi.bio/research/feedstocks-theme/

Project Goals: The goal of this study is to mine tissue-specific transcriptome data sets over development to identify stem preferred genes and their promoter elements that drive strong expression in a stem-specific manner. This study will be expanded by performing experiments to confirm applicability of promoters of these stem-preferred gene to drive strong expression of genes involved in oil biosynthesis for improvement of plant-derived fuel.

Plant-derived fuels present an exciting alternative to fossil fuels. However, a major roadblock to making biofuels a viable alternative is that the amount of energy invested given the amount of return is imbalanced. Increasing the total bio-product content per plant could help to balance the equation, especially when scaled up to the field level. Accumulation of oils or oil precursors in large grass stems is one promising strategy. By identifying stem-preferred genes and exploring their regulatory promoters, we may be able to synthetically drive the strong expression of genes involved in oil biosynthesis and transport in grass stems. Use of native tissue- and developmental stage-specific promoters for genetic engineering may also avoid known negative effects of constitutive promoter expression, such as dwarfing. We performed transcriptome analysis on RNA-seq data from sorghum samples taken from specific tissues over development to identify stem-preferred genes over time and space (McCormick et al, 2018). Using a stringent computational analysis, we identified 14 high confidence, stem-preferred genes, including lipase, a lipid transporter, xylem/phloem pattern formation related proteins, and several transcription factors (TFs). These genes had tissue-specific expression in stems at early developmental stages (juvenile, vegetative); however, during the later stages these genes were expressed at similar levels among stem, flag leaf, and reproductive tissues. Promoter analysis of the stem-preferred genes revealed an over-representation of cis-regulatory elements (CREs) for 14 TF families including MYB, bZIP, C2H2, Dof, and ERF. Subsequent co-expression analysis of the 14 stem-specific genes with all stem-expressed TFs (belonging to the discovered TF family CREs) revealed dynamic regulation of these genes by different TFs at different stages. We are poised to experimentally explore the regulatory role of the discovered TFs and promoter CREs in driving stem-specific gene expression in sorghum.

References

1. McCormick et al., (2018), The Sorghum bicolor reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. Plant J, 93: 338-354. doi:10.1111/tpj.13781

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Modeling Energy Sorghum Emissions for the Rainfed United States

Tara Hudiburg[1] (thudiburg@uidaho.edu), Jeffrey Kent,[1] and Danielle Berardi,[1]

[1]University of Idaho, Moscow, Idaho

https://cabbi.bio/research/sustainability-theme/

Project Goals: We are using new and existing experimental observations to calibrate and develop biogeochemical models with the explicit goal of improving predictions of biofuels yield and associated greenhouse gas emissions. The experimental data will be used to develop and test new mechanistic knowledge about above- and below-ground plant physiological response to changing environmental conditions (e.g. climate, soil nutrients), management practices (e.g. fertilization, tillage, land type) and varied genotypes. We will also develop the models to represent new genotype plant functional types. For example, as the plants are modified to produce more lipids, this will require new pathways for carbon allocation in the biogeochemical models.

The U.S. Renewable Fuel Standard 2 mandates use of increasing volumes of cellulosic biofuels in the U.S. fuel supply. A substantial body of research has been conducted to design agricultural systems capable of producing large amounts of cellulosic feedstocks, primarily from perennial grasses such as switchgrass (*Panicum virgatum*), energy cane (*Saccharum spp.*), and Miscanthus (*Miscanthus x giganteus*). However, the perennial growth habit of these crops necessitates major changes in equipment and agronomic practice and presents multi-year financial risks to farmers considering their cultivation.

Sorghum (*Sorghum bicolor (L.) Moench*) is a species of heat- and drought-tolerant C4 grass widely cultivated for food, feed and forage. Sorghum can be integrated more readily into traditional crop rotations, and can be managed with conventional equipment. Recent breeding efforts have focused on photoperiod-sensitive varieties that continue vegetative growth until harvested or killed by frost.

The biogeochemical impacts of large-scale biomass sorghum cultivation may differ substantially from those of either corn perennial biomass crops. This work advances understanding of those impacts by first calibrating the DayCent ecosystem model to energy sorghum field data from several years and sites spanning the central and eastern U.S., and then comparing simulated emissions from several thousand sites across the rainfed U.S. under either energy sorghum or conventional corn cultivation. Sorghum emissions averaged 21.2 g CO_2-equivalent per MJ of ethanol, somewhat lower than the corn average of 25.2, but was substantially lower-emitting in several states including Indiana, Kentucky, and Texas. The large fraction of biomass removed during sorghum harvest caused losses of soil organic C on many sites, but also reduced inputs of organic N available for microbial transformation to N_2O. Work is ongoing to assess the potential

for no-till management and lower residue collection rates to further improve the emissions profiles of energy sorghum at these sites.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Integrating transcriptomics and metabolomics in the oleaginous yeast *Lipomyces starkeyi* for the production of biofuels and bioproducts

Sujit Sadashiv Jagtap* (jagtap@illinois.edu), Anshu Deewan, Jing-Jing Liu, Hanna Walukiewicz, Matthew Plutz, and **Christopher V Rao**

Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois at Urbana-Champaign

https://cabbi.bio/

Project Goals: The goal of this project is to engineer *Lipomyces starkeyi* for the production of biofuels and bioproducts from lignocellulosic sugars. We are interested in identifying sugar transporters, understanding the lipogenesis process, and integrating transcriptomics and metabolomics in the oleaginous yeast *L. starkeyi*.

Oleaginous yeasts are a promising platform for the production of lipids, polyols, and alcohols [1-3]. The lipids can be utilized directly or can be processed into biolubricants, surfactants, and liquid fuels. *L. starkeyi* has the ability to naturally utilize a variety of carbon sources, especially those from sustainable biomass-derived sugars such as glucose, xylose, and cellobiose. However, *L. starkeyi* has been relatively unexplored due to insufficient knowledge of its physiology as well as the lack of efficient genetic tools [4].

In our study, we have evaluated the growth of *L. starkeyi* on different carbon sources and performed transcriptome and metabolome analysis to understand the underlying mechanisms of sugar metabolism. We also identified the putative transporters for the utilization of glucose, xylose, and cellobiose. We functionally characterized sugar transporters from *L. starkeyi* in *Saccharomyces cerevisiae*. One putative sugar transporter showed promising results in the co-consumption of glucose and xylose. Our results may inform the metabolic engineering of *L. starkeyi* strains for production of biofuels and bioproducts from plant-based sugars.

References

1. Jagtap, S. S.; Rao, C. V., Microbial conversion of xylose into useful bioproducts. *Applied Microbiology and Biotechnology* **2018,** *102* (21), 9015-9036.
2. Jagtap, S. S.; Rao, C. V., Production of D-arabitol from D-xylose by the oleaginous yeast *Rhodosporidium toruloides* IFO0880. *Applied Microbiology and Biotechnology* **2018,** *102* (1), 143-151.
3. Wang, W.; Wei, H.; Knoshaug, E.; Van Wychen, S.; Xu, Q.; Himmel, M. E.; Zhang, M., Fatty alcohol production in *Lipomyces starkeyi* and *Yarrowia lipolytica*. *Biotechnology for Biofuels* **2016,** *9* (1), 227.
4. Riley, R.; Haridas, S.; Wolfe, K. H.; Jeffries, T. W., Comparative genomics of biotechnologically important yeasts. *Proceedings of the National Academy of Sciences of the United States of America* **2016,** *113* (35), 9882-9887.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

A Role for Differential Rca Isoform Expression in C4 Bioenergy Grass Thermotolerance?

Sang Yeol Kim[1,2]* (kim6989@illinois.edu), and **Donald R. Ort**[1,2]

[1]University of Illinois at Urbana-Champaign, Urbana, Ill.; [2]Carl R. Woese Institute for Genomic Biology, Urbana, Ill.

https://cabbi.bio/research/feedstocks-theme/

Project Goal: Enhancement of crop production of C4 bioenergy grass by modifying posttranslational modification of Rubisco activase.

Rubisco activase (Rca) uses energy from ATP hydrolysis to remodel the conformation of Rubisco protein (ribulose-1,5-bisphosphate carboxylase/oxygenase), allowing dissociation of inhibitory sugar phosphates from Rubisco active sites, thereby facilitating CO_2 fixation. In Arabidopsis and rice, a single Rca gene generates two protein isoforms (Rca-α and Rca-β) by alternative splicing. The C-terminal extension (CTE) found in the α-isoform contains the two redox-sensitive Cys residues that are known to regulate activity. In addition, Arabidopsis Rca is phosphorylated at Thr-78 (T78) in the dark, which appears to serve a regulatory role in growth and photosynthesis. Despite the knowledge of Rca in C3 plants, little is known about the role of Rca in C4 plants where Rubisco is confined to bundle sheath chloroplasts. Our studies focus on determining the impact of these posttranslational modifications (PTMs) of Rca on C4 plant performance with the rationale that regulation of Rca activity at low light may restrict the ability of photosynthesis to respond to rapid changes in irradiance and as a result, constrain the light use efficiency of photosynthesis. We found that three C4 plants (sorghum, setaria and maize) contain separate genes for the Rca isoforms and Rca-α contains T78 phosphosite and Cys residues in CTE. Interestingly, the Rca-α isoforms were expressed only at high temperature (>40ºC) and the Rca-α proteins were slowly degraded at 25ºC. Gas exchange data suggest the Rca-α expressing sorghum and setaria performed photosynthesis less effectively than the control. Collectively, the results suggest that Rca-α of C4 may play a role at high-temperature tolerance via a redox-sensing mechanism.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Improved High-Throughput Quantification of Nitrogen Cycling Genes in Bioenergy Crop Soils

Jaejin Lee,[1,2] Jinlyung Choi,[1,2] Lanying Ma[1,2*](lanyingm@iastate.edu), Santosh Gunturu,[3,4] Yang Ouyang,[3,4] Jim Cole,[3,4] Lisa Tiemann,[3,4] **Andy VanLoocke**,[2,5] **Emily Heaton**,[2,5] and **Adina Howe**[1,2]

[1]Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, Iowa; [2]DOE Center for Advanced Bioenergy and Bioproducts Innovation (CABBI), Urbana, Ill.; [3]Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, Mich.; [4]Great Lakes Bioenergy Research Center, Madison Wis.; [5]Department of Agronomy, Iowa State University, Ames, Iowa

https://cabbi.bio/research/sustainability-theme/; https://www.germslab.org

Project Goals: We will characterize plant-microbe interactions associated with nitrogen cycling in miscanthus and corn soils. Denitrification genes will be compared between miscanthus and corn crops over a growing season, under different fertilization rates, and with various stand ages. Our results will help us to understand how plant, soil, and microbe interactions provide nutrients to varying feedstocks and how we may sustainably manage their productivity.

Metagenomic sequencing has helped us to understand the immense diversity of microbial strains that participate in nitrogen (N) cycling in soils. However, this diversity represents a challenge for identifying and measuring the key drivers of nitrogen cycling, especially as metagenome sequencing is not practical for large number of samples. Consequently, we still have a limited understanding on the abundance of the microorganisms carrying out these processes and related interactions between soil microbes and bioenergy crops. The development of high-throughput qPCR presents a novel opportunity to characterize nitrogen genes in complex soil environments but is limited by selection of probes for representative gene targets. Our objective was to develop genomic probes for characterizing denitrification in bioenergy crops. Among N cycle pathways, denitrification is an important process in agricultural soils because it removes fixed N that can otherwise be used for primary production and produces N_2O (one of the major global warming gases). Our results confirmed that currently available probes, designed based on well-characterized isolates, cannot cover diverse denitrification genes in our soils. Thus, we developed gene probes that can cover diverse denitrification genes present in agricultural soils and are also appropriate for high-throughput qPCR, to quantify denitrification genes in bioenergy crop soils. A novel primer-design tool, EcoFunPrimer, was used to redesign 384 novel primer sets targeting *nap*A, *nar*G, *nir*K, *nir*S, and *nos*Z genes based on the abundances of gene clusters enumerated in 1,950 publicly available soil metagenomes. These probes were initially tested against a subset of miscanthus soil samples from the CABBI Long-Term Assessment of Miscanthus Productivity and Sustainability (LAMPS) site. Our results show that new primers

can detect a significantly higher proportion of denitrification genes in soil systems and justifies using these probes for all 700 soil samples available from the LAMPS experiment. Further, we are expanding probe design to other N cycle genes.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

CROPSR: An Automated Platform for Complex-Genome Wide CRISPR gRNA Design and Validation

Hans Müller Paul,[1,2,3]* (hmpaul2@illinois.edu), Jacob Heldenbrand,[1,4] and **Matthew Hudson**[1,2,3,4]

[1]University of Illinois at Urbana-Champaign, Urbana, Ill.; [2]DOE Center for Advanced Bioenergy and Bioproducts Innovation, Urbana, Ill.; [3]Carl R. Woese Institute for Genomic Biology, Urbana, Ill., [4]National Center for Supercomputing Applications, Urbana, Ill.

https://cabbi.bio/research/feedstocks-theme/

Project Goals: RNA guided technologies such as CRISPR/Cas9 are a revolutionary tool for crop engineering and synthetic biology. To take advantage of this technology, scientists use online guide RNA design aid software; generally designed based on human and mouse genomes. Differences in genomic architecture between these genomes and the complex, paleopolyploid genomes of bioenergy and bioproduct crops are impacting the utility of the technology. Our goal is to develop a unified software platform to direct edits to specific genes or groups of genes using the CRISPR system. Experimental validation is key to improving both guide RNAs and crop traits, but polymerase chain reaction (PCR) validation can be challenging in complex intergenic regions or genes with multiple copies. The software will design guide RNAs genome-wide, evaluate them based on experimental data including plant genomes, and provide the optimal guides along with unique primers for PCR validation.

Recent advances in genome editing technology such as the CRISPR/Cas9 system have changed the pace and face of functional genomics research. This RNA guided technology is a groundbreaking innovation and an excellent tool for plant biotechnology due to its efficiency, accuracy, and ease of use (1, 2). The key components can often be assembled in a single plasmid, facilitating a quick turnaround time in the experimental procedure. Currently, the default pipeline includes the scanning of the target gene sequences for protospacer adjacent motif (PAM) sites, where the Cas9 enzyme cuts in the DNA, designing and validating the target specificity of the guide RNA, introducing the gRNA and the Cas9 in to the plant and then validating the mutations with PCR and sequencing. Currently, these design steps are mostly done manually by bench scientists, using various software products to help design the gRNA and the PCR primers. This process is labor intensive and can introduce inconsistency in experimental efficiency. We have designed and are building a suite of software that will maximize success probability by controlling all the parameters of CRISPR/CAS9 genome editing construct design and validation, from target selection to accurate validation and interpretation of edits. Our approach consists of firstly pre-computing all potential target PAM sites in the genome, designing a guide RNA for each site, and storing them in a data frame, where we can easily update each entry to add information. We then evaluate each gRNA for their on-site and off-site activities utilizing a machine-learning approach that combines support vector machine (SVM) with L1 logistic regression. The currently accepted standard methods utilize human and mouse genomes for the

training and evaluation steps (3-5), whereas we have also introduced data from the rice genome (6) to improve the predictive power of the model in energy crops such as *Sorghum bicolor*. The scored gRNAs are then associated with the genomic features present in their genomic positions, allowing for easy selection by the user. Our tool also provides PCR primers for validation with uniqueness in the genome, confirmed by Burrows-Wheeler alignment. Our highly parallelized software leverages the power of high-performance computing to provide the user with a wide array of options to choose for each target, allowing them to proceed to the next stages of genome editing experiments with increased confidence of success.

References

1. L. Arora, A. Narula, Gene Editing and Crop Improvement Using CRISPR-Cas9 System. *Front Plant Sci* **8**, 1932 (2017).

2. X. Liu, S. Wu, J. Xu, C. Sui, J. Wei, Application of CRISPR/Cas9 in plant biology. *Acta Pharm Sin B* **7**, 292-302 (2017).

3. J. G. Doench *et al.*, Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* **32**, 1262-1267 (2014).

4. J. G. Doench *et al.*, Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* **34**, 184-191 (2016).

5. J. Listgarten *et al.*, Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs. *Nature Biomedical Engineering* **2**, 38-47 (2018).

6. Y. Lu *et al.*, Genome-wide Targeted Mutagenesis in Rice Using the CRISPR/Cas9 System. *Mol Plant* **10**, 1242-1245 (2017).

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Genomics-Assisted Improvement of Abiotic Stress Tolerance and Yield in the Biomass Feedstock Crops *Miscanthus* and *Saccharum*

Nicholas LaBonte,[1] Liang Xiao,[1,2] Lindsay V. Clark,[1] Brian Baldwin,[3] Jesse Morrison,[3] Kankshita Swaminathan,[4] and **Erik J. Sacks**[1]* (esacks@illinois.edu)

[1]University of Illinois at Urbana-Champaign; [2]Hunan Agricultural University, Changsha, China; [3]Mississippi State University; [4]HudsonAlpha Institute for Biotechnology

https://cabbi.bio/research/feedstocks-theme/

Project Goals:

- **Develop climate-resilient *Miscanthus* and energycane biomass feedstocks by targeting abiotic stress tolerance traits:**
  - ***Miscanthus* adapted to periodic flooding**
  - ***Miscanthus* that can be cultivated on saline marginal lands**
  - **Improve winter-hardiness and biomass yield of *Miscanthus* for the Midwest**
  - **Improve biomass yield and adaptation of *Miscanthus* for the Mid-South**
  - **Introgress cold-tolerance genes from *Miscanthus* into *Saccharum* to expand the potential area of energycane production in the U.S.**
- **Use genomics to improve breeding efficiency and increase rates of genetic gain:**
  - **Identifying associations between genetic polymorphisms and key traits in *Miscanthus* for use in marker-assisted selection**
  - **Development and implementation of genomic selection in *Miscanthus* to shorten breeding cycles**

Abstract:

Cultivars of bioenergy crops that are tolerant of abiotic stress must be developed in order to utilize marginal lands that do not compete with food production. To address this need, our lab is developing *Miscanthus* and *Saccharum* for high biomass production and improved adaptation to marginal lands. We are building and using genomics tools to greatly improve breeding efficiency of these long-lived perennials. In particular, genomic selection will enable us to identify and cross *Miscanthus* selections within ~1.5 years, whereas under traditional phenotypic selection, a breeding cycle for this crop is ~4 years.

In the Midwest, large areas of croplands are subject to seasonal flooding and experience yield-losses due to standing water. These flood-prone lands are broadly-dispersed in the Midwest, on low edges or depressions ('potholes') within otherwise productive fields. Flood-tolerant *Miscanthus* is an opportunity for farmers to gain income from the currently unprofitable portions of their fields that flood in most years. In an interconnected F1 population (n = 809) of three high-yielding *M. sinensis* crossed to a single high-yielding *M. sacchariflorus* ssp. *lutarioriparius*, we have identified genotypes that grow well after being completely submerged, and quantitative trait loci (QTL) for flooding tolerance will be mapped. A high-density SNP dataset for this population is being generated in cooperation with JGI. *M. sacchariflorus* ssp.

lutarioriparius is indigenous to seasonally flooded edges of lakes and rivers in the Yangtze River watershed, so we expected it to be a good source of submergence-tolerance. We have also conducted preliminary experiments on salinity tolerance in this population and discovered that some of the parents and progeny can survive and grow for more than one month when given highly saline irrigation water (300 mM; i.e. about half the salinity of seawater). QTL mapping for salinity tolerance will be conducted in 2019. Additionally, replicated field trials of this population were established in Illinois and Mississippi during 2018, to assess yield.

Though *M.* ×*giganteus* 'Illinois' has high yield-potential (~25 T/ha) and the environmental benefits of a long-lived perennial, it is the only genotype of *Miscanthus* for biomass production currently available in the U.S. for farmers to grow, and it is insufficiently winter-hardy to establish reliably every year in the northern Midwest. In addition to winter-hardiness, new *Miscanthus* cultivars must be high-yielding; however, phenotypic evaluation for yield is slow, requiring field evaluations over three years. To address these problems, we have begun to implement genomic selection for yield, and overwintering ability within *M. sinensis* and *M. sacchariflorus* (the parental species of *M.* ×*giganteus*), and assess the effect of this selection on their 3x *M.* ×*giganteus* progeny from controlled crosses. To ensure reliable testing of overwintering ability each year, we have begun field-testing at the University of Wisconsin's Spooner Agricultural Research Station (USDA hardiness zone 3b), with cooperation from The Great Lakes Bioenergy Center. In previous DOE-funded studies, we identified seven genetic groups of *M. sinensis* and six genetic groups of *M. sacchariflorus*, and developed initial genomic selection models for each species. In 2018, we initiated seed increases for each genetic group. In 2019, we will grow populations of three genetic groups to test and improve the initial genomic selection models, and conduct population improvement for yield and overwintering ability.

Sugarcane (*Saccharum*) is currently the world's most important bioenergy crop. Sugarcane production in the U.S. is currently limited to a small area of southern Florida, Texas, and Louisiana because this tropical crop is susceptible to damage by cold. However, we expect that cane cultivation for bioenergy feedstocks (energycane) could be expanded substantially northward in the southern U.S. (into hardiness zone 8) by introgressing cold tolerance from temperate *Miscanthus* into the tropical canes. During the current autumn-winter crossing season, we have made >20 crosses between canes and highly cold-tolerant *Miscanthus*, in our greenhouse in Illinois; many of these crosses used *Miscanthus* parents from eastern Russia (hardiness zones 3 and 4). If successful, offspring of these crosses will be backcrossed to sugarcane and used to develop cold-tolerant energycane cultivars with the potential to dramatically expand the landscape of bioenergy feedstocks in the U.S.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Cooperative Asymmetric Reactions Combining Photocatalysis and Enzymatic Catalysis

Yajie Wang[1]* (ywang345@illinois.edu), and **Huimin Zhao[1]**

[1]University of Illinois at Urbana-Champaign

https://cabbi.bio/research/conversion-theme/

Project Goals: The goal of this work was to develop a novel cooperative chemoenzymatic reaction capable of stereoconvergent reduction of a mixture of alkenes.

Living organisms rely on simultaneous reactions that are catalyzed by mutually compatible and selective enzymes to synthesize complex natural products and other metabolites. To combine the advantages of these biological systems with the reactivity of artificial chemical catalysts, chemists have devised sequential, concurrent, and cooperative chemoenzymatic reactions that combine enzymatic and artificial catalysts. Cooperative chemoenzymatic reactions consist of interconnected processes that generate products in yields and selectivities that cannot be obtained from the two reactions run sequentially with their respective substrates. This feature of cooperative reactions makes them particularly valuable among chemoenzymatic systems. However, such reactions are difficult to develop because chemical and enzymatic catalysts generally operate in different media at different temperatures and can deactivate each other. As a result of these constraints, the vast majority of cooperative chemoenzymatic processes that have been reported over the last 30 years can be divided into just two categories: 1) chemoenzymatic dynamic kinetic resolutions of alcohols and amines, and 2) enzymatic reactions requiring the simultaneous regeneration of a cofactor. New approaches to the development of chemoenzymatic reactions are needed to enable valuable chemical transformations beyond dynamic kinetic resolutions of alcohols and amines and cofactor regenerations. University of Illinois researchers can report a new class of cooperative chemoenzymatic reactions that combine photocatalysts that isomerize alkenes with ene-reductases that reduce carbon-carbon double bonds to generate valuable enantioenriched products. This method enables the stereoconvergent reduction of E/Z mixtures of alkenes or reduction of the unreactive stereoisomers of alkenes in yields and enantiomeric excess (ee's) that match those obtained from the reduction of the pure, more reactive isomers. The system affords a range of enantioenriched precursors to biologically active compounds. More generally, these results show that the compatibility between photocatalysts and enzymes enables chemoenzymatic processes beyond cofactor regeneration and provides a general strategy for converting stereoselective enzymatic reactions into stereoconvergent ones. We are developing a cooperative system by coupling photocatalyst-catalyzed co-factor regeneration and isomerization with enzymatic reduction in one pot. By doing so, we can realize stereoconvergent reduction of a mixture of alkenes without using an expensive and unstable co-factor such as NADPH.

This work was funded by National Science Foundation under the CCI Center for Enabling New Technologies through Catalysis (CENTC, CHE-1205189 to H.Z. and J.F.H.) and the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any

opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Toward a Dynamic Photosynthesis Model to Guide the Yield Improvement for C4 Energy Crop

Yu Wang[1] (yuwangcn@illinois.edu) and **Stephen Long**[1]

[1] University of Illinois at Urbana Champaign

https://cabbi.bio/research/feedstocks-theme/

Project Goals:

To meet the increasing societal need of energy, one of the missions of the Center for Advanced Bioenergy and Bioproducts Innovation (CABBI) is developing efficient ways to increase the biomass productivity of bioenergy crops and improve the efficiency of conversion from biomass into valuable chemicals.

A dynamic metabolic model from cell to canopy could help us identify potential targets for increasing photosynthesis, water, and nitrogen use efficiency, and improve the biomass productivity of bioenergy crops in various environmental conditions.

1. **Develop a dynamic metabolic model for general C4 plants.**
2. **Parameterize the model using measured data of bioenergy crops, such as sugarcane and sorghum, and predict potential targets for increasing photosynthesis.**
3. **Develop leaf model of C4 bioenergy crop, including 3D leaf anatomy and dynamic metabolic model.**
4. **Develop a canopy model for C4 bioenergy crop, and identify targets for increasing canopy photosynthesis.**

Sorghum and sugarcane, the focal crops of CABBI, belong to the monophyletic grass tribe Andropogonae. The tribe includes the most productive crops and wild species known, which is associated with their use of C4 photosynthesis. All members, as would be expected with their common evolutionary ancestry use Malic enzyme as their primary decarboxylase; this type of C4 being known as C4-ME. Despite high productivities, in sorghum and sugarcane these fall well short of the theoretical maximum crop C4-ME solar conversion efficiency of 6% (Zhu et al., 2008). Understanding the basis of these inefficiencies is key to achieving bioengineering and breeding strategies to increase sustainable productivity and approach the high theoretical efficiencies of C4-ME crops.

Achieving the high potential efficiency of C4-ME photosynthesis requires coordination of a great many metabolic and anatomical features. To quantify the impact of each feature and identify limiting factors, we are developing a generic dynamic systems model of C4 photosynthesis. Previously, we built a C4-ME metabolic model for maize, a close relative of sorghum and

sugarcane, simulating the fluxes of C4 metabolic pathways and their flexibility (Wang et al., 2014ab). In the CABBI version of the model, we are extending the model to include all individual steps in carbon metabolism, and their inter-cellular and inter-organelle trafficking factors affecting dynamic photosynthetic rate, which include posttranslational regulation and temperature response of enzyme activities, dynamic stomata conductance, and detailed light reaction. The model is also being re-parameterized with data from the sorghum and sugarcane materials being used in CABBI. The model outputs are tested in vivo against measured rates of CO2 uptake, water vapor flux and electron flow measured by infra-red gas analysis and modulated chlorophyll fluorescence. Once satisfactory validation is achieved, optimization routines will be used to identify the most promising points for genetic/bioengineering intervention to achieve increased photosynthetic efficiency at the leaf and canopy level following procedures we have already successfully field validated in C3 crops (Zhu et al. 2007; Srinivasan et al., 2017).

References

1. Srinivasan, V., Kumar, P., & Long, S. P. (2017) Decreasing, not increasing, leaf area will raise crop yields under global atmospheric change. Global change biology, 23(4), 1626-1635.
2. Wang, Y., Bräutigam, A., Weber, A. P., & Zhu, X. G. (2014) Three distinct biochemical subtypes of C4 photosynthesis? A modelling analysis. Journal of experimental botany, 65(13), 3567-3578. 2.
3. Wang, Y., Long, S. P., & Zhu, X. G. (2014) Elements Required for an Efficient NADP-ME Type C4 Photosynthesis. Plant physiology, pp-113.
4. Zhu, X. G., de Sturler, E., & Long, S. P. (2007) Optimizing the distribution of resources between enzymes of carbon metabolism can dramatically increase photosynthetic rate: a numerical simulation using an evolutionary algorithm. Plant physiology, 145(2), 513-526.
5. Zhu, X. G., Long, S. P., & Ort, D. R. (2008) What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? Current opinion in biotechnology, 19(2), 153-159.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

Lead PI: Evan DeLucia
PI Organization: University of Illinois at Urbana-Champaign
Program manager: Kent Peters
Presenter Organization: Lawrence Berkeley National Laboratory
Poster Abstract Funding Area: Bioenergy Research Center
Abstract file naming convention: DeLucia_**Yoshikuni**.pdf

Title: Population Genomics of *Issatchenkia orientalis* Reveals the Genetic Mechanisms of its Multi-Stress Tolerance

Ping-Hung Hsieh[1], Zhiying Zhao[2], Yusuke Sasaki[2], **Yasuo Yoshikuni (yyoshikuni@lbl.gov)[1,2,3]**

[1]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory
[2]US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory
[3]Biological Systems Engineering Division, Lawrence Berkeley National Laboratory
[4]University of Illinois at Urbana-Champaign

Project Goals: We aim to use population genomics analysis to understand how multi-stress tolerance evolved in *Issatchenkia orientalis* so that this non-model yeast may be further developed for biotechnology applications.

Issatchenkia orientalis is an emerging non-model ascomycetes yeast that has unparalleled ability to tolerate multiple stresses, including low pH and high concentrations of lignocellulosic inhibitors. Because of these unique characteristics, CABBI considers *I. orientalis* one of its flagship yeast species and is exploring development of *I. orientalis* as a chassis for production of bioproducts from lignocellulosic biomass. To understand how this species evolved to acquire multi-stress tolerance ability, we are performing population genomics analyses. First, we collected 160 *I. orientalis* strains globally and re-sequenced them using the sequencing pipeline at the JGI. Subsequent sequence analyses identified an initial set of 298,334 SNPs and 24,414 InDels. Our pilot analysis of 30 strains indicated that 21% of these strains are triploids and 79% are diploids. CNVs of large chromosome fragments (>80kb) were shown in 4 of the 30 strains. Loss of heterozygosity (LOH) events were detected among some strains. The next step of our project is to perform GWAS. We expect that GWAS will help us identify functionally important genetic variants and provide insights into how to further develop *I. orientalis* strains more suitable for biotechnological applications.

Funding statement: *This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.*

Plant-Microbe Interfaces: From gene discoveries to molecular, genetic and biochemical validations in *Populus*

Jay Chen[1*] (chenj@ornl.gov), Wellington Muchero,[1] David J. Weston,[1] Xiaohan Yang,[1] Udaya C. Kalluri,[1] Sara S. Jawdy,[1] Lee E. Gunter,[1] Jessy L. Labbé,[1] Gerald A. Tuskan,[1] Dale A. Pelletier,[1] Timothy J. Tschaplinski,[1] and **Mitchel J. Doktycz**[1]

[1]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

http://PMI.ornl.gov

Project Goals: The goal of the Plant-Microbe Interfaces (PMI) SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. *Populus* and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic *Populus*-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

One of the primary objectives of PMI research is to identify genetic loci regulating *Populus*-microbial interactions and to dissect the signals and pathways responsible for initiating and maintaining beneficial relationships between the host and its associated microbes. In order to accomplish these goals, we have developed and leveraged two key genomic and genetic resources including ~1,000 *P. trichocarpa* natural variants and ~400 *P. trichocarpa* × *P. deltoides* pseudo-backcross pedigree, which have enabled Genome-Wide Association Studies (GWAS) and Quantitative Trait Locus (QTL) mapping, respectively. By taking these approaches, we have identified a number of genetic loci associated with *Populus*-microbial interactions. Validating the function of these genetic loci represents a rate-limiting step in elucidating the signaling cascades from the perception of microbial signal to the manifestation of biological responses. In order to address these challenges, we have developed two key systems including the *Populus* protoplast transient expression system and the *Populus* hairy root transient expression system to enable robust molecular, cellular and biochemical validation. To further validate the function of targeted genes at the organismal level, we apply both a heterologous expression system using the model plant *Arabidopsis* for rapid generation of transgenic plants and a *Populus* transformation system for generating stable transgenic lines over- or under-expressing the gene of interest. These genetically modified materials and *P. trichocarpa* natural variants with high-impact single nucleotide polymorphisms (SNPs) (*i.e.*, resulting in opening read frame shift, deletion or insertion) are subject to microbial inoculation, microscopic examination and colonization analysis to validate the function of selected genes at the physiological level. Furthermore, transcriptomics, metabolomics and proteomics are used to reveal global changes at the transcript, metabolite and protein levels, respectively. Taken together, this is a general strategy to identify genetic loci regulating *Populus*-microbial interactions and to functionally validate these loci. We continue to explore other efficient means to accelerate our discoveries and functional validations.

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.

Plant-Microbe Interfaces: Detection and characterization of signaling polypeptides in plant-microbial systems using high-performance tandem mass spectrometry

Robert Hettich[1*] **(hettichrl@ornl.gov),** Paul E. Abraham,[1] and Richard J. Giannone[1]

[1]Chemical Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

http://PMI.ornl.gov

Project Goals: The goal of the Plant-Microbe Interfaces (PMI) SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. *Populus* and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic *Populus*-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

Small polypeptides are emerging as key signaling molecules that mediate cell-cell communication and various biological processes that occur within and between plants and microbes. These polypeptides can be secreted or non-secreted and are generally categorized into two classes: 1) small polypeptides (~5 to 20 amino acids in length) that are post-translationally cleaved from larger precursor proteins and 2) gene-encoded small polypeptides (i.e., small proteins that comprise of ~40 to 100 amino acids).

The discovery and characterization of these polypeptides is challenging because their post-translational processing is not well understood nor predictable. Many polypeptides can harbor post-translational modifications (PTMs) such as tyrosine sulfation, proline hydroxylation, hydroxyproline, and arabinosylation. In some cases, these polypeptides can form cyclic structures via disulfide or isopeptide bonds. Adding to this complexity, previous studies have shown that small open reading frames (sORFs) embedded in larger ORFs can encode additional sources of functional polypeptides. As such, gene expression levels provide little, if any, value in the characterization of these small polypeptides. Instead, high performance liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) is an effective means to identify and characterize these endogenous peptide molecules.

A traditional peptide sequencing measurement (LC-MS/MS) benefits greatly by having prior knowledge of the potential sequences. However, given the extent to which these small polypeptides can be derived and post-translationally modified, traditional bioinformatic approaches have limited applicability for their identification. Because these small polypeptides are a critical signaling component in plant-microbe interactions, there is an increasing need for better methods to facilitate their detection and identification. To this end, we have developed new MS measurement approaches and analysis platforms for *de-novo* sequencing of these small

polypeptides using high-performance LC-MS/MS instrumentation (ThermoFisher Q-Exactive Orbitrap mass spectrometry).

To effectively measure these polypeptides, we optimized a sample preparation protocol (molecular cutoff filtering) for a customized hybrid MS approach that integrates full mass spectral scans (for global measurements of the full set of sample components) with targeted scans that focus on m/z regions encompassing multiple forms of the specific polypeptides of interest. By employing HCD fragmentation methods in a QE-MS instrument, we found that most cyclic peptides usually can be fragmented in sufficient detail to enable sequence tag identifications that can verify suspected amino acid linkages. In particular, we utilized a software approach entitled DirectTag to identify sequence tags in these tandem mass spectra, but are examining more sophisticated approaches that are designed to characterize unusual and cyclic peptide forms.

Recently, it has been shown that plants harbor genes encoding signaling polypeptides that can be expressed in one organ, for instance the shoot, only to be transported to another organ, like the root. Therefore, we evaluated our MS approach to identify novel polypeptide species across several plant sources (i.e., leaf, xylem sap, and root). Overall, using this new approach, we were able to identify 23,572 polypeptide sequence tags (with lengths of at least 7 amino acids) that serve as a foundation to identify the full-length polypeptide species and proteins of origin. Hierarchical clustering revealed distinct groupings of these polypeptides within root, leaf, and stem tissues. We have noted that several of these do not match *de-novo* sequence annotations, indicating the presence of unknown post-translational modifications, and thus have begun more extensive characterizations.

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.

Plant-Microbe Interfaces: Integrated omics and phytobiome analysis of the adaptation of *Populus deltoides* to acute progressive drought and cyclic drought

Daniel Jacobson[1*] (jacobsonda@ornl.gov) Benjamin Garcia,[1] Jessy Labbé,[1] Piet Jones,[1,2] Paul Abraham,[1] Ian Hodge,[3] Sharlee Climer,[4] Sara Jawdy,[1] Lee Gunter,[1] Gerald Tuskan,[1] Xiaohan Yang,[1] Timothy Tschaplinski,[1] and **Mitchel Doktycz**[1]

[1] Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [2] University of Tennessee Knoxville, Knoxville, TN; [3] Stanford University, Palo Alto, CA; [4]Department of Mathematics and Computer Science, University of Missouri – St. Louis, St. Louis, MO

Project Goals: The goal of the PMI SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. *Populus* and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic *Populus*-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

Plant drought stress causes systematic changes to photosynthesis, metabolism, growth, and potentially the phytobiome. Additionally, drought affects plants in both a species-specific and water-deficit-driven manner, causing the response to drought to be dependent both on how drought is being experienced and on any adaptation to prior drought exposure. Thus, understanding the effect of drought on plants requires assessing drought response in multiple conditions, such as progressive acute drought and recurrent cyclic drought, and at different levels of severity. In this study, we have utilized RNA sequencing, GC-MS and LC-MS in order to identify changes to the plant transcriptome, metabolome, proteome and the phytobiome during both acute progressive drought and cyclic drought at multiple severities. Co-analysis of these omic layers with the phytobiome, allows for the identification of novel associations that would not be possible otherwise. We have identified that the drought response ranges from increased transcripts related to photosynthesis and metabolic activity in mild acute drought to decreased transcripts related to photosynthesis and metabolic impairment in severe drought. Moreover, while water deficit is a main driver of transcriptional responses in severe drought, there are increases in reactive oxygen species (ROS) metabolism and photosynthetic transcripts in cyclic severe drought compared to acute severe drought, independent of water deficit. The phytobiome exhibits alternate responses to drought when compared to the transcriptome. Specifically, the phytobiome is affected more by the cyclic or acute nature of the drought rather than the severity of the drought, with the phytobiome having an increase in taxa under cyclic drought that are often reported to have beneficial effects on the plants. Lastly, we have identified associations between taxa in the phytobiome with expression of disease response, ROS metabolism, and photosynthesis transcripts suggesting interplay between the host plant and its phytobiome in response to drought.

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.

Plant-Microbe Interfaces: Assembly and function of the ectomycorrhizome of *Populus trichocarpa*

Brian Looney[1]*(brian.looney@duke.edu), Alejandro Rojas-Flechas,[2] Khalid Hameed,[1] Christopher Schadt,[3] Francis Martin,[4] Jessy Labbé,[3] and **Rytas Vilgalys[1]**

[1]Department of Biology, Duke University, Durham, NC; [2]Department of Plant Pathology, University of Arkansas, Fayetteville, AR; [3]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [4]INRA and the Université de Lorraine, Labex ARBRE, UMR 1136, Champenoux, France

http://PMI.ornl.gov

Project Goals: The goal of the Plant-Microbe Interfaces (PMI) SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. *Populus* and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic *Populus*-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

The ectomycorrhizal fungi (EMF) are mutualistic associates of some of the most dominant and speciose trees on the planet, including pines, oaks, and eucalypts. Typically, an ectomycorrhizal plant species will play host to hundreds of different EMF species, however, there are cases where a restricted community of EMF can be found with specific host species. *Populus trichocarpa* is an unusual ectomycorrhizal host in that it harbors a root microbial community of diverse symbiotic guilds, including arbuscular mycorrhizae, EMF, rhizobia, and endophytic fungi. One aim of the PMI project is to characterize the fungal communities associated with *P. trichocarpa* to begin to assess the functional roles of different components in the root microbiome and how they impact plant health and resistance to pathogens. Field surveys of natural *P. trichocarpa* communities in the Pacific Northwest have yielded a specimen and culture collection of the EMF associates of *P. trichocarpa*. This EMF community is mostly restricted to a dozen or so core species including several that are host specific with *Populus*. Our objective is to sequence the genomes of this "core community" to test whether these core members are functionally partitioned through coevolution of their host or whether they are functionally redundant. Using established *in vitro* and *in planta* inoculation and mycorrhization systems, our goal is to reassemble the core EMF community to assess assembly and functional redundancy of core EMF communities.

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.

Plant-Microbe Interfaces: Measuring rates and consequences of horizontal gene transfer in the *Populus* rhizosphere

Joshua K. Michener* (michenerjk@ornl.gov), Dawn Klingeman, Dale A. Pelletier, and **Mitchel J. Doktycz**

Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

http://PMI.ornl.gov

Project Goals: The goal of the Plant-Microbe Interfaces (PMI) SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. *Populus* and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic *Populus*-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

Microbes have larger population sizes and shorter generation times than trees. As a result, the plant-microbe interaction is constantly being challenged by microbial mutants with new phenotypes. We are investigating how these mutually-beneficial interactions are maintained despite evolutionary challenges. In particular, horizontal gene transfer (HGT) can rapidly move genes and pathways between physically-adjacent microbes. When host-microbe recognition is mediated by specific molecular signals, those interactions are vulnerable to hijacking by unrelated microbes that acquire the necessary pathway through HGT.

We are developing sequencing-based techniques for tracking HGT in microbial communities, using next-generation capture-C to associate plasmids and hosts in a function-independent fashion. This approach is capable of differentiating strains with variable plasmid occupancy, though further optimization is ongoing to decrease the fractional detection limit.

In addition, we are identifying gene clusters in sequenced rhizosphere isolates and transferring them into naïve hosts. Focusing initially on gene clusters involved in degradation of salicylates, plant defense compounds common in *Populus*, we are genetically modifying non-degrading strains to introduce new aromatic degradation pathways. These modified strains will then be tested for colonization efficiency to determine whether pathway acquisition affects root colonization.

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.

Plant-Microbe Interfaces: Investigating microbial recruitment and colonization dynamics in the poplar microbiome

Jennifer Morrell-Falvey[1,2*](morrelljl1@ornl.gov), Amber Bible,[2] Yunzi Li,[1] Jayde Aufrecht,[3] Scott Retterer,[1,3] and **Mitchel Doktycz**[1,3]

[1]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [2]Department of Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville TN; [3]Bredesen Center, University of Tennessee, Knoxville, TN

http://PMI.ornl.gov

Project Goals: The goal of the PMI SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. *Populus* and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic *Populus*-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

The organization and behavior of plant-microbe communities are dictated by a complex network of physical and chemical interactions between organisms within the rhizosphere. These interactions produce spatially and temporally localized micro-environments that change as organisms grow, consume and produce materials and alter their surroundings. The goal of this research is to understand how the chemical environment of the poplar rhizosphere influences microbial recruitment and community organization. Previous results have shown that bacteria colonize the plant host to differing extents, ranging in abundance by orders of magnitude (10^8-10^3 cfu/gram root). To better understand these patterns, we are utilizing methods to image and measure the dynamics of microbial colonization along plant roots with increasing spatial and temporal resolution. The results show that some bacterial species reproducibly colonize specific regions of the plant root, suggesting that the bacteria are sensing and responding to physical and chemical cues from the host plant. We are currently performing screening experiments using capillary assays and microfluidic habitats to determine how specific chemicals influence growth and chemotactic behaviors based on the hypothesis that some bacteria are recruited to the host plant via the secretion of compounds that act as chemoattactants. By correlating these data with visualization of microbial colonization dynamics using imaging chambers and with data from ongoing efforts to sample the chemical environment of the rhizosphere, we hope to gain important insights into how the microbiome is recruited and organized.

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.

Plant-Microbe Interfaces: Transcriptional co-regulation of *MYB46* and *WRKY33* by ANGUSTIFOLIA modulates *Arabidopsis* resistance towards biotrophic and necrotrophic pathogens

Meng Xie,[1,2,3] **Wellington Muchero**[1,2]* (mucherow@ornl.gov), Jin Zhang,[1,2] Anthony C. Bryan,[1] Yunqiao Pu,[1,2] Jessy Labbé,[1,2] Dale A. Pelletier,[1] Nancy Engle,[1,2] Jennifer L. Morrell-Falvey,[1] Jeremy Schmutz,[4,7] Arthur J. Ragauskas,[1,2,5,6] Timothy J. Tschaplinski,[1,2] Gerald A. Tuskan,[1,2] and **Jin-Gui Chen**[1,2]

[1]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; [3]Department of Plant Sciences, University of Tennessee, Knoxville, TN; [4]U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA; [5]UT-ORNL Joint Institute for Biological Science, Oak Ridge National Laboratory, Oak Ridge, TN; [6]Department of Chemical and Biomolecular Engineering & Department of Forestry, Wildlife, and Fisheries, University of Tennessee, Knoxville, TN and [7]HudsonAlpha Institute for Biotechnology, Huntsville, AL

http://PMI.ornl.gov

Project Goals: The goal of the Plant-Microbe Interfaces (PMI) SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. *Populus* and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic *Populus*-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

ANGUSTIFOLIA (AN) is involved in the development of numerous plant organs. The primary role of the animal homolog of AN, C-TERMINAL BINDING PROTEIN (CtBP), is the transcriptional co-repression of tumor suppressor and pro-apoptotic genes (1,2). However, the nuclear function of AN remains unstudied in plants. Here, we found that AN can accumulate in the nucleus and functions as a transcriptional repressor. By interacting with another nuclear protein TYROSYL-DNA PHOSPHODIESTERASE1 (TDP1), AN imposes transcriptional repression on *MYB46* which encodes a key transcription factor regulating the phenylpropanoid biosynthesis pathway (3), while releasing the TDP1-imposed transcriptional repression on *WRKY33* which encodes a critical transcription factor regulating the ethylene/jasmonic acid (ET/JA) signaling pathway (4,5). Consistent with these molecular results, genetic analyses on transgenic *Arabidopsis* plants showed that AN is capable of regulating the expressions of *WRKY33*, *MYB46*, as well as their downstream genes involved in salicylic acid (SA) and ET/JA signaling pathways. Meanwhile, plant defense capability against biotrophic and necrotrophic

pathogen infection was altered by AN. Collectively, these findings indicate that the transcriptional co-regulation of *MYB46* and *WRKY33* by AN may play an important role in the integration of SA and ET/JA signaling, as well as defenses against biotrophic and necrotrophic pathogens.

References

1. Stankiewicz TR, Gray JJ, Winter AN, & Linseman DA (2014) C-terminal binding proteins: central players in development and disease. Biomol Concepts 5(6):489-511.
2. Chinnadurai G (2009) The transcriptional corepressor CtBP: a foe of multiple tumor suppressors. Cancer Res 69(3):731-734.
3. Xie M, Zhang J, Tschaplinski TJ, Tuskan GA, Chen JG, & Muchero W (2018) Regulation of lignin biosynthesis and its role in growth-defense tradeoffs. Frontiers in plant science 9: 1427.
4. Birkenbihl RP, Diezel C, & Somssich IE (2012) Arabidopsis WRKY33 is a key transcriptional regulator of hormonal and metabolic responses towards Botrytis cinerea infection. Plant physiology:pp. 111.192641.
5. Liu S, Kracher B, Ziegler J, Birkenbihl RP, & Somssich IE (2015) Negative regulation of ABA signaling by WRKY33 is critical for Arabidopsis immunity towards Botrytis cinerea 2100. Elife 4:e07295.

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.

Plant-Microbe Interfaces: Cultivating a representative *Populus* microbiome?

Dale A. Pelletier* (pelletierda@ornl.gov), Dana L. Carper, David J. Weston, Tse-Yuan S. Lu Aditya Barde, Collin M. Timm, Leah H. Burdick, Sara S. Jawdy, **Mitchel J. Doktycz**

[1]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [2]Graduate School of Genome Science and Technology, University of TN, Knoxville, TN

http://pmiweb.ornl.gov/

Project Goals: The goal of the PMI SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. *Populus* and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic *Populus*-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

Environmental samples of *Populus* tissues harbor a diverse microbiome. Many of the functional attributes of *Populus* are a result of both its genetic potential and its microbiome. The integral role of the microbial community in plant performance is now widely recognized, and increasingly, the constituents of plant microbiomes are being defined. While phylogenetic surveys have revealed who is present in the *Populus* microbiome, and show that this composition can depend on, and respond to, environmental conditions and stresses, the challenge shifts to determining why particular microbes are selected and how they collectively function in concert with their host.

To dissect and understand the complexity of natural plant and microbial communities, we are developing experimental plant-microbial community systems amenable to laboratory conditions. We targeted the isolation of representative bacterial strains from environmental samples of *Populus* roots using a direct plating approach. This resulted in 3205 unique bacterial isolates from 4 phyla, 9 classes, 16 orders, 46 families and 118 genera. All isolates have been identified with 16S rRNA sequencing, with a portion (14%) having genome sequences. The greatest number of isolates are from 5 genera; *Pseudomonas* (18.2%), *Bacillus* (13.7%), *Rhizobium* (13.4%) *Streptomyces* (11.8%) and *Variovorax* (5.9%). The majority of the isolates were cultured from the root endosphere (50%) followed by rhizosphere (37.3%) and unsterilized root tissues (12.7%) of *Populus*. The representativeness of bacterial isolates has been assessed by comparison of 16S sequences to 16S amplicon sequences from environmental surveys. Many of the bacterial strains represent relatively abundant genotypes present within the *Populus* rhizosphere and endosphere. While this culture collection represents a large fraction of the abundant phyla present within the *Populus* rhizosphere, efforts are underway to isolate additional phyla. Additionally, with the initiation of a JGI-CSP, genome sequencing has been carried out on representatives from all 4 phyla and 9 classes, encompassing 12 orders, 34 families and 84 genera. These isolates and genome sequences have facilitated comparative genomic analyses and experimental plant-microbial community approaches to understand the dynamics of microbiome organization. Reference microbiomes comprised of low diversity microbial communities representing abundant taxa from

environmental samples were created and tested for their ability to reproducibly colonize *Populus*. We found that colonization is reproducible across replicates and host plants, despite isolation from different host genotypes. Community member abundances in general are similar to single strain inoculation indicating potential for unique niche site colonization. Communities increased root growth compared to non-inoculated host plants. Further, these studies are beginning to uncover characteristics of host-microbe selectivity and identify molecular mechanisms underlying community assembly and function. This work allows us to determine how individual microbes contribute in a community and enables a mechanistic understanding of how plant and microbial genetics lead to complex community phenotypes.

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.

Plant-Microbe Interfaces: Accessing the hidden diversity of the poplar microbiota through targeted metagenomics and cultivation.

Mircea Podar* (podarm@ornl.gov), Joel Turner, Leah Burdick, Dawn Klingeman, Zamin Yang, Dale Pelletier and **Mitchel Doktycz**

Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

http://PMI.ornl.gov

Project Goals: The goal of the Plant-Microbe Interfaces (PMI) SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. Populus and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic Populus-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

The bulk of the microbial taxonomic diversity from soils and the *Populus* rhizosphere has remained uncultured. Currently, we only have representatives of 4 of the 21 phyla with abundance of >0.1% of the rhizosphere community. Competition and inhibition could restrict the growth of some organisms in the presence of others. It is also likely that many of the uncultured bacterial taxa require chemical compounds synthesized either by the plant or by other members of the community. A comprehensive understanding of the role of microbes in *Populus'* physiology requires expanding the range of bacterial taxa used in the assembly of controlled communities used for laboratory mechanistic studies. For yet-unculturable microbes, obtaining genomic information will aid inferring their physiological potential and dependencies, which in turn could lead to cultivation.

We developed flow cytometric fractionation of the complex *Populus* microbiome, based on various cellular characteristics including size and morphology, DNA content and cell wall constituents. Following sorting of populations containing 1000-5000 cells and genomic amplification by multiple displacement, we obtained mini-metagenomes enriched in low abundance microbial constituents of *Populus* rhizosphere, including Thaumarchaeota and candidate bacterial phyla OD1, TM6 and TM7. Metagenomic sequencing resulted in draft genomes that are currently being used in comparative genomic analyses with relatives from other microbiomes.

We also applied single cell isolation and high throughput cultivation using flow cytometry cell sorting. We were successful in isolating in pure culture a member of Acidobacteria (*Terriglobus* sp.) and one of Verrucomicrobia, (*Roseimicrobium* sp), the first representatives of those phyla from the poplar rhizosphere. We also isolated novel representatives of Actinobacteria, Proteobacteria and Bacteroidetes. Those organisms are currently being characterized physiologically and at genome level. Ongoing effort is also directed towards isolation of candidate bacterial phyla representatives.

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

Plant-Microbe Interfaces: Effects of variation in host secondary chemistry across *Populus* genotypes on the composition of the rhizosphere microbiome

Allison M. Veach[1], Reese Morris[1], Daniel Z. Yip[1], Zamin K. Yang[1], Nancy L. Engle[1], Melissa A. Cregger[1], Timothy J. Tschaplinski[1], **Christopher W. Schadt[1,2*](schadtcw@ornl.gov)**

[1]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [2]Department of Microbiology, University of Tennessee, Knoxville, TN

http://pmi.ornl.gov

Project Goals: The goal of the Plant-Microbe Interfaces (PMI) SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. *Populus* and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic *Populus*-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

Abstract

Plants have developed defense strategies for phytopathogen and herbivore protection via coordinated metabolic mechanisms. Low-molecular weight metabolites produced within plant tissues, such as salicylic acid, represent one such mechanism which likely mediates plant – microbe interactions above and below ground. Salicylic acid is a ubiquitous phytohormone at low levels in most plants, yet are concentrated defense compounds in *Populus*, likely acting as a selective filter for rhizosphere microbiomes. We propagated twelve *Populus trichocarpa* genotypes which varied an order of magnitude in salicylic acid (SA)-related secondary metabolites, in contrasting soils from two different origins. After four months of growth, plant properties (leaf growth, chlorophyll content, and net photosynthetic rate) and plant root metabolomics specifically targeting SA metabolites were measured via GC-MS. In addition, rhizosphere microbiome composition was measured via Illumina MiSeq sequencing of 16S and ITS2 rRNA-genes.

Soil origin was the primary filter causing divergence in archaeal/bacterial and fungal communities with plant genotype secondarily influential. Both archaeal/bacterial and fungal evenness varied between soil origin and with at least one SA metabolite (archaea/bacteria: populin; fungi: salicin and salicylic ACID). The production of individual salicylic acid derivatives - tremuloidin, populin, salicortin - that varied by host genotype, resulted in compositional differences for archaea/bacteria within a soil origin. Contrastingly, overall salicylic acid levels and its major derivative, salicortin, were associated with shifts in fungal

community composition, but to a lesser degree. Using regression-based analyses, it was revealed that changes in several dominant bacterial operational taxonomic units (OTUs) were drivers of these metabolite-community relationships; and most of these bacterial OTUs exhibited positive correlations with SA derivatives. No fungal OTUs were detected which significantly varied with metabolites regardless of overall compositional differences.

These results indicate microbial communities diverge most among soil origin. However, within a soil origin, bacterial communities are responsive to plant SA production, particularly populin, within greenhouse-based rhizosphere microbiomes. Fungal microbiomes are impacted by root SA-metabolites, but to a lesser degree within this experimental context. These results suggest plant defense strategies, such as SA and its secondary metabolites, may partially drive both archaeal/bacterial and fungal taxa-specific colonization and assembly.

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.

Plant-Microbe Interfaces: Systematic discovery and prediction of novel LuxI-type quorum sensing signals in members of the *Populus* microbiome

Amy L. Schaefer[1]* (amyschae@uw.edu), Omar Demerdash,[2] Dale A. Pelletier,[2] **Caroline S. Harwood**,[1] **E. Peter Greenberg**,[1] and **Mitchel J. Doktycz**[2]

[1]Department of Microbiology, University of Washington, Seattle; [2]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

http://PMI.ornl.gov

Project Goals: The goal of the Plant-Microbe Interfaces (PMI) SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. *Populus* and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic *Populus*-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

Acyl-homoserine lactone (acyl-HSL) quorum sensing (QS) has received considerable interest as a possible target for controlling microbial activities and as a model for communication system involved in coordinating activities of groups of bacteria. Recent advances have shown there are two related families of acyl-homoserine lactone quorum-sensing signal synthesis enzymes (LuxI-type enzymes), which differ in whether they acquire the organic acid reaction substrate as an acyl-acyl carrier protein (ACP) intermediate or an acyl-coenzyme A (CoA) intermediate. The acyl-CoA-utilizing enzymes are of particular interest because one has served as a model to understand the reaction mechanism of acyl-homoserine lactone synthases, and because there might be a great diversity in signals synthesized by this family of enzymes. Using a bioinformatics approach, we found the CoA class of LuxI homologs is common in genomes of α-Proteobacteria isolated from *Populus* tree roots. To systematically study acyl-HSL diversity among these isolates, we developed an experimental pipeline that includes the following steps: i) determination of the potential CoA substrate inventory encoded in a given genome; ii) utilization of a radiolabel ^{14}C-methionine protocol to detect acyl-HSLs synthesized in the presence of potential CoA substrates; iii) AiiA lactonase treatment of the ^{14}C-product to confirm it is an acyl-HSL compound; and iv) purification and structural identification of the acyl-HSL using high-resolution mass spectrometry. There is also an opportunity to leverage these CoA-type LuxI enzymes for computational based structural analysis and molecular docking experiments designed to predict the substrates for these enzymes. Because the potential CoA-substrates are 'simple' (relative to the ACP-linked substrates), we can create a tailored library of potential CoA substrates to dock with a predicted protein model of a given LuxI-type synthase. The experimental protocol will validate the computational results and aid in prioritizing signaling

partners for further evaluation. Taken together, these approaches should be useful in expanding the range of acyl-HSL signal diversity among plant-associated bacteria.

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.

Plant-Microbe Interfaces: The nature of the progression of drought stress drives differential metabolomic responses in *Populus deltoides*

Timothy J. Tschaplinski[1*] (tschaplinstj@ornl.gov), Paul Abraham[2], Sara Jawdy[1], Lee Gunter[1], Madhavi Martin[1], Nancy Engle[1], Xiaohan Yang[1], and Gerald Tuskan[1]

[1]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [2]Chemical Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

http://PMI.ornl.gov

Project Goals: The goal of the Plant-Microbe Interfaces (PMI) SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. *Populus* and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic *Populus*-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

Use of woody crops for Quad-level (10^{15} BTU) energy production will require require use of marginal agricultural lands, where periods of water stress are frequent, especially given the predictions of increased frequency and severity of droughts associated with the predicted global climate change. Our previous research demonstrated that some *Populus* sp. genotypes have the capacity to increase dehydration tolerance by lowering the osmotic potential via osmotic adjustment, the active accumulation of solutes under stress, which allows turgor and growth maintenance under mild to moderate stress and facilitates growth recovery after stress relief. Despite the large number of drought stress studies that have been published, few studies have contrasted the degree and nature of solute accumulation if the nature of drought stress progression is varied (e.g., cyclic vs acute, short-term vs long term, moderate vs severe stress). The aim of the study was to determine how the inherent genetic potential of a given clone interacts with the nature of the stress experienced to determine the degree of the biochemical response.

A drought stress study on *Populus deltoides* 'WV94' was conducted in a greenhouse and the resulting metabolomic profiles of leaves were determined for plants subjected to cyclic mild (-0.5 MPa predawn leaf water potential) drought vs cyclic severe (-1.26 MPa) drought after 2 or 4 drought cycles in contrast to well-watered controls (-0.1 MPa), and in contrast with plants subjected to acute drought, where plants were not rewatered, but were allowed to dessicate for up to 8 days. Leaves were rapidly sampled, fast frozen on dry ice, ground with liquid N and then twice extracted with aqueous ethanol (80%). Dried aliquots of extracts were analyzed for metabolites by gas chromatography-mass spectrometry (GCMS) with electron impact ionization (70 eV) following trimethylsilylation.

The nature of drought onset (cyclic vs acute), frequency of drought (number of cycles), and the severity of drought (mild vs severe), all interact to dictate the degree of osmotic adjustment and the nature of the

organic solutes that accumulate. Acute onset of prolonged, severe drought induced the greatest osmotic adjustment after withholding water for 7 days (1.42x) with the greatest accumulations in the large, complex higher-order salicylate conjugates. Organic solute accumulation under cyclic stress relative to well-watered controls was moderate (1.20x) and was largely constituted by soluble sugars, organic acids, and amino acids. In contrast, acute onset of prolonged drought induced the greatest osmotic adjustment and the largest responses in secondary metabolism, with hydroxycinnamic acid conjugates of salicin; the populosides, playing a key role in drought tolerance of *P. deltoides* by lowering osmotic potential and increasing the likelihood of recovery following stress alleviation. Current studies are attempting to integrate these metabolomic responses with proteomic and transcriptomic responses measured on the same plants to obtain a more complete understanding of drought tolerance in *P. deltoides* and how these responses affect plant-microbe interactions.

Publications

Tschaplinski, T.J., S.S. Jawdy, L.E. Gunter, M.Z. Martin, N.L. Engle, X. Yang, and G.A. Tuskan. 2019. The nature of the progression of drought stress drives differential metabolomic responses in *Populus deltoides*. Ann. Bot. (in press) https://doi.org/10.1093/aob/mcz002

Garcia, B.J., J.L. Labbé, P. Jones, P.E. Abraham, I. Hodge, S. Climer S.S. Jawdy, L.E. Gunter, G.A. Tuskan, X. Yang, T.J. Tschaplinski, and D.A. Jacobson. 2018. Phytobiome and transcriptional adaptation of Populus deltoides to acute progressive drought and cyclic drought. Phytobiome J. 4 December 2018 https://doi.org/10.1094/PBIOMES-04-18-0021-R.

Abraham, P.E., B. Garcia, L.E. Gunter, S.S. Jawdy, D.A. Jacobson, N.L. Engle, X. Yang, R.L. Hettich, G.A. Tuskan, and T.J. Tschaplinski. 2018. Quantitative proteome profile of water-deficit stress responses in *Populus deltoides* leaves. Plos One 13(2): e0190019. https://doi.org/10.1371/journal.pone.0190019

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.

Plant-Microbe Interfaces: Functional characterization of small secreted proteins in poplar in relation to P*opulus-Laccaria* symbiosis

Xiaohan Yang* (yangx@ornl.gov), Hua Cheng, Linling Li, Rongbin Hu, Degao Liu, Jennifer Morrell-Falvey, Jessy Labbé, and **Mitchel Doktycz**

Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

http://PMI.ornl.gov

Project Goals: The goal of the Plant-Microbe Interfaces (PMI) SFA is to characterize and interpret the physical, molecular, and chemical interfaces between plants and microbes and determine their functional roles in biological and environmental systems. *Populus* and its associated microbial community serve as the experimental system for understanding the dynamic exchange of energy, information, and materials across this interface and its expression as functional properties at diverse spatial and temporal scales. To achieve this goal, we focus on 1) defining the bidirectional progression of molecular and cellular events involved in selecting and maintaining specific, mutualistic *Populus*-microbe interfaces, 2) defining the chemical environment and molecular signals that influence community structure and function, and 3) understanding the dynamic relationship and extrinsic stressors that shape microbiome composition and affect host performance.

Symbiosis is a mutually beneficial interaction between two or more organisms. Many fungal lineages have evolved elaborate protein-based signals to influence their hosts in order to support their metabolic requirements during symbiosis (Martin and Kamoun, 2011). Poplar (i.e., *Populus* species) plants are fast-growing trees with great potential for biofuel production and ecosystems service (Yang et al., 2009). Mycorrhizal colonization of poplar roots by ectomycorrhizal fungus *Laccaria bicolor* has a positive impact on the overall tree health and growth. A deep understanding of the molecular mechanism underlying poplar-*Laccaria* interaction will have important implications for increasing bioenergy and ecosystem sustainability.

Recently, we performed RNA sequencing of *P. trichocarpa* roots in symbiosis with *L. bicolor* and identified 417 plant-encoded putative small secreted proteins (SSPs) that were significantly regulated during this poplar-*Laccaria* interaction and four of the five poplar SSPs tested in an *in-vitro* feeding experiment can enter *L. bicolor* nuclei (Plett et al., 2017). Based on this preliminary work, the present study investigates the function of poplar SSPs using both computational and experimental approaches, with a focus on the candidate poplar SSPs that have potential of regulating gene expression in the fungus. Specifically, through computational analysis of DNA-binding potential and secretion signals in the 417 poplar SSPs, we predicted 14 secreted DNA-binding SSPs. Furthermore, for gain-of-function analysis, the poplar genes encoding 12 of these secreted DNA-binding SSPs were over-expressed in multiple plant species (*i.e.*, poplar, *Arabidopsis*, tobacco) through *Agrobacterium*-mediated plant stable transformation. Also, for loss-of-function analysis, two poplar SSP genes were knocked-out using CRISPR/Cas9

technology (Liu et al., 2016). Our preliminary analysis demonstrated that the over-expression in poplar of one of these SSPs (PtSSP1), greatly increased the symbiosis between poplar and *L. bicolor*, along with the PtSSP1 protein moving from the transgenic plants into the nuclei of fungal hyphae. The transgenic poplar, *Arabidopsis* and tobacco plants will be characterized in various aspects such as protein movement from plant to fungal cells, regulation of fungal gene expression, and effect of over-expression/knockout on symbiosis. In summary, we are identifying and characterizing candidate poplar SSPs that could potentially regulate fungal gene expression to promote symbiosis.

References

Liu, D., Hu, R., Palla, K.J., Tuskan, G.A. and Yang, X. (2016) Advances and perspectives on the use of CRISPR/Cas9 systems in plant genomics research. *Current Opinion in Plant Biology* **30**, 70-77.

Martin, F. and Kamoun, S. (2011) *Effectors in plant-microbe interactions*:John Wiley & Sons.

Plett, J.M., Yin, H., Mewalal, R., Hu, R., Li, T., Ranjan, P., Jawdy, S., De Paoli, H.C., Butler, G., Burch-Smith, T.M., Guo, H.-B., Ju Chen, C., Kohler, A., Anderson, I.C., Labbé, J.L., Martin, F., Tuskan, G.A. and Yang, X. (2017) *Populus trichocarpa* encodes small, effector-like secreted proteins that are highly induced during mutualistic symbiosis. *Scientific Reports* **7**, 382.

Yang, X., Kalluri, U.C., DiFazio, S.P., Wullschleger, S.D., Tschaplinski, T.J., Cheng, M.Z.-M. and Tuskan, G.A. (2009) Poplar genomics: State of the science. *Critical Reviews in Plant Sciences* **28**, 285-308.

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

Approaches to Increasing Terpenoid Production in *Zymomonas mobilis* by Improving Activity of IspG and IspH of the MEP Pathway

Isabel Askenasy[1,2*] (askenasyflor@wisc.edu), Jyotsna Misra[1,2], and **Patricia Kiley**[1,2,3]

[1] DOE Great Lakes Bioenergy Center, University of Wisconsin-Madison, Madison, WI; [2] Department of Biomolecular Chemistry, University of Wisconsin-Madison, Madison, WI; [3] University of Wisconsin School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI

https://www.glbrc.org

Project Goals: Our goal is to improve the production of terpenoids in *Zymomonas mobilis* by the MEP pathway through the engineering of IspG and IspH. These enzymes are the bottleneck in the pathway due to the nature of their active site that contains an iron-sulfur cluster. To improve IspG and IspH activity we propose to engineer the proteins itself, as well the ability of *Z. mobilis* to produce and deliver the iron-sulfur clusters.

Terpenoids can be used as a substitute for petroleum in the production of plastics and biofuels. The bacterium *Zymomonas mobilis* produces terpenoid precursors, dimethylallyl diphosphate (DMADP) and isopentenyl diphosphate (IDP) from glucose via the methyl erythritol phosphate (MEP) pathway. Because the MEP pathway is proposed to be limited by the activity of the enzymes IspG and IspH, we are testing ways to increase their activity in *Z. mobilis*. Both IspG and IspH contain a [4Fe-4S] cluster in their respective active sites, which makes them prone to O_2 damage, and recent data indicate that pathway intermediates accumulate in response to O_2. We have initiated several lines of investigation to study the role of IspG/H activity in MEP pathway function. To test the O_2 sensitivity of the *Z. mobilis* IspG and IspH variants, we took advantage of *E. coli* tester strains that conditionally require IspG or IspH activity for growth. We found that under either aerobic or anaerobic conditions, *Z. mobilis* IspG fully complemented growth of the relevant *E. coli* tester strain. In contrast, *Z. mobilis* IspH only fully restored growth of the relevant strain under anaerobic conditions. These results suggest that *Z. mobilis* IspH is more O_2 sensitive than its *E. coli* counterpart. To test this hypothesis, we will compare the effect of O_2 on the stability of the [4Fe-4S] cluster in isolated *E. coli* and *Z. mobilis* IspH, use this tester strains to identify O_2 resistant IspH variants, and test the ability of these variant proteins to increase the ability of *Z. mobilis* to produce terpenoids by the MEP pathway. In a second approach, we are testing if improving [4Fe-4S] cluster delivery or synthesis would increase the occupancy, and accordingly, the activity of IspH. When each of the 3 predicted iron-sulfur cluster carriers from *Z. mobilis* were co-expressed with *Z. mobilis* IspH in the *E. coli* tester strain, none of them improved growth under aerobic conditions. We are also testing if expression of the Fe-S cluster biosynthetic machinery affects *Z. mobilis* IspH activity, since this bacterium contains a homolog of the *E. coli* [2Fe-2S]-containing transcription factor, IscR. Purification of the *Z. mobilis* IscR homolog has shown that it carries an iron-sulfur cluster, which is destroyed

by O_2. To test the role of this IscR homolog in expression of the Fe-S cluster-biosynthetic machinery, we are constructing a strain of *Z. mobilis,* which lacks this gene. Combined, we predict that these studies will allow us to test if *Z. mobilis* strains with increased Fe-S cluster occupancy of IspH can be used to improve terpenoid production.

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409.

The Response of Bacterial Communities to Nitrogen Fertilization Depends on Temporal and Spatial Scale

Lukas Bell-Dereske[1,2]* (belldere@msu.edu), **Lisa Tiemann**[1,3], **Maren Friesen**[4], **Jim Cole**[3], and **Sarah Evans**[1,2,5,6]

[1]DOE Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI,[2]W.K. Kellogg Biological Station, Michigan State University, Hickory Corners, MI,[3]Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI , [4]Department of Plant Pathology & Department of Crop and Soil Sciences, Washington State University, Pullman, WA, [5]Department of Integrative Biology, Michigan State University, East Lansing, MI, [6]Department of Microbiology & Molecular Genetics, Michigan State University, East Lansing, MI

https://rhizosphere.msu.edu/
https://www.glbrc.org/research/sustainable-cropping-systems

Project Goals: We are interested in exploring the role microbial communities play in sustainable growth of dedicate bioenergy crops on marginal lands. Here we are exploring the response of soil microbial communities associated with switchgrass (*Panicum virgatum*) to nitrogen fertilizer across spatial and temporal scales. We will use this information to inform efforts to develop microbial communities with traits that improve bioenergy crop productivity and tolerance to stresses.

Understanding plant-microbial interactions is critically important for producing cellulosic energy crops. Plants interact with a diverse microbial community throughout their life cycle and this community can be utilized to increase yields with reduced application of fertilizers. However, microbial communities vary temporally and spatially, and we are only beginning to understand the scale and drivers of these processes. We have begun to explore the role of the microbial community in the response of the biofuel crop species *Panicum virgatum* to nitrogen fertilization across spatial scales from <1m to >500km (spanning the GLBRC Marginal Land Experimental sites in Michigan and Wisconsin) and temporal scales of weeks to years (spanning two growing seasons). We surveyed the bacterial community along with soil chemistry, weather data, and bacterial traits. We found that nitrogen fertilization had a weak effect on the bacterial community composition and this effect was dependent on spatial and temporal scale. The effect of fertilizer on the bacterial community was strongest and most variable at the southern Michigan Marginal Land site (Lux Arbor) and the effect peaked after application in 2016. This peak in fertilizer effect on the bacterial community corresponded with a peak in soil nitrate. Across all the environmental factors measured, soil moisture and variability in soil temperature were the major environmental drivers of bacterial community composition. However, models including distance, time, and environmental factors were only able to explain between ~18% and 25% of the bacterial community composition. Overall, we found that the response of the microbial community to nitrogen fertilization is dependent on short term flushes and soil nitrogen is likely not the major driver of the bacterial community in switchgrass bioenergy systems.

This material is based upon work supported in part by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award

Numbers DE-SC0018409 and DE-FC02-07ER64494 in addition to Award Number DOE DE-FOA-0001207 to the MMPRNT project.

Vertical niche analysis of rhizosphere and soil fungal and bacterial communities: A deep look into sustainable biofuel cropping systems

Gian Maria Niccolò Benucci[1,2], Xinxin Wang[1,2], Pedro Beschoren da Costa[1,2]* (beschore@msu.edu), and **Gregory Bonito**[1,2]

[1] DOE Great Lakes Bioenergy Research Center, East Lansing, MI; [2] Michigan State University, Department of Plant, Soil, and Microbial Sciences, East Lansing.

https://www.glbrc.org/

Project Goals: The main goal of the Great Lakes Bioenergy Research Center is to develop sustainable biofuels and bioproducts from all usable portions of dedicated energy crops grown on marginal or non-agricultural lands. One tool we are exploring to sustainably improve the productivity of switchgrass (*Panicum virgatum*) biomass is the plant microbiome. Here we report on amplicon sequence analysis of fungi and bacteria from soils and roots sampled across different depths at the GLBRC Biofuel Cropping System Experiment. The aim of this research is to assess how the switchgrass rhizobiome relates to other plant species, how it correlates to plant biomass, and how it changes according to soil depth. Ultimately, we seek to uncover key microbes that can be used and manipulated to improve plant productivity, health and resilience in marginal lands.

Plant-microbe interactions can be extensively exploited for sustainable biofuel cropping. It is well established that fungi and bacteria living in and around plant tissues directly affect plant physiology and development, and that understanding the details in the dynamics of microbial communities is critical for 21st century agriculture. Advances obtained through Next Generation Sequencing (NGS) technologies have shown, however, that decrypting the plant microbiome is a long-term endeavor. This is due to the inherent complexity of microbial communities and plant-microbial interactions. Further, soils deeper than 20cm compose a poorly understood and characterized environment, but are likely to have important implications on switchgrass development, as *Panicum virgatum* develops long and deep root systems. Here, we utilize integrated research resources from the GLBRC Biofuel Cropping System Experiment to explore the microbial community composition over a soil-depth profile, with the ultimate goal of modeling how soil microbial networks affect switchgrass sustainable production. The hypotheses we are testing are that: (1) soil depth will create a non-linear frequency gradient of specific microbial taxa; (2) higher connectivity between soil and root compartments across the depth gradient will be positively correlated to plant health and microbial biodiversity; (3) significant correlations between microbe OTU frequency and crop yield will be higher in native species than in exotic species, especially in deeper soil layers less exposed to perturbations; (4) communities with multiple plant species will have more complex co-occurrence OTU networks compared to monocultures, with an inverse relation between soil depth and network complexity; (5) occurrence and relevance of indicator genera across soil depth will be correlated to higher

network connectivity of the same genera. To test these hypotheses, soil cores with depths of 0 to 10cm, 10 to 25 cm, 25 to 50cm, and 50 to 100cm were collected from plots with native switchgrass (*Panicum virgatum*, var. Cave-in-rock) monocultures, exotic hybrid popular ("NM-6", *Populus nigra* x *Populus maximowiczii*) monocultures, and restored prairie 17-species mix. Three representative samples were obtained from each of the 5 plot replicates, totaling 180 samples that were divided between soils and root fragments. Total DNA was extracted from these samples and amplified with the primer set 515F/806R targeting the V4 region of prokaryotic 16S rDNA and ITS1f/ITS4 targeting fungal internal transcribed spacer (ITS) rDNA region. Amplicon libraries were sequenced on the Illumina MiSeq platform. A total of 20,515,462 raw reads were obtained with the ITS1f/ITS4 primer in 390 samples (including controls), while 25,205,138 raw reads were obtained with the 515F/806R primer set in 387 samples (including controls). Data filtering, standardization, and preliminary analysis are currently underway.

References

1. Zhang B, Penton CR, Xue C, Quensen JF, Roley SS, Guo J, Garoutte A, Zheng T, Tiedje JM (2017) Soil Biology & Biochemistry Soil depth and crop determinants of bacterial communities under ten biofuel cropping systems. Soil Biol Biochem 112:140–152. doi: 10.1016/j.soilbio.2017.04.019.

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-SC0018409 and DE-FC02-07ER64494.

Advanced Biofuels of the Future: Atom−Economical or Energy–Economical?

Peyman Fasahati[1,2]* (fasahati@wisc.edu), **Christos T. Maravelias** [1,2]

[1] DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI;

[2] Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI.

Project Goals: This study aims to identify critical factors that are important in the assessment and selection of new conversion strategies and biofuels. We specifically assess interplay of three key factors: (1) carbon yield, (2) conversion energy efficiency at biorefinery, which indirectly determines excess electricity production, and (3) biofuel type and quality.

Biofuels are sustainable energy sources which could lead to the reduction of GHG emissions by displacing fossil fuels consumption. However, limited resource availability warrants that optimal strategies for efficient utilization of biomass are exploited. This study aims to identify the critical factors that must be considered in the selection of target biofuels and development of new deconstruction and conversion strategies. To this end, a high level carbon and energy analysis of two idealized representative biomass−to−fuel strategies is performed to identify critical trade-offs. The impact of process energy requirements and byproduct (electricity) production, as well as the impact of fuel quality are studied by extending the boundary of analysis to include the end−use of the produced fuel and electricity in the transportation sector. Results show that the interplay of three critical factors (i) carbon efficiency, (ii) conversion energy efficiency at biorefinery, and (iii) biofuel quality must be considered to fairly understand the merits and potential of new conversion strategies. In addition, results from detailed technoeconomic analysis reported in the literature are used to validate our assumptions and modeling approach, as well as conclusions and insights.

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409.

Modulating hemicellulose to improve bioenergy crops

Sang-Jin Kim (sjkim@msu.edu)[1,2], Starla Zemelis-Durfee[1], Jacob Krüger Jensen[1,4], Mingzhu Fan[1,4], Curtis G. Wilkerson[1,3,4], Kenneth Keegstra[1,2,3], **Federica Brandizzi**[1,2,4]

[1]Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI; [2]MSU-DOE Plant Research Lab, Michigan State University, East Lansing, MI; [3]Department of Biochemistry & Molecular Biology, Michigan State University, East Lansing, MI; [4]Department of Plant Biology, Michigan State University, East Lansing, MI.

https://www.glbrc.org

Project Goals : The goal of Bioenergy Plant Design team in Great Lakes Bioenergy Research Center (GLBRC) is to increase the quantity and quality of bioenergy crop biomass per hectare of land, which is crucial for the sustainable and economically viable production of lignocellulosic-derived bioproducts.

The plant cell wall (CW) is composed of polysaccharides and lignin, which have specific roles during plant's growth. Importantly, the CW constitutes the majority of the biomass destined for conversion. Therefore, we want to improve quantitative and qualitative CW traits such as elevating the production and accumulation of cell wall polysaccharides, including mixed linkage (1,3;1,4)-β-glucan (MLG), a low-recalcitrance glucose polymer. Thus, characterizing and engineering MLG synthases are also required to produce MLG with high efficiency. To achieve our goals, we used *Brachypodium distachyon* as a model grass species. MLG is one of the major components of cereal grains, and MLG biosynthesis depends on the biochemical activity of membrane spanning glucan synthases encoded by the CSLH and CSLF cellulose synthase-like gene families. As the first step of the project, we demonstrated the topology of CSLF6 protein derived from Brachypodium (BdCSLF6) using heterologous expression systems. We reported that a functional YFP fusion of BdCSLF6 is localized to the Golgi apparatus and that the Golgi localization of BdCSLF6 is sufficient for MLG biosynthesis using its catalytic domain in the cytoplasm. The localization of BdCSLF6 was further confirmed in its native environment, and MLG was detected in the Golgi, post-Golgi structures and in the cell wall. Accordingly, analyses of a functional fluorescent protein fusion of CSLF6 stably expressed in Brachypodium demonstrated that the enzyme is localized in the Golgi as we have seen in tobacco. We also established that overproduction of MLG causes developmental and growth defects in Brachypodium as in barley. To overcome the growth defect by over-accumulation of MLG, we generated plants with improved biomass by expressing the endoplasmic reticulm (ER) stress sensor IRE1 and crossed them with MLG over-accumulating Brachypodium plants to utilize the unfolded protein response, thereby alleviating stress caused by overexpression of BdCSLF6. We found that the transgenic lines for both IRE1OX and CSLF6OX maintained the increased biomass trait of the IRE1OX lines, supporting that IRE1 overexpression suppresses the growth penalty induced by BdCSLF6 overexpression. We also found an increased MLG level in the crossed line compared to wild type. To understand the tissue- and development-specific regulation of MLG synthesis and IRE1 expression, we have employed a new *in-situ* hybridization technique. Applying this technique in the Brachypodium seeds indicated that MLG synthesis is specifically activated in the endosperm with a tissue specific manner. With this

result, we expect to apply this technique in the other stem tissue to understand MLG synthesis, degradation and impact of plant growth.

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-SC0018409 and DE-FC02-07ER64494.

Biomimetic Lignin Cleavage Using Small Thiols

Grace E. Klinger,[1,2]* (klingerg@msu.edu), James, E. Jackson,[1] and **Eric L. Hegg**[1,2]

[1]Michigan State University, East Lansing, MI; [2]Great Lakes Bioenergy Research Center, Biological Research Center, East Lansing, MI;

https://www.glbrc.org/research

Project Goals: Depolymerization of lignin and lignin-like structures using small organic thiols that mimic the glutathione in the enzymatic β-aryl ether cleavage pathway.

Lignin valorization for the replacement of petrochemicals is challenging due, in part, to the expensive inputs and/or caustic depolymerization techniques typically required. Lignin-degrading enzymes provide an excellent model for a simple and green cleavage method. In the β-aryl ether cleavage pathway of wood degrading bacteria, the cleavage of β-O-4 bonds occurs in a three-step process: 1) oxidation of the α-carbon hydroxyl group, 2) glutathione nucleophilic attack on the β-carbon, displacing the phenoxide, and 3) reduction of the above-formed glutathione covalent intermediate's S-C bond with a second glutathione, releasing the second lignin fragment and the glutathione disulfide. This work focuses on mimicking this enzymatic pathway using small organic thiols without the aid of proteins or metals. Oxidized lignin dimer models treated with β-mercaptoethanol and related thiols undergo cleavage with yields ranging from 30% to 100% depending on the specific reaction conditions. The sensitivity of this cleavage to lignin-relevant functional groups was probed to determine the scope of the reaction. As in the enzymatic pathway, oxidation of the alpha hydroxy group was found to be necessary. Primary and secondary beta aryl ether bonds were cleaved while tertiary sites were unreactive as expected for S_N2 reactivity. Lastly, the reaction tolerates a variety of functional groups on both the phenolic and keto aromatic rings. Having mechanistically analyzed the reactions in well-defined model dimers, we applied this process to real lignin, achieving approximately 68% molecular weight reduction. This work exemplifies a reductive biomimetic approach to lignin depolymerization by mimicking the nucleophilic thiol-mediated ether cleavage found in the enzymatic β-aryl ether cleavage pathway.

References

1. Hegg, E. L.; Jackson, J. E.; Klinger, G. E. Depolymerizing lignin by reacting lignin compound with thiol compound to depolymerize lignin compound and form depolymerized lignin product having reduced molecular weight relative to lignin compound prior to reacting. U.S. Patent WO2018195000-A1, Oct. 25, 2018.

This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494 and DE-SC0018409).

Engineering Alternative Model Yeast Species for Biofuel Production

David J. Krause[1,2]* (dkrause2@wisc.edu), Russell L. Wrobel,[1,2] and **Chris Todd Hittinger**[1,2]

[1]DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI; [2]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

Project Goals: The Great Lakes Bioenergy Research Center is developing the production of sustainable biofuels and bioproducts from dedicated energy crops grown on marginal lands. Our research is focused on engineering various species of yeast to convert sugars derived from lignocellulosic biomass into advanced biofuels such as isobutanol. By studying alternative model systems, we aim to identify and bypass engineering roadblocks, gain insight into the robustness and breadth of applicability of metabolic engineering strategies, and discover and develop new biological tools.

Saccharomyces cerevisiae has long been considered the model yeast species in many arenas, including the production of biofuels. Producing large quantities of isobutanol from sugars found in lignocellulosic hydrolysates will likely require the deletion of genes in competing pathways that normally direct most of pyruvate flux into ethanol fermentation: pyruvate decarboxylase (*PDC*). Deletion of all three *PDC* isoforms in *S. cerevisiae* renders the mutant unable to grow on glucose.[1] This presents a significant roadblock to further genetic engineering of this species to produce alternative end products. In contrast, the yeast *Kluyveromyces lactis* contains only a single isoform of *PDC*, and deletion of this gene does not prohibit growth on glucose, making further engineering of alternative metabolism on glucose feasible. Here we report initial characterization of the Kl*pdc1Δ* strain during growth on glucose, and its potential for use to bypass the growth deficiency of its *S. cerevisiae* counterpart.

After the removal of competing reactions that would otherwise divert carbon flux into non-target products, the next step of rationally designed engineering approaches is often to introduce new genes or upregulate existing genes for the desired pathway. For the production of isobutanol, these are the *ILV* genes that direct pyruvate through branched-chain amino acid (BCAA) biosynthesis, as well as a decarboxylase and alcohol dehydrogenase. These approaches to isobutanol production in *S. cerevisiae* have shown some success, and applying similar strategies to *K. lactis* will inform us of the robustness of these engineering approaches when applied to other yeast species.

Finally, we are exploring alternative approaches to rational design for engineering organisms to produce specific end products. *K. lactis*, as well as several other yeast species, produce a red, iron-binding molecule during growth called pulcherrimin. We recently identified and characterized a secondary metabolite gene cluster in this species responsible for the production and utilization of this iron-binding siderophore.[2] The evolution of this gene cluster suggested an

ecological role for this molecule, possibly in competition for iron in the environment. Because the carbon in pulcherrimin is derived from two leucine molecules, its production is a result of carbon flux through BCAA biosynthesis. We are exploring the utility of this molecule to link secondary metabolism with primary metabolism and develop alternative strategies to increasing BCAA flux in yeasts.

References

1. Flikweert M.T., van der Zanden L., Janssen W.M.T.M., Steensma H.Y., van Dijken J.T., Pronk J.T. 1996. *Yeast.* 12:247-257.
2. Krause D.J., Kominek J., Opulente D.A., Shen X.X., Zhou X., Langdon Q.K., DeVirgilio J., Hulfachor A.B., Kurtzman C.P., Rokas A., Hittinger C.T. 2018. *Proc. Natl. Acad. Sci. USA.* 115:11030-11035.

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-SC0018409 and DE-FC02-07ER64494.

Genetics in *Zymomonas mobilis* ZM4

Piyush Behari Lal,[1,2]* (plal2@wisc.edu), Fritz Wells,[2] and **Patricia Kiley**[1,2]

[1]Great Lakes Bioenergy Research Center, University of Wisconsin – Madison, Madison, WI;
[2]Department of Biomolecular Chemistry, University of Wisconsin – Madison, Madison, WI.

http://www.glbrc.org/

Project Goals:

The goal of this work is to develop genetic tools for *Zymomonas mobilis* ZM4, and create a platform strain with increased efficiency to accept foreign genes.

Abstract

The ethanologenic bacterium *Zymomonas mobilis* ZM4 (*Z. mobilis*) has emerged[1,2,3] as a promising candidate for microbial conversion of plant biomass (lignocellulosic biomass/hydrolysate) into biofuel and other value-added products. The present work aims to expand genetic techniques for *Z. mobilis* so that strain optimization is feasible and efficient. To remove genes from *Z. mobilis*, we have developed a method for generating markerless gene deletions. This is a two-step process that involves integration of an engineered suicide plasmid into the genome via homologous recombination, and a subsequent recombination event that leads to loss of target gene. A key feature of this strategy is that GFP expressed from integrated vector allows easy identification of cells that have lost the integrated plasmid by fluorescence activated cell sorter (FACS). We have also improved transconjugation of plasmids from *Escherichia coli* into *Z. mobilis* by deleting *Z. mobilis* restriction systems. We propose this work will provide a platform for genetic engineering in *Z. mobilis* to build better biofuel producing strains.

References

1. *He et al*., 2014.
2. *Motamedian et al*., 1977.
3. *Swings and Ley*, 1977.

Funding statement.

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-SC0018409 and DE-FC02-07ER64494.

Kinetic and mechanistic insights into hydrogenolytic production of lignin monomers in a continuous flow-through system

Yanding Li,[1,2,†]* (yli594@wisc.edu) Benginur Demir,[2,3,†] Leida M. Vázquez Ramos,[2,3] Mingjie Chen,[2] **James A. Dumesic**,[2,3] **John Ralph**[1,2,4]

[1]Department of Biological Systems Engineering, University of Wisconsin–Madison, Madison, WI. [2]DOE Great Lakes Bioenergy Research Center, University of Wisconsin–Madison, Madison, WI. [3]Department of Chemical and Biological Engineering, University of Wisconsin–Madison, Madison, WI. [4]Department of Biochemistry, University of Wisconsin–Madison, Madison, WI.

http://www.glbrc.org

Project Goals: The mission of the Great Lakes Bioenergy Research Center (GLBRC) is grand, but simply stated: to perform the basic research that generates technology to convert cellulosic biomass to advanced biofuels. The project combines aims for the "design" of superior plant lignins with methods that allow their depolymerization to the highest yields of phenolic monomers. Lignin depolymerization by catalytic hydrogenolysis is one of the most promising depolymerization methods to produce platform lignin monomers with both high yield and selectivity. This study focuses on the kinetic and mechanistic insights into the lignin hydrogenolysis reaction in a continuous flow-through system. Understanding the function of each reaction parameter is the first and essential step to upscaling this method.

Environmental issues caused by the unrestrained use of fossil energy require a renewable carbon-neutral substitute. Lignocellulosic biomass is one of the most promising alternatives as a widely available sustainable resource. It typically consists of 40-50% cellulose, 25-30% hemicelluloses, and 15-25% lignin. It can be fractionated into these three component streams, and each stream can be upgraded to fuels and chemicals. However, the recalcitrance of the plant cell wall, lignin in particular, makes biomass fractionation and upgrading inefficient. Most of the current biorefinery processes focus on the conversion of the polysaccharides, in which the biomass fractionation using harsh acids, bases, or additional chemicals, destroys the native lignin structure.[1] The lignin undergoes irreversible condensation reactions, such that essentially the only option left for the condensed lignin is to burn it to generate low-value heat.

Native lignin is biosynthesized from its three monolignols: *p*-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol, via combinatorial radical coupling reactions. The most common type of native lignin unit is characterized by its β–O–4 ether units (50-60%), and the rest are minor C–C and diaryl ether units. Full cleavage of the β-ethers leads to effective lignin depolymerization. For this purpose, reductive catalytic fractionation (RCF) of lignin simultaneously fractionates the biomass

and cleaves the β-ether linkages.[2] Lignin hydrogenolysis has been studied since 1938 and, with the evolution of the technique, high monomer yields with high selectivity can be achieved. However, the reaction kinetics and mechanism of lignin hydrogenolysis are still not well understood. Elucidation of the fundamentals of lignin hydrogenolysis was sought to guide us not only to create active and robust catalysts, but also to model selective reaction environments for the complex upgrading process.

Lignins are often obtained from bulk biomass fractionation processes. Kraft and organosolv lignins are the major technical lignins currently produced on a large scale by the pulp and paper industry. However, the high sulfur content in the former and the often extremely condensed structure of especially the latter make them unsuitable for further upgrading. A high-quality lignin stream is essential to add value to the biorefinery process. Researchers have emphasized the limiting factors towards the lignin depolymerization process, and rationalized a few basic principles of a high-quality lignin ideotype for depolymerization.[3] Briefly, when wild-type biomass is used, pretreated lignin with a high β-ether content, such as Cu-AHP lignin, aldehyde-protected lignin, and GVL-lignin, is more valuable for downstream lignin upgrading.

We propose herein a fully continuous biomass refinery as a combination of a traditional biomass pretreatment process with a continuous lignin hydrogenolysis pathway in a flow-through system using a Pd/C catalyst. We describe a method to obtain value-added platform monomers from the downstream processing of the isolated lignin stream to enhance the biorefinery. With the help of lignin model compounds, the reaction pathways of lignin hydrogenolysis were investigated and a kinetic model developed. Understanding the reactivity of lignin model compounds helps us design the reaction conditions for actual lignin hydrogenolysis.

References

1. Schutyser, W.; Renders, T.; Van den Bosch, S.; Koelewijn, S.-F.; Beckham, G. T.; Sels, B. F., Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chem Soc Rev* **2018,** *47* (3), 852-908.
2. Renders, T.; Van den Bosch, S.; Koelewijn, S.-F.; Schutyser, W.; Sels, B., Lignin-first biomass fractionation: the advent of active stabilisation strategies. *Energ Environ Sci* **2017,** *10* (7), 1551-1557.
3. Li, Y.; Shuai, L.; Kim, H.; Motagamwala, A. H.; Mobley, J. K.; Yue, F.; Tobimatsu, Y.; Havkin-Frenkel, D.; Chen, F.; Dixon, R. A.; Luterbacher, J. S.; Dumesic, J. A.; Ralph, J., An "ideal lignin" facilitates full biomass utilization. *Science Advances* **2018,** *4* (9), eaau2968, 1-10.

Funding was provided by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-SC0018409).

Regulation of biomass yield and composition of energy sorghum stems

Brian McKinley, Joel Oliver, Bill Rooney, Starla Zemelis-Durfee, Curtis Wilkerson, Federica Brandizzi, and John Mullet

Sorghum bicolor is a resilient, sustainable, low cost source of specialty biofuels and bio-products due to the crop's high biomass yield, drought and heat tolerance, excellent genetic and genomic platform, and highly diverse germplasm. Our research is focused on increasing stem biomass yield and optimizing stem composition and value. During the stem growth phase, substantial carbon is allocated to cell wall biosynthesis associated with cell proliferation and internode expansion. Following floral initiation, stem growth and cell wall biosynthesis declines and non-structural carbohydrates accumulate in the pith parenchyma cells of sorghum stems. To better understand carbon partitioning during the stem growth phase we are investigating the function of Dw2/KIPK, a locus/gene that regulates stem growth in sorghum using transcriptomic and phosphoproteomic analyses. A better understanding of this gene and its regulatory network will facilitate the engineering of stem growth and enhance biomass yield. Energy sorghum can accumulate up to 50% of its stem biomass as non-structural carbohydrates (sucrose, starch, MLG) that can be readily converted to specialty biofuels and bioproducts. We have identified candidate genes and mapped pathways involved in the synthesis and degradation of the non-structural carbohydrates that accumulate in sorghum stems. The identification of these genes was facilitated by the collection of transcriptome data from most sorghum organs/tissues during plant development in collaboration with the Joint Genome Institute. The information obtained through these studies comprises a sorghum transcriptome compendium which is being used to design plants with higher biomass yield and improved composition.

Development of Biosensors for High-throughput Enzyme Screens

Anthony Meger[1,3]* (ameger@wisc.edu), Kyle Nishikawa,[1] Megan Leander,[1] and **Srivatsan Raman**[1,2,3]

[1]Department of Biochemistry [2]Department of Bacteriology [3]Great Lakes Bioenergy Research Center, University of Wisconsin-Madison

Project Goal: Engineer sugar-responsive biosensors to enable high-throughput screens of cellulase mutants for improved enzymatic activity.

Allosteric TFs (aTF) are useful metabolite sensors which enable directed evolution of enzymes and biosynthetic pathways. Expanding aTFs to respond to new molecules would greatly increase their utility, facilitating the biosynthetic production of fine chemicals, intermediates, and fuels. However, inducer recognition and transcriptional response in aTFs are tightly coupled through allostery, making redesign toward new inducers challenging. Using computational protein design guided by evolutionary analysis, we engineer variants of the *E. coli* lac repressor, LacI, to respond to a variety of fermentable sugars while preserving allosteric function. These sugar biosensors are highly valuable for screening cellulase variants with broader substrate specificities and higher catalytic efficiencies. Superior cellulases have the potential to improve cellulose recovery from biomass and reducing the overall cost of commercial production of biofuels.

References

1. Walker, J. A., Takasuka, T. E., Deng, K., Bianchetti, C. M., Udell, H. S., Prom, B. M., Kim, H., Adams, P. D., Northen, T. R., … Fox, B. G. (2015). Multifunctional cellulase catalysis targeted by fusion to different carbohydrate-binding modules. Biotechnology for biofuels, 8, 220. doi:10.1186/s13068-015-0402-0
2. Taylor, N. D., Garruss, A. S., Moretti, R., Chan, S., Arbing, M. A., Cascio, D., Rogers, J. K., Isaacs, F. J., Kosuri, S., Baker, D., Fields, S., Church, G. M., … Raman, S. (2015). Engineering an allosteric transcription factor to respond to new ligands. Nature methods, 13(2), 177-83.

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409.

Microbial Conversion of Chemically Depolymerized Lignin Into Valuable Compounds

Miguel Perez[1,2]* (perez8@wisc.edu), Wayne Kontur[2], Manar Alherech[2,3], Steven Karlen[2], Shannon S. Stahl[2,3], **Timothy J. Donohue**[2,4], Daniel R. Noguera[1,2]

[1] Department of Civil and Environmental Engineering, University of Wisconsin-Madison, Madison, WI.; [2] DOE Great Lakes Bioenergy Research Center, Madison, WI.; [3] Department of Chemistry, University of Wisconsin-Madison, Madison, WI.; and [4] Department of Bacteriology, University of Wisconsin-Madison, Madison, WI.

https://www.glbrc.org/

Project goals: The project aims to valorize the lignin fraction of plant biomass via chemical fractionation and depolymerization followed by conversion of the resulting mixtures of aromatic compounds into single valuable chemicals by genetically engineered bacteria.

Plant cell wall consists mainly of a combination of three different types of chemical polymers: cellulose, hemicellulose, and lignin. Lignin is a heteropolymer of different types of aromatic compounds whose chemical properties make it highly insoluble and recalcitrant to chemical and biological breakdown. This presents a major challenge for a full valorization of lignocellulose.

Recently developed chemical approaches for lignin deconstruction result in mixtures of different aromatic compounds that share common aromatic structures of three types: Syringyl (S), Guaiacyl (G), or *p*-hydroxy (H). On the other hand, some microbial strains have evolved to utilize multiple lignin-derived aromatic compounds as their source of carbon and energy for growth via metabolic funneling into a few aromatic intermediates before complete degradation. This natural capability presents an attractive opportunity for upgrading aromatic compounds via metabolic engineering of suitable strains.

Novosphingobium aromaticivorans DSM12444 has the ability to catabolize multiple S, G, and H type aromatic compounds present in oxidized and formic acid induced depolymerized lignin. Of particular interest, this strain can transform guaiacyl and syringyl-derived diketones, which can result from oxidative chemical depolymerization, but are not found in natural environments. In this work, we analyze the metabolism of these and other lignin-derived aromatics by *N. aromaticivorans*, identify genes required for metabolism of S, G and H aromatic units and demonstrate how the use of mutants to transform these compounds into potentially valuable products.

This work was supported by U.S. Department of Energy (DOE) Great Lakes Bioenergy Research Center grants (DOE Office of Science BER DE-FC02-07ER64494 and DE-SC0018409).

Investigating Mechanisms of Soil Carbon Accrual and Protection in Bioenergy Cropping Systems

Matthew L. Reid[1,*] (reidmat1@msu.edu), Jenifer A. Gonzalez[2], Fiona M. Käch[1], Darian N. Smercina[1], and **Lisa K. Tiemann**[1]

[1]Michigan State University, East Lansing, MI; [2]University of the District of Columbia, Washington D.C.

https://www.glbrc.org/

https://tiemann.psm.msu.edu/carbon-accrual-in-managed-systems/

Project Goals: The goals of this study were to understand mechanisms controlling the protection and accrual of soil carbon in bioenergy cropping systems. Specifically, we aimed to characterize differences in potential for soil C contributions versus loss as CO_2 of different plant residue types. We use controlled, laboratory soil incubation experiments to quantify the respiration from decomposition of soil organic matter and plant residues. Additionally, we quantify the importance of physical versus chemical (bioaccessibility) protection of soil organic carbon in soil aggregates of various size fractions from bioenergy cropping systems. These data will help us better understand which bioenergy systems and soils are best suited for promoting soil C accrual with the ultimate goal of creating C neutral or C negative systems.

Perennial bioenergy cropping systems offer the potential to accrue soil carbon, potentially making them a carbon-neutral energy source. We seek to understand patterns of soil carbon accrual using a series of laboratory incubations to characterize soil carbon dynamics in bioenergy cropping systems. In our first incubation, we characterized differences in soil respiration for soils amended with plant tissue residues. We used soils from a Conservation Reserve Program (CRP) field that has been dominated by the C3 grass smooth brome (*Bromus inermis*) and residues from C4 grass bioenergy crops, corn (*Zea mays*) and switchgrass (*Panicum virgatum*). The soil organic matter has a $\delta^{13}C$ signature reflective of the C3 grass of -27‰, while the corn and switchgrass residues have a $\delta^{13}C$ signature of -13‰. This isotopic difference allows us to differentiate between the C respired during decomposition of soil organic matter and that respired from decomposition of the residues. We added 0.45 g of either corn or switchgrass roots, shoots, or roots and shoots to 30 g of CRP soil. We assessed respiration rates for 39 days, and found that across both species, residue addition of any kind stimulated total respiration compared to control soils with no residues additions. Total C respired from soils and residues was highest in soils plus shoots; 32% greater than soil plus root and 17% greater than soil plus roots and shoots. Across all treatments, residue additions increased peptidase activity and decreased chitinase activity. Isotope analyses indicated that across all treatments residues were the primary source of C respired, and addition of residues suppressed respiration from soil

organic carbon, relative to soil only controls. In our second incubation, we characterized the physical and chemical protection of soil carbon using soils collected from large-scale experimental fields that have been in corn, switchgrass or maintained as CRP land for the past 9 years. We partitioned 100 g of soil from these fields into three aggregate size fractions (<0.5 mm, 0.5-2 mm, >2 mm). We initiated a soil incubation with aggregates from these different size classes, using crushed and un-crushed aggregates to assess physical protection, while chemical protection was assessed via addition of glucose. Glucose provides a very accessible high-energy C source that would potentially prime soil bacteria to break down less energetically favorable (chemically protected) SOC. After 109 days of incubation, we see no consistent effect of crushing aggregates on soil respiration rates, indicating that physical protection of C in these systems may not be an important mechanism for C accrual. During the first 109 days of the incubation we've added glucose twice. Overall, glucose addition did increase respiration rates, but this effect was inconsistent across crop species and aggregate sizes. Within the aggregates from corn, glucose addition increased soil respiration rates in the <0.5 mm and >2 mm aggregates, but not the 0.5-2 mm aggregates. There was no effect of glucose addition on any aggregate sizes from both switchgrass and CRP fields. These incubations have thus far revealed that patterns of soil C protection and accrual are affected by plant residue inputs and additions of glucose, a common root exudate, but these effects are highly dependent on the crop type. A better understanding these patterns of soil C dynamics is necessary to achieve the potential of making bioenergy crops a carbon-neutral energy source.

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409.

The Other Lignocellulosic Leftover: Conversion Residue Valorization with a Microbial Community

Matthew J. Scarborough[1,2*] (mscarborough@wisc.edu), Daniel R. Noguera[1,2], and **Timothy J. Donohue**

[1]Great Lakes Bioenergy Research Center, Madison, Wisconsin; [2] University of Wisconsin - Madison;

https://www.glbrc.org/research/highlights/microbial-community-members-play-distinct-roles-biosynthesis

Project Goals:

The carboxylate platform has emerged as a promising technology to produce carboxylic acids from complex organic wastes (1). Limitations persist in the ability to direct production from short-chain (C1-C5) products to medium-chain (C6-C12) products, which have higher value, are more energy dense, and are easier to recover. I will discuss implementing the carboxylate platform on a waste stream from lignocellulosic biorefining to recover additional carbon and energy from biorefinery "leftovers." We evaluated process stability and economics, and performed metagenomic, metatransciptomic, and thermodynamic analyses of the microbial community to reconstruct metabolic networks. Lastly, we propose strategies to further improve our understanding of industrial microbiome systems and increase production of valuable products from conversion residue.

Abstract:

Starting with an inoculum from a wastewater treatment digester, we enriched a community of microorganisms that produced a variety of volatile fatty acid end products, including medium chain fatty acids (MCFAs) (Figure 1, adapted from (2)). In total, the microbial community converted 18% of the reducing equivalents (measured as chemical oxygen demand, COD) to the MCFAs hexanoic and octanoic acid. Based on this data, we performed a technoeconomic analysis of valorizing conversion residue and determined that the minim ethanol selling price could be reduced by 18% in a biorefinery producing MCFAs as a coproduct to ethanol.

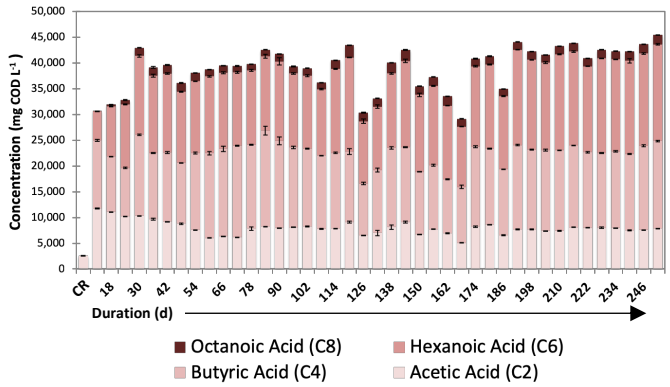


Figure 1. Production of MCFAs from organic material in conversion residue (CR). The microbial community converted xylose and other carbohydrates in conversion residue to acetic acid, butyric acid, hexanoic acid, and octanoic acid.

Amplicon sequencing of the 16s rRNA gene revealed a stable community consisting of members related to five key genera making up > 95% of the sequenced reads: *Lactobacillus*, *Olsenella*, *Atopobium*, *Roseburia*, and *Pseudoramibacter*. To elucidate the roles of these community members, we performed metagenomic and metatransciptomic analyses. This resulted in the recovery of 10 metagenome-assembled genomes: five related to *Lactobacillus* (LAC1-5), three related to the *Coriobacteriaceae* (COR1-3), a member of the *Eubacteriaceae* (EUB1), and a member of the *Lachnospiraceae* (LCO1). Based on highly abundant transcripts, we reconstructed the active metabolic networks in each of these populations. We then performed thermodynamic analyses to refine the predicted roles of each population. In total, we predict the flow of substrates in conversion residue to MCFAs as shown in Figure 2 (adapted from (3)) . Strategies to increase production of MCFAs include constructing a synthetic community to direct more carbon to lactate (rather than acetate) and utilizing hydrogen gas as a supplemental electron donor to further elongate short-chain products.

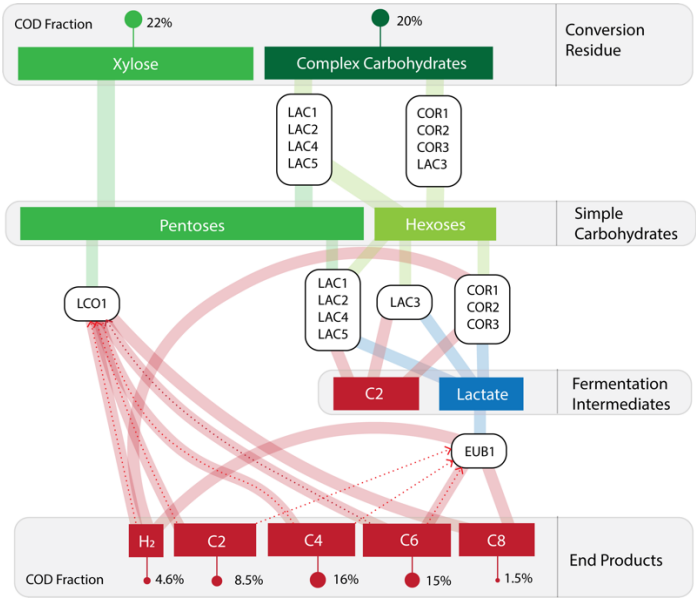


Figure 2. Based on thermodynamic and transcriptomic analyses we proposed functions of individual MAGs within the community. In total, the microbiome converts carbohydrates remaining in conversion residue to a variety of monocarboxylic acid products, including MCFAs. Lactate is expected to be a key intermediate. Future work will develop strategies to drive production of hexanoate and octanoate over acetate and butyrate.

References

1. Angenent LT, *et al.* (2016) Chain Elongation with Reactor Microbiomes: Open-Culture Biotechnology To Produce Biochemicals. *Environ. Sci. Technol.* 50(6):2796-2810.
2. Scarborough MJ, *et al.* (2018) Increasing the economic value of lignocellulosic stillage through medium-chain fatty acid production. *Biotechnology for Biofuels* 11(1):200.
3. Scarborough MJ, Lawson CE, Hamilton JJ, Donohue TJ, & Noguera DR (2018) Metatranscriptomic and thermodynamic insights into medium-chain fatty acid production using an anaerobic microbiome. *mSystems*.

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-SC0018409 and DE-FC02-07ER64494.

Overexpression and Metabolic Regulation of *Z. mobilis* MEP Pathway Enzymes

Mehmet Tatli[1,2*] (mtatli@wisc.edu) and **Daniel Amador-Noguez**[1,2]

[1]Department of Bacteriology, University of Wisconsin-Madison, Madison, Wisconsin, USA;
[2]DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, Wisconsin, USA

https://www.glbrc.org/

Project Goals: Understanding metabolic regulation of *Z. mobilis* MEP pathway and engineering ZM4 strains to overproduce targeted isoprenoids.

Zymomonas mobilis, a facultative anaerobe, can convert 96% of the glucose consumed to ethanol at high yields. This highly catabolic metabolism can also be redirected towards generation of isoprenoid-derived biofuels via the 2-C-Methyl-D-erythritol 4-phosphate (MEP) pathway. Here, we have individually over-expressed the MEP pathway enzymes (DXS, DXR, IspDF, and IspE) of *Z. mobilis* in the ZM4 strain to better understand the metabolic regulations in the first part of the MEP pathway, and investigate the effect of the enzymes on directing carbon flux into the MEP pathway and down to the cyclic intermediate, MEcDP. Initial results showed that DXS2 over-expression increases flux through MEP pathway, leading to a 70 fold increase in intracellular MEcDP levels, and also increases levels of the two end products of this pathway, IDP/DMADP. These results have indicated that DXS2 is a rate limiting enzyme in the MEP pathway of *Z. mobilis*. Moreover, coupling DXS2 with isoprene synthase, IspS, allows for production of isoprene in ZM4. This over-expression strategy also revealed interesting metabolic changes in the MEP pathway, which might bring new insights into understanding metabolic regulation of the MEP pathway.

This work supported by the Great Lakes Bioenergy Research Center, U.S. 769 Department of Energy, Office of Science, Office of Biological and Environmental Research 770 under Award Number DE-SC0018409.

Using Engineered *Streptomyces* for Production of Fatty Acids and Isoprenoids from Lignocellulosic Biofuel Conversion Residue

Kurt O. Throckmorton (kthrockmorto@wisc.edu),[1*] Caryn S. Wadler,[1] and **Michael G. Thomas**[1]

[1]DOE Great Lakes Bioenergy Research Center, University of Wisconsin, Madison

https://www.glbrc.org/

Project Goals:

- **Construct reporter systems to assess metabolic flux toward fatty acid (melanin) and isoprenoid (lycopene) in *Streptomyces* species**
- **Conjugate these reporter systems into a phylogenetically diverse *Streptomyces* strain collection**
- **Screen for growth on lignocellulosic biofuel conversion residue and production of the reporter metabolites**
- **Conduct transcriptomic analyses to identify genetic elements enabling each of these properties and recombine in an ideal host strain**
- **Produce target fatty acid and isoprenoid compounds from conversion residue using these strains**

The process of lignocellulosic bioethanol fermentation leaves behind a substantial portion of the energy-containing organic material, called conversion residue (CR). This CR can contain up to two-thirds of the chemical oxygen demand for biomass hydrolysates produced from AFEX-treated corn stover. Using microbes to convert this CR to value-added bioproducts instead of burning the CR to generate electricity or use of CR in animal feed would increase the value of CR and improves the economics of lignocellulosic biorefineries. Members of the genus *Streptomyces* represent good candidates to fill this role as they are genetically tractable, catabolically versatile, and produce a wide variety of compounds of the types that would be industrially useful, e.g. fatty acids and isoprenoids. Using our transferable reporter systems, we screened a diverse strain collection of *Streptomyces* spp. for their ability to grow on CR and to produce these compounds at high levels. Using comparative transcriptomic analysis, we aim to identify genetic elements that distinguish the high performers in this screen and use genetic and metabolic engineering to optimize their abilities to catabolize CR and to shift more metabolic flux toward fatty acids or polyketides. Using these genetic elements and the identified host strains, we will target high-value fatty acid and isoprenoid for production, chosen in accordance with technoeconomic analysis.

Funding statement.

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409.

The Structure and Specificity Landscape of a Glycoside Hydrolase Family

Kirk A. Vander Meulen,[1,2] Evan M. Glasgow,[1,2] Craig A. Bingman,[1,2] Taichi E. Takasuka,[1,2,3] Christopher M. Bianchetti,[1,2,4] Lai F. Bergeman,[1,2] Samuel Deutsch,[5] and **Brian G. Fox**[1,2]

[1]Great Lakes Bioenergy Research Center, Madison, WI; [2]Department of Biochemistry, University of Wisconsin – Madison; [3]Research Faculty of Agriculture, Hokkaido University, Sapporo, Japan; [4]Department of Chemistry, University of Wisconsin – Oshkosh, USA; [5]DOE Joint Genome Institute, Walnut Creek, CA, USA

https://www.glbrc.org

Project Goals: Great Lakes Bioenergy is a U.S. Department of Energy-funded research center dedicated to developing sustainable biofuels and bioproducts. Studying enzymes capable of hydrolyzing the numerous polysaccharide types contained in plant cell walls will facilitate development of biomass deconstruction strategies.

To deepen the understanding of the structural and evolutionary driving forces underlying specificity patterns in β-1,4 endo-acting GH family 5, we have characterized structure and function across subfamily 4 (GH5_4) and closely related enzymes. GH5_4 is an expansive subfamily consisting of three major clades, and here we mapped activity profiles and structures onto these phylogenetic groupings. In a quantitative enzymatic screen of the catalytic core domains of 243 enzymes on multiple polysaccharide substrates, members from one of these clades (Clade 3) possessed consistently and significantly elevated activities. A subfamily-wide correlation between lichenase and xylanase specific activity values was also observed, suggesting that the ancestral enzyme's structural framework may enforce a linkage between changes in either activity. Crystal structures were determined for 11 members of subfamily 4, and in concert with previous structures depict a diverging binding cleft morphology. Across this backdrop, two cleft residues correlate with the most active enzymes, a histidine residue that hydrogen bonds with the -1 glycosidic subunit and a +1-subsite stacking tryptophan.

References

Glasgow, EM, Vander Meulen, KA, Takasuka, TE, Bianchetti, CM, Bergeman, LF, Deutsch, S, and Fox, BG. Extent and Origins of Functional Diversity in a Subfamily of Glycoside Hydrolases. Journal of Molecular Biology. In Press.

This material is based upon work supported in part by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-SC0018409 and DE-FC02-07ER64494. EG is supported by the NIGMS Biotechnology Training Program (NIH 5 T32 GM008349) at the University of Wisconsin – Madison.

Producing Valuable Bioproducts from the Waste Residue of Biofuel Production

John F. Wolters[1,2,*] (jwolters@wisc.edu), Dana Opulente[1,2], Jacek Kominek[1,2], Rachel M Schneider[1,2], **Chris Todd Hittinger**[1,2]

[1] DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI;

[2] 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

Project Goals: The Great Lakes Bioenergy Research Center is developing the production of sustainable biofuels and bioproducts from dedicated energy crops grown on marginal lands. Our research is focused on identifying non-model yeasts that are capable of utilizing the esidual carbon that is not metabolized by the primary fermentation of ethanol from cellulosic feedstocks. For species that are capable of growth using the waste residue from cellulosic ethanol production we will identify valuable bioproducts produced, such as lipids that can be used for biodiesels or as substitutes for other oils. Incorporating the conversion of the waste residue into these products will increase the cost efficiency of the biofuel production pipeline.

During cellulosic ethanol production cellulosic feedstocks are hydrolyzed to produce sugar streams which are fermented by a primary microorganism to produce ethanol. Some microbes, such as engineered strains of the yeast *Saccharomyces cerevisiae* or the anaerobic bacterium *Zymomonas mobilis*, are capable of metabolizing multiple sugars present in hydrolysate including xylose. However, even after consumption of glucose and xylose a large fraction of the carbon from the original feedstocks remains in the waste residue after ethanol distillation. The use of non-model microbes that can metabolize the complex and uncharacterized compounds remaining in the waste residue could turn this waste product into a valuable feedstock for production of additional bioproducts.

The fungal subphylum *Saccharomycotina* contains over one thousand budding yeast species including many with metabolic capacities not present in model species such as *S. cerevisiae*. New genomics resources being built for these species can enable predictions regarding which species may be able to metabolize the complex substrates found in the waste residue[1]. We demonstrate that the ability to grow on the waste residue is widely distributed across yeast species. The most robust growth profiles are found among basal lineages which include biotechnologically important species including *Yarrowia lipolytica* and *Lipomyces starkeyi* known for their capacity to produce and store high levels of intracellular lipid. We show that lipid production triggered by nitrogen limitation is variable among these species and demonstrate considerable lipogenesis phenotypes in several species capable of robust growth on the waste residue. We consider engineering strategies to better understand the triggers of lipogenesis and how to achieve high lipid titres during growth on the waste residue.

References

1. Shen, X.-X. *et al.* Tempo and Mode of Genome Evolution in the Budding Yeast Subphylum. *Cell* **175,** 1533–1545 (2018).

This material is based on work supported by the Great Lakes Bioenergy Research Center, U.S., Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-SC0018409 and DE-FC02-07ER64494

Comparative Studies of Diverse Lignocellulosic Feedstocks for Microbial Biofuel Synthesis and the Impact on the Biorefinery

Yaoping Zhang[1*] (yzhang8@wisc.edu), Lawrence G. Oates,[1] Jose Serate,[1] Dan Xie,[1] Edward Pohlmann,[1] Yury V. Bukhman,[1] Steven D. Karlen,[1] Megan K. Young,[1] Alan Higbee,[1] Dustin Eilert,[1] Gregg R. Sanford,[1] Jeff S. Piotrowski,[1] David Cavalier,[2] John Ralph,[1] Joshua J. Coon,[1] Trey K. Sato,[1] and Rebecca G. Ong[3]

Great Lakes Bioenergy Research Center, [1]University of Wisconsin-Madison, Madison, WI, [2]Michigan State University, East Lansing, MI, [3]Department of Chemical Engineering, Michigan Technological University, Houghton, MI.

https://www.glbrc.org

Project Goals: Feedstocks can vary widely in terms of their biomass yields and quality characteristics (chemical composition, moisture content, etc.), which can affect microbial biomass-to-biofuel conversion efficiency. Previously, we found that elevated levels of lignocellulose-derived inhibitors in hydrolysates produced from pretreated switchgrass harvested from a drought year greatly inhibited the growth of *Saccharomyces cerevisiae*. We then expanded our studies to five different feedstocks, including two annuals (corn stover and energy sorghum) and three perennials (switchgrass, miscanthus, and mixed prairie), and evaluated the variability in microbial response and fermentation performance related to levels of lignocellulose-derived inhibitors. We found that the process ethanol yield varied among these five feedstocks during the microbial fermentation process using both yeast and bacterium (*Zymomonas mobilis*). However, both biomass quality and biomass yield affected the field-scale ethanol yields, the amount of ethanol produced per acre. Biomass quality was the main driver of the ethanol yields for high yielding crops, such as miscanthus. Biomass yield was the main driver for the ethanol yields for low-productivity crops, such as mixed prairie. Our results suggest that a lignocellulosic refinery may use a variety of feedstocks with a range of quality without a major negative impact on ethanol yields.

A variety of biomass feedstocks could be used for producing bio-based chemicals and biofuels. However, refineries that convert biomass into fuels often rely on just one feedstock. It would greatly benefit refinery operation if more than one feedstock can be used. To investigate how the diversity of potential biofuel cropping systems and feedstock supply might affect process and field-scale ethanol yields, we processed and experimentally quantified ethanol production from five different herbaceous feedstocks: two annuals (corn stover and energy sorghum) and three perennials (switchgrass, miscanthus, and mixed prairie). We previously found that *Saccharomyces cerevisiae* were significantly impaired for anaerobic fermentation of hydrolysates produced from pretreated switchgrass harvested from a drought year compared to non-drought years, as well as from pretreated corn stover from the same drought year. To avoid interannual variability, we have used lignocellulosic hydrolysates derived from five different plant feedstocks harvested from the same location and timeframe, as well as pretreated in an identical manner. This allowed us to identify feedstock-specific differences. We studied how the

levels of lignocellulose-derived inhibitors varied in these different feedstock hydrolysates and how their impact on microbial response and fermentation performance using both a yeast (*Saccharomyces cerevisiae* Y128) or a bacterium (*Zymomonas mobilis* 8b). Overall, the process ethanol yield showed some variability across years and feedstocks. A low process ethanol yield for corn stover was determined to result from inhibition of xylose utilization by unusually elevated levels of hydroxycinnamates (*p*-coumaric and ferulic acids) in the untreated biomass and their amide derivatives in the resulting hydrolysates. However, the field-scale ethanol yield from each feedstock was dependent on both biomass quality and cropping system productivity. Biomass yield had a greater influence on the ethanol yield for low-productivity crops, while biomass quality was the main driver for ethanol yields from high-yielding crops. Most feedstocks fall within a similar range of process ethanol yield, particularly for the more resistant strain *Z. mobilis* 8b. This supports the claim that the refinery can successfully diversify its feedstock supply, enabling many social and environmental benefits that can accrue due to landscape diversification.

References

1. Ong, R. G., A. Higbee, S. Bottoms, Q. Dickinson, D. Xie, S. A. Smith, J. Serate, E. Pohlmann, A. D. Jones, J. J. Coon, T. K. Sato, G. R. Sanford, D. Eilert, L. G. Oates, J. S. Piotrowski, D. M. Bates, D. Cavalier, and Y. Zhang. 2016. Reaction products of drought stressed bioenergy crops inhibit microbial biofuel production. Biotechnology for Biofuels. 9:237.
2. Y. Zhang, L. G. Oates, J. Serate, D. Xie, E. Pohlmann, Y. Bukhman, S. D. Karlen, M. K. Young, A. Higbee, D. Eilert, G. R. Sanford, J. S. Piotrowski, D. Cavalier, J. Ralph, J. J. Coon, T. K. Sato, and R. G. Ong. 2018. Diverse lignocellulosic feedstocks can achieve high field-scale ethanol yields while providing flexibility for the biorefinery and landscape-level environmental benefits. GCB Bioenergy 10:825-840.

This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-SC0018409 and DE-FC02-07ER64494).

Principles for Successfully Engineering Microbial Community Functioning

Michaeline B.N. Albright[*1] (malbright@lanl.gov) & **John Dunbar**[1]

[1]Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico

https://www.lanl.gov/science-innovation/science-programs/office-of-science-programs/biological-environmental-research/sfa-microbial-carbon.php

Project Goals

1) Test the relative importance of the environment, inoculation frequency, inoculation dose, the resident consortia, and the introduced consortia in driving microbial composition and functioning after microbial community introductions
2) Establish predictive links between microbial composition and functioning, in particular with relation to carbon flow during decomposition

Abstract

With increasing frequency, humans are manipulating microbial communities, introducing microbial consortia to new locations to achieve desired functional outcomes. Examples include modification of human gut flora for better human health, bioreactor microflora to improve waste degradation, and agricultural soil microbes for pest and disease control and/or increased carbon sequestration. Feasibility depends on the predictable establishment, persistence, and performance of inoculated microbial communities. However, the parameters for successfully engineering microbial community functioning are not well understood. Consequently, manipulation of microbial consortia continues to be a trial-and-error endeavor with low success rates. Here, we targeted the application of microbial control of plant litter decomposition, a focus of the LANL SFA program in Microbial Carbon Cycling, which aims to inform climate modeling and enable carbon management in terrestrial ecosystems. In this study we used a high-throughput laboratory microcosm experiment to test the relative importance of the environment, inoculation frequency, inoculation dose, the resident consortia, and the introduced consortia in driving microbial composition and functional outcomes after microbial community introductions. We found that fungal and bacterial introduction dynamics differ, where resident consortia played a larger role in shaping bacterial communities, while introduced consortia were more important in fungal community assembly. In terms of functional outcomes, CO_2 production was driven by dose, as well as both resident and introduced consortia, but not their interactions, while dissolved organic carbon (DOC) production was driven primarily by the environment and to a lesser extent by dose, and invader consortia. In addition, greater CO_2 production was correlated with increased bacterial richness and decreased fungal richness. Overall, identifying general ecological principles surrounding the establishment and resilience of introduced microbial communities can be used to advance applications in engineering microbial communities.

Funding Statement: This work was supported by the U.S. Department of Energy Biological System Science Division, through a Science Focus Area Grant (2015SFAF260 and 2019SFAF255)

and by Los Alamos National Laboratory through a Laboratory Directed Research and Development Director's Postdoc Fellowship to MBNA.

Role Of Geographic Scale In Likelihood Of Microbial Driven Variation In Litter Decomposition

Michaeline B.N. Albright[1], Jason Gans[1], Danielle Marias[1], Renee Johansen[1], La Verne Gallegos-Graves[1], **John Dunbar**[1]

[1]Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico

https://www.lanl.gov/science-innovation/science-programs/office-of-science-programs/biological-environmental-research/sfa-microbial-carbon.php

Project Goals:

The LANL SFA program in Microbial C Cycling aims to inform climate modeling and enable carbon management in terrestrial ecosystems. To achieve these aims, our program develops and uses community genomics approaches to discover microbiological processes that control carbon storage and release in temperate biome soils.

Abstract

In terrestrial ecosystems, stochasticity in the assembly of surface litter decomposer communities is widely believed to shape community composition. The compositional variation may drive functional variation, resulting in different patterns of carbon flow from decomposing litter. Is this important for climate feedbacks? The importance depends foremost on 1) the magnitude of functional variation and 2) the likelihood of substantial functional variation to occur within ecosystems.

Here, we examined the likelihood of microbial driven variation in surface litter carbon (C) flow as a function of geographic scale. Our null hypothesis is that stochastic assembly of decomposer communities creates only minor functional variation (e.g. *a few* percent difference in carbon flow). Consequently, *substantial* functional variation among decomposer communities is likely to be found only among communities over large geographic scales (e.g. >100km), where climate and ecosystem gradients can create persistent functional differences between distant microbial communities.

We performed a test of this hypothesis with a collection of over 400 soil samples from locations representing varied geographic scales (meters to 1000km). We suspended the soil microbial communities suspended in water, transferred aliquots to laboratory microcosms with sterile plant litter, and measured carbon flow (CO_2 and DOC) arising from 45 days of decomposition. We present the likelihood of substantial functional variation among communities as a function of the original distance between the communities, ranging from <1cm (replicate microcosms derived from the same gram of soil) to >1000km.

This work was supported by grants 2015SFAF260 and 2019SFAF255 from the OBER Genomic Sciences program.

Microbial lineages linked to decomposition outcomes across multiple litter types

Renee Johansen[1] (Rjohansen@lanl.gov), Michaeline Albright[1], Deanna Lopez[1], La Verne Gallegos-Graves[1], Cheryl Kuske[1], Andreas Runde[1], Thomas Yoshida[1], **John Dunbar**[1]

[1]Los Alamos National Laboratory, Los Alamos, NM

Project goals:
Determine which bacterial and fungal community members are consistently important to carbon fate in soils during the decomposition of leaf litter. This work was undertaken as part of the LANL Genomic Science SFA, which aims to inform carbon modeling and enable carbon management in terrestrial ecosystems. For more information about the SFA, see here: https://www.lanl.gov/science-innovation/science-programs/office-of-science-programs/biological-environmental-research/sfa-microbial-carbon.php

During plant litter decomposition, a portion of photosynthetically fixed plant carbon is released to the atmosphere, and a portion is retained in soils. An increase in the fraction released as CO_2 may exacerbate global warming, whereas an increase in soil storage may create a carbon sink. A relationship between carbon fate and microbial community composition has been frequently postulated, but disentangling the impacts of community composition from co-varying abiotic factors and litter quality is challenging. Over the last two years, we have presented evidence from experiments with over 800 decomposition microcosms demonstrating that microbial composition can drive substantial differences in carbon flux patterns, at least in the early phase of plant litter decomposition, even when environmental conditions are constant. Using dissolved organic carbon (DOC) abundance as an indicator of carbon flow, we define two contrasting outcomes – high versus low DOC accumulation – and explore the community features linked to each outcome. Here, we identify bacterial and fungal lineages that show consistent links with the two decomposition outcomes among three litter types. Although the litter types generally selected different decomposer communities, numerous bacterial and fungal lineages consistently changed in average relative abundance across all three litters in comparisons of 'high DOC' vs 'low DOC' communities. This work is a first step towards our goal of identifying universal community features that can influence carbon cycling in disparate environments.

This work is supported by grants 2015SFAF260 and 2019SFAF255.

Dissolved organic carbon (DOC) abundance links to microbial and DOC composition in soil

Danielle E. M. Ulrich[1]* (daniellem@lanl.gov), Malak Tfaily[2], Jaron Thompson[3], Brian Munsky[3], Jason Tyoda[4], **John M. Dunbar**[1]
[1]Bioscience (B-11), Los Alamos National Laboratory, Los Alamos NM, USA
[2]Department of Soil, Water and Environmental Science, The University of Arizona, Tucson AZ, USA
[3]Chemical and Biological Engineering, Colorado State University, Fort Collins CO, USA
[4] Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland WA, USA

https://www.lanl.gov/science-innovation/science-programs/office-of-science-programs/biological-environmental-research/sfa-microbial-carbon.php

Project Goals: The project goals are to 1) determine if there are dominant compounds underlying patterns in DOC abundance, and 2) investigate the microbial taxa governing DOC composition.

Carbon flow from soil microbial decomposition processes significantly influences climate. Soil microbial decomposition of litter results in CO_2 released to the atmosphere contributing to global warming and dissolved organic carbon (DOC) potentially sequestered in soil. Therefore, manipulation of microbes during decomposition can help us increase C storage. However, controls on decomposition processes and C partitioning to each of these fluxes (CO_2 and DOC) are not completely understood. We know microbial composition alters DOC abundance and composition, yet we do not know 1) if there are dominant compounds underlying patterns in DOC abundance, and 2) the microbial taxa governing DOC composition. Here, we address both of these knowledge gaps. We extracted natural microbial communities from 206 soil samples from the Southwestern United States, inoculated them on plant litter in homogeneous laboratory microcosms, and compared cumulative carbon flow into two divergent pools (CO_2 and DOC) after 44 days. We sequenced the microbial communities using an Illumina MiSeq sequencer at the Los Alamos National Laboratory (LANL) and characterized DOC composition using Fourier transform ion cyclotron resonance mass spectrometry (FTICR MS) and nuclear magnetic resonance (NMR) spectroscopy at the Environmental Molecular Sciences Laboratory (EMSL). The combination of these state-of-the-art tools provides unprecedented detail into both microbial and DOC composition. By doing so, we reveal types of compounds driving DOC abundance and the types of microbes associated with these compounds.

Funding Statement: The information presented in this poster was supported by the U.S. Department of Energy Biological System Science Division, through a Science Focus Area Grant (2015SFAF260 and 2019SFAF255).

High-throughput Chemical Imaging for Optimizing Biofuel Synthesis Using Synthetic Biology

Fengyuan Deng[1], Jean-Baptiste Lugagne[2], Nathan Tague[2], Haonan Lin[1], Wilson Wong[2], Ji-Xin Cheng[1,2], and **Mary J. Dunlop[2*] (mjdunlop@bu.edu)**

[1]Department of Electrical and Computer Engineering, Boston University, Boston, MA;
[2]Department of Biomedical Engineering, Boston University, Boston, MA

Project Goals: Develop a chemical imaging platform to provide a direct, high-throughput method for imaging biofuel production. Major goals include: (1) Establish a chemical imaging platform for rapid quantification of biofuel production. (2) High-throughput, multiplexed strain engineering enabled by chemical imaging. (3) Optimizing biofuel production using imaging for cell sorting and selection.

Recent advances in the fields of synthetic biology and metabolic engineering have resulted in an unprecedented ability to engineer genomes and design and build gene circuits for improving biofuel production. However, the ability to design and build these genetic variants has far outpaced methods for assessing biofuel production. Current methods for quantifying production rely on low-to-medium throughput approaches such as GC-MS, or indirect measurements such as those using biosensors coupled with fluorescent reporters. A direct, high-throughput method for imaging biofuel production *in vivo* has the potential to greatly advance our ability to rapidly design, build, and test strains for enhanced biofuel synthesis. This project addresses this gap by introducing a technology for directly measuring synthesis of biofuels in living cells.

We use a chemical imaging method called stimulated Raman scattering (SRS) microscopy. SRS uses photons to produce a vibrational spectrum on the microsecond time scale. Such high speed allows real-time chemical mapping of a sample at sub-micron diffraction-limited spatial resolution. This capacity is significant for quantifying biofuel synthesis because it can directly reveal production levels and further distinguish between different structures of chemical bonds. In this project we focus on fatty acids, which are biodiesel candidates and can serve as precursors to high value oleochemicals. We are developing a high-throughput platform for chemical imaging of biofuel production, which will be used in concert with multiplexed genome engineering and gene circuit design strategies to improve *E. coli* fatty acid production. Overall, our goal is to develop SRS imaging as a new technology for directly measuring chemical signatures in *in vivo* samples for the engineering and optimization of biofuel production strains. Here, we present results towards these goals including SRS imaging of fatty acid biosynthesis for different engineered strain backgrounds, and a multiplexed CRISPRi-based approach to strain generation. In addition, we will present advances in the imaging methodology that dramatically increase image acquisition speeds, enabling the potential for high-throughput imaging.

This research was supported by the Office of Science (Biological and Environmental Research), US Department of Energy.

Genetics and Genomics of Pathogen Resistance in Switchgrass

Serge Edmé[1]* (serge.edme@ars.usda.gov), Gautam Sarath[1], **Nathan Palmer**[1], Rob Mitchell[1], Satyanarayana Tatineni[1], Gary Yuen[2], Anthony Muhle[2], and Chowda Reddy[1]
[1]USDA-ARS Wheat, Sorghum and Forage Research Unit, Lincoln, NE; [2]Plant Pathology Department, University of Nebraska, Lincoln, NE

Project Goals: This project aims at applying genomic selection in switchgrass, a key bioenergy crop in the US, to define the genic regions causally linked to high biomass, ethanol yield, and durable resistance to fungal and viral diseases. Specific objectives are: (1) Understand the genetic and genomic bases of pathogen response in regionally-adapted upland and lowland switchgrass breeding populations with contrasting disease symptoms under field conditions, (2) Dissect the molecular underpinnings of the broad resistance or tolerance to pathogens in 'Kanlow' vs 'Summer', and (3) Discover the molecular differences that permit systemic viral infections in some switchgrass plants, but not in other genetically-related plants.

Genetics (classical) and genomics (genomic selection, RNAseq) approaches were applied to incorporate durable resistance to fungal (rust caused by *Puccinia emaculata* and *Uromyces graminicola*) and viral (mosaic caused by *Panicum mosaic virus* as PMV and satellite as sPMV) diseases in a switchgrass (*Panicum virgatum*) population being bred for bioenergy. Switchgrass feedstocks, developed for the lignocellulosic industry, will need to have higher biomass yield, higher quality (e.g. low lignin content), and durable disease resistance to be economically competitive and sustainable. This study covers genomic selection on two out of three generations of a switchgrass population derived from crossing two ecotypes ('Kanlow' as lowland and 'Summer' as upland) to leverage their differential performance in terms of biomass yield, quality, disease resistance, and winter survivability. Parental and halfsib progeny populations were screened for rust and mosaic in field and/or lab and phenotyped for biomass yield and quality traits. These populations were genotyped with the DArTseq platform to develop SNP (0, 1, 2) and in-silico (presence/absence) DArT markers. Rust and mosaic inoculation techniques were developed and successfully applied on switchgrass. Original populations (Kanlow and Summer) were sequenced with RNAseq to capture the gene expression profiles across serial time-points. Constructs of PMV and sPMV mosaic virus from different origins were made and tested alone and in combination first on proso millet and later on switchgrass to determine any potential synergistic action between the two.

Multivariate genetic analyses were performed on these populations to appraise the proportion of existing additive genetic variation to jointly improve these traits (i.e. biomass and ethanol yields, Klason lignin, and rust and mosaic, to estimate their phenotypic and genetic correlations, and to predict the breeding values of parents and progeny. Ample additive genetic variance for all traits exists in these populations and the genetic correlations are favorable to allow joint improvement in biomass yield, Klason lignin, ethanol yield, and disease resistance. Heritability estimates were

0.53 for dry matter yield, 0.36 for KL, 0.39 for ETOH, and 0.49 for mosaic ratings. Dry matter yield was genetically but negatively correlated with mosaic ratings (-0.44) and KL (-0.12ns), indicating that higher yielding genotypes were more resistant/tolerant to the virus. To maximize genetic gains, the selection index has been updated to include yield and disease traits by taking their variances and correlations into consideration. Including the pedigree (spanning 3 generations) into the analyses further improved the accuracy of prediction of the breeding values.

We were successful in developing SNP and in-silico markers for switchgrass: more precisely 99,941 in-silico DArT (dominant-type) markers and 66,021 SNP for a grand total of 159,962 markers (non-filtered), called across all chromosomes of the switchgrass genome (version 4). A large proportion of the markers fall within or near gene coding regions in the Pvi4 genome, based on overlapping the sequences with protein coding regions (CDS), non-coding gene regions (UTRs/introns), ±5kb up/downstream of a gene coding region, or 5-10kb up/downstream of a gene coding region. Multidimensional scaling plot indicates that Kanlow and Summer are genetically different with their hybrids located midway between the two parents and that the first 2 generations represent a panmictic (mixed) population, still with a large variation along the 2nd dimension.

The synergistic interaction between PMV and sPMV was investigated, using infectious cDNA clones of NE and TX isolates of PMV and clones of KS and TX isolates of sPMV. Both PMV-NE and TX elicited mild mosaic symptoms on proso millet and switchgrass whereas co-infection by PMV-NE+sPMV-K elicited severe mosaic, yellowing, and stunting symptoms, compared with moderate symptoms by PMV-TX+sPMV-TX. The severe symptoms caused by PMV-NE or PMV-TX with sPMV-K indicated that sPMV-K was the main contributor to the interaction. The genome sequences of sPMV-K and sPMV-TX differ by 11 nucleotides with four non-synonymous and three synonymous changes in the coat protein ORF. These genomic differences between sPMV isolates provide the basis for the differential synergistic interaction with PMV.

Analyses of the RNAseq data on the Kanlow and Summer populations, collected before and after infection by rust, indicates that expression was detected over a total of 34,381 genes, with 84% being differentially expressed (FDR < 0.05 and |log2FC| > 1) based on time or population. A subset of the Kanlow genes are not present in Summer and most of the gene-coding regions have relatively similar genomic resequencing coverage in both populations. Summer had 521genes, including 15 putative transcription factors, 8NB-LRRs, and 4 RLKs (receptor-like kinases). Kanlow (more resistant) had 1,402genes, with 23 putative transcription factors, 38 NB-LRRs, and 4 RLKs. Of the Kanlow-specific genes, approximately178 genes (including 7 NB-LRRs and 2 RLKs) were not found in the Summer genome. Substantial differences exist in the genetic regions controlling rust resistance in the two ecotypes.

Funding statement: This research project is supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. *DE-SC0016108.*

Identification and Characterization of Rust Effectors that Affect Host Immunity

Liang Hu[1], Lipika R. Pal[1], Ying Wu[1], James F. Parsons[1], Gen Li[1], Gary Coleman[1,2], Shunyuan Xiao[1,2] and **Edward Eisenstein**[1,3,*] (eisenstein@umd.edu)

[1]Institute for Bioscience and Biotechnology Research (IBBR), University of Maryland, Rockville, MD; [2]Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD; [3]Fischell Department of Bioengineering, University of Maryland, College Park, MD

Project Goals: An analysis of the *Melampsora larici-populina* genome reveals a large number of genes encoding candidate secreted effector proteins (CSEPs) that may play many roles in host infection. A screen was developed in tobacco leaves and poplar protoplasts for CSEPs that affect host immunity. Two subsets of effectors were targeted for analysis: one with significant homology to CSEPs in other fungal pathogens of plants, and another family, containing unique members with no similarity to any proteins in sequence databases. The results suggest that homologous CSEPs target conserved host factors and immune mechanisms to inflict disease. A platform was developed to further explore structure-function relationships within CSEP families. Gateway cloning of effector genes into a bacterial expression vector facilitated purification, and the use of fusions with a form of maltose binding protein modified for enhanced crystallization may greatly facilitate structure analysis.

Abstract: The genomes of fungal pathogens of plants contain genes for small, cysteine-rich secreted proteins that are specifically up-regulated for expression during infection, signifying a key role in host colonization and disease. Computational analysis of the 1,524 candidate secreted effector proteins (CSEPs) in *Melampsora larici-populina* shows that they belong to 807 structural families, including 600 single-member families, another 206 families that contain 2-36 members, and one family that has undergone a remarkable expansion to 117 members. To reveal the molecular interactions of poplar host factors and rust effectors a screen was developed to identify CSEPs that can suppress host immunity in tobacco and poplar. Two subsets of *Melampsora* CSEPs were targeted for analysis. The first group contained 160 *Melampsora larici-populina* CSEPs belonging to 67 structural families, all of which share significant homology among pathogenic rust, *Septoria* or powdery mildew fungi. The other subset was the unique 117-member family, whose members show virtually no similarity to any protein in sequence databases. Bootstrapped phylogenetic trees of this particular family were constructed using the maximum likelihood approach for both the DNA and protein sequences under best-fitted evolutionary models. The trees were used to identify family members common to branches on both trees that most likely diverged recently. This approach identified a modest subset of 10-15 relatively-distinct members for an analysis of their effect on host immunity.

The two subsets of CSEPs were expressed transiently in tobacco along with well-known 'autoactive' domains of *R* genes or an *R* gene-*AvrP* pair that promotes a hypersensitive response (HR) or HR-like cell death. Attenuation (or enhancement) of

salicylate levels stemming from changes in HR was measured quantitatively using LC-MS to assess the impact of *Melampsora larici-populina* CSEPs on mounting an immune response. For the first group, the screen led to the identification of 72 CSEPs that affect host immunity. These 'reactive' CSEPs belong to 49 of the 67 structural families, which suggests that homologous pathogen effectors may disrupt conserved plant immune responses and trigger disease. Transient expression of the reactive CSEPs in poplar protoplasts is being used to confirm the results in tobacco, and if validated, will enable tobacco-based large-scale screens for additional *Melampsora larici-populina* CSEPs that disrupt host factors and immunity, providing a foundation for engineering durable rust resistance in *Populus* spp. Transient expression of the second CSEP group will reveal the impact of the relatively unique members on host immunity, and shed light on how pathogen gene expansion and amplification can overcome host resistance.

The use of a recombination-based (Gateway) cloning system has enabled the facile transfer of CSEP genes from binary vectors for functional studies in plants to bacterial vectors for expression and purification. Production of *Melampsora larici-populina* CSEPs as fusions to a form of maltose binding protein that has been modified for enhanced solubility, affinity purification and containing amino acid substitutions that enhance crystallization has been piloted on three candidates. Initial results suggest promise of this approach for large-scale investigation of structure-function relationships among effector families important for plant-pathogen interactions.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0017886

Conversion of Natural Grassland into Biofuel Crops (Switchgrass, Panicum virgatum L.) is Associated with Reduction in Methane Consumption

Colin T. Bates,[1]*(colin.bates@ou.edu), Jialiang Kuang[1], Arthur Escalas,[2] Lauren Hale,[3] Yuan Wang,[4] Don Herman,[5] Ying Fu,[1] Renmao Tian[1], Daliang Ning[1], **Liyou Wu,[1] Jennifer Pett-Ridge,[6] Malay Saha,[4] Kelly Craven,[4] Jizhong Zhou,[1]**and **Mary Firestone**[7]

[1]The University of Oklahoma, Norman, OK; [2]MARBEC, University of Montpellier, Montpellier, France; [3]USDA-ARS, Parlier, CA; [4]Noble Research Institute, Ardmore, OK; [5]Lawrence Berkeley National Laboratory, Berkeley, CA; [6]Lawrence Livermore, National Laboratory, Livermore, CA; [7]University of California, Berkeley, CA;

Project Goals: Switchgrass (SG; *Panicum virgatum* L.) is a perennial C_4 grass native to the tallgrass prairies and a promising feedstock in the U.S. for bioenergy production. Capable of abundant biomass yield with minimal fertilizer or water, SG can survive on marginal soils, and even thrive once established. We hypothesize that successful establishment and sustainable cultivation of SG in marginal soils is in part enabled by beneficial plant-microbial interactions. We are investigating the succession of rhizosphere microbial communities, and ecosystem-scale effects of high- and low-performing SG plants grown in nutrient-limited soils in southern Oklahoma. The outcome of this research will provide a better genomic basis for SG cultivation in marginal soils, expand our knowledge of the interactions between soil microbiomes, plants and ecosystems, and ultimately guide efforts for translation into agronomic row crops.

SG has been shown to input carbon (C) into the soil, and thus has the potential to increase C sequestration and to improve soil productivity over long-term cultivation. However, to fully evaluate the sustainability of SG-based biofuel production, it is crucial to understand the impacts of SG establishment on biotic and abiotic characteristics of nutrient-poor soils. Here, we aim to characterize the ecosystem-scale consequences of SG cultivation in comparison with native annual grassland fallows at two sites in Oklahoma. We hypothesize that SG sustainability relates to improvements of soil quality mediated by its influence on the soil microbial communities and the activation of beneficial plant-microbe interactions. This will ultimately result in measurable effects on key ecosystem functions such as C sequestration and greenhouse gas (GHG) production. To understand the impact of SG establishment on soil characteristics and microbiology, we setup multi-year monthly field measurements of greenhouse gases fluxes (CO_2 and CH_4), soil chemical properties, and characterization of soil microbial communities' structure and function.

We selected two study sties in southern Oklahoma that are marginally suitable for row-crop cultivation. These two sites, a silt-loam soil close to the Red River (RR) and a clay loam soil (3rd Street Farm, 3S), share a low N and P nutrient availability but differ in other physio-chemical variables (pH, soil texture, soil organic matter (SOM) content). At both sites, two plots (27 x 22m) were established: a native annual grassland fallow (FL) and a SG plot contacting 500 genetically distinct seedlings (Alamo variety) in a honeycomb design. During the first two growing seasons following SG planting, these four plots were monitored monthly. We measured topsoil chemistry (plant-available P and N, and SOM content), GHG fluxes (CO_2 and CH_4), and microbial

community composition (16S rRNA genes). GHG concentrations were measured with a high sensitivity (parts-per-billion), high resolution (every 2s) cavity ring down spectrometer (Picarro G2508 analyzer). Five-minute resolution meteorological data such as soil temperature, air temperature, soil moisture, and precipitation were provided by the Oklahoma Mesonet network (http://climate.ok.gov/index.php/climate). We found that SG significantly increased plant-available P levels, SOM content, and soil C/N at the RR site when compared to the FL. No significant changes in soil chemistry were observed between SG and FL plots at the 3S site. Similar seasonal microbial successional patterns were observed for all plots, but community structures differed between sites and plot types. SG cultivation did not influence significantly annual CO_2 fluxes compared to the FL plots. However, SG significantly reduced the annual CH_4 consumption, potentially impacting ecosystem carbon balance. A machine learning approach was used to identify important factors for predicting soil trace gas dynamics and we found that soil temperature, soil moisture, and month of the year were the most influential variables. Further investigations are underway to elucidate linkages between the microbial communities (by time-series) and GHG emissions. These data are also being used to parameterize an ecosystem model.

Funding statement.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Number DE-SC0014079 to the UC Berkeley, Noble Research Institute, the University of Oklahoma, the Lawrence Livermore National Laboratory and the Lawrence Berkeley National Laboratory. Part of this work was performed at Lawrence Berkeley National Lab under contract DE-AC02-05CH11231 and at Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

Characterizing the Multitrophic Interactions that Mediate Carbon Flow in Soil and their Responses to Drought

Javier A. Ceja-Navarro (JCNavarro@lbl.gov)[1*], Abelardo Arellano[1], Leila Ramanculova[1], Katerina Estera[2], Peter K. Weber[4], Eoin L. Brodie[1,2], **Neo D. Martinez**[3], **Rosemary Gillespie**[2], **Jennifer Pett-Ridge**[4], **Mary Firestone**[2]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA; [2]University of California, Berkeley, CA; [3]University of Arizona, Tucson, AZ, [4]Lawrence Livermore National Laboratory, Livermore, CA.

Project Goals: Our project addresses both a fundamental understanding of carbon (C) cycling as mediated by multi-trophic interactions in the rhizosphere and the potential impacts of altered precipitation regimes on these interactions. Specifically, our work employs the use of stable isotopes to identify and quantify pathways of C-flow through multi-trophic interactions in the rhizosphere. Primary goals are to broaden knowledge of the roles played by multi-trophic interactions in terrestrial C cycling and to discover if drought alters the interactions and/or C cycling. This research will substantially expand our knowledge of soil microbial ecology, belowground food web, and terrestrial C cycling under a changing climate.

The objective of the research presented here is to illuminate the contribution of bacterial, archaeal, fungal, and other eukaryotic communities to carbon and nutrient cycling in soil by applying multi-omics approaches, stable isotope tracing, and field manipulations to the development of mechanistic network theory of material flows in the soil. In all terrestrial ecosystems, numerous populations of organisms such as protists, nematodes and arthropods interact with the soil's free-living and plant-associated microbes, defining biogeochemical and nutrient cycling processes. To predict the responses of C and nutrient cycling to environmental change, it is important to recognize that these environmental processes are the result of the interactions of multiple groups of organisms that in concert shape an ecosystem and its processes. We have optimized multiple approaches to characterize the belowground populations from all trophic levels from a Mediterranean annual grassland, while testing approaches to measure their contributions to C cycling.

Soil samples were collected from a field experiment where replicated plots were manipulated by the addition of different levels of water (50 and 100% local recorded precipitation), aiming to induce normal vs drought conditions. Arthropods were separated from soil samples using berlese funnels, quantified, and morphologically characterized. Nematode and protozoan populations were extracted from fresh soil samples using gradient centrifugation. Specimens were concentrated by filtration and their DNA extracted for amplicon sequence libraries. Sequences were analyzed and groups identified against our locally developed and manually curated dataset. Bacterial and fungal populations were characterized by amplicon sequencing from DNA extracted directly from soil samples. Both diversity and community structure of the analyzed populations showed changes in response to the different water treatments. After five months of differential water regimes, the diversity of nematodes and fungi remained unchanged between treatments, while their community structure changed with at least 10-15% of data variation explained by the type of treatment. Protozoa and bacterial communities, on the other hand, showed higher diversity in the soils under normal conditions without changes in community structure. Predicted functional roles of the

identified nematode and protozoan groups suggest direct interactions with the roots of plants in addition to bacterial and fungal predation. Co-occurrence analyses are being conducted to suggest possible interactions among groups detected from all trophic levels. Results from the co-occurrence analyses will inform laboratory experiments in which populations of organisms isolated from the field will be combined to further define their interactions and fine tune our food-web models.

Besides field water manipulations, our field experiment also included the incubation of *Avena fatua* plants with $^{13}CO_2$. Rhizosphere samples were collected and attached soil was separated, allowing us to physically separate nematode and protozoan populations. Isolated specimens were fixed and deposited onto a filter, imaged using electron microscopy, and ^{13}C levels of different organisms were screened using NanoSIMS isotopic imaging.

The approaches developed here will provide the foundation for molecular approaches to quantitative study soil trophic networks, and also have potential applications as diagnostic tools to identify and intervene for the early control of plant pathogenic organisms in bioenergy cropping systems.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Number DE-FOA-0001458 and DE-AC02-05CH11231 to UC Berkeley. Work at Lawrence Berkeley National Laboratory was performed under the auspices of the U.S. Department of Energy Contract No. DE-AC02-05CH11231, and work at Lawrence Livermore National Laboratory under U.S. Department of Energy Contract DE-AC52-07NA27344.

Unravelling Rhizosphere-microbial Interactions in the Rhizosphere of Alamo Switchgrass (*Panicum virgatum*) under Abiotic Stresses

Na Ding [1*](nding@noble.org), Nameer Baker [2], Yonatan Sher [2], Don Herman [2, 3], **Wolf Scheible** [1], **Michael Udvardi** [1], **Jennifer Pett-Ridge** [4], and **Mary Firestone** [2,3]

[1] Noble Research Institute, Ardmore, OK, [2] University of California, Berkeley, CA, [3] Lawrence Berkeley National Laboratory, Berkeley, CA, [4] Lawrence Livermore National Laboratory, Livermore, CA

Project Goals: Our project works towards a fundamental understanding of the key molecular mechanisms driving beneficial plant-microbial interactions in superior switchgrass genotypes adapted to a range of resource limitations. Plant-microbe interactions are examined during establishment to gain insight into how symbiotic and associative microbes improve plant performance and carbon stabilization in marginal soils. We will combine focused (single plant-microbe pairing) and 'community' systems biology approaches to examine the complex interplay among plants, microbes, and their physio-chemical environment.

In the rhizosphere, root exudation is a key process for C transfer into the soil, influencing the role of soil microbial communities in organic matter decomposition and in nutrient cycling. Root exudates have been shown to increase the number and activity of soil microbes and fauna found in the rhizosphere through a myriad of complex interactions. Soil microorganisms depend upon plant C and, in turn, potentially provide plants with nitrogen (N), phosphorus (P) and other mineral nutrients in part through decomposition of soil organic matter. We grew Alamo switchgrass (SG) in two types of greenhouse experiments to investigate how SG transcriptomes and exudates shape rhizosphere microbial community metagenomes and soil characteristics, as well as how these interactions are affected by abiotic stresses. The first set of experiments explore the gene-to-metabolite networks responsible for coping with N or P starvation by investigating the Alamo SG transcriptome, metabolome and physiology under a range of N and P supply conditions. The second experimental approach applied $^{13}CO_2$ stable-isotope labeling to trace Alamo SG photosynthate through root exudates and debris into the metagenomes of the microbial communities that consume them in 1-m soil profiles recreated in the greenhouse.

Sand-based: Alamo seedlings SG were grown in sand culture over a 4-week period under various N or P stresses with nutrient media containing KNO_3 (0.2-6.0 mM) and KH_2PO_4 (20-600 µM). Treatment with 6mM·N and 600 µM·P supply was served as control. Compared to control plants, moderately- (200 µM P_i) and severely-reduced (20 µM P_i) P-supply resulted in a decrease of biomass by 35% and 96%, respectively. Plants grown under severe P-limitation had comparable primary root length, but more and longer root hairs than P-replete plants. P-limitation induced the expression of genes involved in regulation of transcription, phospholipid degradation and phosphate transport in shoots and roots, phosphorylation in shoots, tryptophan and glycolipid biosynthetic processes in roots. Conversely, genes involved in phospholipid synthesis and secondary metabolism in shoots and roots, and photosynthesis, chlorophyll biosynthesis, starch and cell wall degradation in shoots were repressed. Switchgrass invested more organic acids (oxalate) and sugars (sucrose and trehalose) in roots relative to shoots in response to P-limitation. Shoots and roots showed distinct P-stress acclimation patterns at the transcriptional and metabolic levels, which reflect, in part, changing priorities for C-allocation to root growth and P-acquisition at the expense of C- and P-investment in shoot growth.

As expected, N levels strongly impacts plant biomass, root to shoot ratio and length of primary seminal roots. N starvation reduced total N and protein compared to control with replete N supply. Providing 2.0 mM N to switchgrass satisfied with the N requirement of plant growth, but reduced total N, total protein and modified metabolites profiles in relative to control. Metabolites composition was altered depending on plant organs and N availability. Some organic acids were elevated in N-stressed roots compared to control, such as malic acid, phosphoric acid, glycolic acid, hydracrylic acid, octanoic acid as well as quininic acid. Sugar metabolism was also influenced by N availability, with the increase of sucrose accumulation and slight decrease in glucose upon N starvation. The number of induced or repressed genes had negative relationship with the degree of nutrient limitation. Some high and low affinity transporters were highly induced during N starvation. In addition, genes encoding nitrate reductase (NR) and nitrite reductase (NIR) were also induced during N deficiency.

Soil-based: Alamo SG were grown in marginal soil with N and/or P amendments and under two watering regimes. We observed strong effects of root biomass in our soils, particularly in the +N/+P treatment where enhanced root biomass may have strongly affected soil water potential. Nitrogen fertilization also affected soil chemistry, as was evident in observed pH and dissolved organic carbon (DOC). Reduced pH observed in soils that received N amendments probably resulted from microbial ammonia oxidation of the coated urea fertilizer, a process that releases protons. However, enhanced [DOC] observed under nitrogen additions may be due to increased release of exudates by switchgrass roots.

We also observed significant differences between control and nutrient-amended soils in extracellular polysaccharide (EPS) production. EPS protects and binds soil microbial communities together, maintains favorable soil-water relations, and acts as an adhesive agent to increase soil aggregate stability. The highest soil EPS content and concentration of ^{13}C-enriched EPS was found in our +N/+P treatment, and was significantly correlated with observed root biomass. In addition, we did not observe significant differences between treatments in bulk soil microbial biomass as measured by phospholipid fatty acid analysis. However, the high mannose content of recovered EPS indicates that these polysaccharides are microbial in origin rather than being a direct root product. Notably, these high-DOC, high-EPS soils also contained more water-stable aggregates, indicating that enhanced root biomass and the resulting exudates and mucilage were potentially mediated through the microbial community to form a polysaccharide matrix that has enhanced a critical aspect of soil health in a marginal soil.

We observed a significant shift in density between the microbial DNA recovered from our $^{13}CO_2$-labeled mesocosms and that recovered from $^{12}CO_2$-labeled controls. This enables us to isolate microbial genomes within our recovered metagenomes that became significantly enriched in ^{13}C as a result of consuming ^{13}C-enriched photosynthate over the course of our labeling experiment. We aim to investigate whether these microbes contain traits that would facilitate the conversion of plant photosynthate into microbial EPS, as well as other traits that could be of potential mutualistic benefit to the plant providing C to the rhizosphere microbial community.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Number DE-SC0014079 to the UC Berkeley, Noble Research Institute, the University of Oklahoma, the Lawrence Livermore National Laboratory and the Lawrence Berkeley National Laboratory. Part of this work was performed at Lawrence Berkeley National Lab under contract DE-AC02-05CH11231 and at Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

Following Soil Carbon Cycling under Reduced Precipitation in a California Annual Grassland

Katerina Estera-Molina[1*] (rinaest@berkeley.edu), Mengting Yuan,[1] Ricardo Feliciano,[2] Donald Herman,[1] Ella Sieradzki,[1] Marissa Lafler,[3] Peter Weber,[3] **Javier Ceja Navarro**,[4] Steven Blazewicz,[3] **Jennifer Pett-Ridge,**[3] and **Mary K. Firestone**[1]

[1]University of California, Berkeley, California, [2] University of Puerto Rico, Mayaguez Puerto Rico, [3] Lawrence Livermore National Laboratory, Livermore, California, [4] Lawrence Berkeley National Laboratory, Berkeley, California

Project Goals: Our project addresses both a fundamental understanding of carbon (C) cycling as mediated by multi-trophic interactions in the rhizosphere and the potential impacts of altered precipitation regimes on these interactions. Specifically, our work employs the use of stable isotopes to identify and quantify pathways of C-flow through multi-trophic interactions in the rhizosphere. Primary goals are to broaden knowledge of the roles played by multi-trophic interactions in terrestrial C cycling and to discover if drought alters the interactions and/or C cycling. This research will substantially expand our knowledge of soil microbial ecology, belowground food web, and terrestrial C cycling under a changing climate.

The soil carbon (C) pool is larger than the atmospheric and terrestrial vegetation pools combined. Within soil C there exists a zone of dynamic biological activity and intense C turnover: the rhizosphere, the area of soil directly interfacing the root. Within the rhizosphere, there is a complex soil food web comprised of plant roots, fungi, bacteria, archaea, viruses and fauna that interact with each other and also utilize, transform, and transfer C from root exudates and root debris. Therefore, the food web interactions that occur within the rhizosphere influence whether C that is fixed via photosynthesis will return back to the atmosphere or remain in the soil for some period of time. While the concept of the soil food web is generally well recognized, the importance of food web pathways and their influence on soil C-cycling is much less known.

Climate models have predicted changes in the number of drought-impacted months in California, as in much of the rest of the world. Drought often reduces C efflux from terrestrial ecosystems, but the mechanisms underlying altered C cycling may be distinct for different periods after rainfall reduction. Understanding how reduced precipitation will impact food web dynamics and the subsequent effects on terrestrial C cycling is essential for anticipating future ecosystem responses. Accomplishing our research objectives requires: 1) Following the flow of carbon from the atmosphere through plants into different trophic levels of the soil food web in a field grassland ecosystem, and 2) Modifying the precipitation regime of the grassland ecosystem,

We hypothesize that under reduced precipitation, a higher proportion of plant derived C will be carried into soil by arbuscular mycorrhizal fungi (AMF), and mineralization via phage and faunal interactions will also be affected. To test these hypotheses, we constructed 16 “trenched”, rainout plots in an annual California grassland located at the Hopland Research and Extension Center (HREC) in Hopland, California. These plots contain a community of mixed California annual grasses and forbs. Precipitation inputs to these plots have been manipulated so that half receive a

50% reduction of the 65-year rainfall average, and the other half receive the full average rainfall amount. Two in-field $^{13}CO_2$ pulse labeling events were conducted in the Spring of 2018 to trace the pathways of C after it was fixed by the plant and delivered belowground in the form of exudates and fresh roots, and consumed by rhizosphere residents. Ecosystem respiration from the plots was estimated by measuring the overnight CO_2 accumulation rate in the labeling chambers during both labeling events. Soils and plants were sampled at multiple time points following ^{13}C pulse labeling to track the location and persistence of the recently fixed C. DNA and RNA have been extracted from rhizosphere soil for stable isotope probing (SIP) enabled-omic analyses. Together, these results will help elucidate how C travels through the rhizosphere food webs.

During the growing season, rainfall reduction significantly decreased soil moisture by 35% and ecosystem respiration by 20%. There was a significant difference in the amount of enrichment of rhizosphere-influenced soil between the $^{13}CO_2$ and $^{12}CO_2$ labeled samples. The $^{13}CO_2$ samples had a $\delta^{13}C$ enrichment of 67‰ higher compared to the $^{12}CO_2$ samples.. While ^{13}C enrichment in rhizosphere-influenced soils was not affected by precipitation treatment, rhizosphere soil DNA was significantly ^{13}C enriched, with an atom percent excess ranging from 22-30%. Individual organisms (nematodes, AMF hyphae, protists) drawn from the root zone were also enriched, up to 3500‰. In sum, this large-scale multi-day field $^{13}CO_2$ labeling of California annual plants successfully labeled belowground communities to a degree that will enable us to track the flow and fate of root C into and through multiple soil communities.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Numbers DE-SC0010570 and DE-SC0016247 to MF at UC Berkeley. Work at LLNL was performed under the auspices of OBER award SCW1589 and the U.S. Department of Energy under Contract DE-AC52-07NA27344.

Performance of Switchgrass 'Alamo' Population and Selections for Sustainable Production on Marginal Soils under Low-Input Management

Malay Saha[1] (mcsaha@noble.org), Josh Barbour[1], Melissa McMahon[1], Yuan Wang[1], Na Ding[1], Erin Nuccio[2], Colin Bates[3], Lauren Hale[3], **Jennifer Pett-Ridge[2], Kelly Craven[1], Wolf Scheible[1], Michael Udvardi[1], Jizhong Zhou[3], Mary Firestone[4*]**

[1]Noble Research Institute, Ardmore, Oklahoma; [2]California Lawrence Livermore National Laboratory, Livermore, California; [3]University of Oklahoma, Norman, Oklahoma; [4]University of California, Berkeley, California.

Project Goals: Switchgrass (SG; *Panicum virgatum* L.) is a perennial C_4 grass native to the tallgrass prairies and a most promising feedstock in the U.S. for bioenergy production. Capable of abundant biomass yield with minimal fertilizer or water, SG can survive on marginal soils, and even thrive once established. We hypothesize that successful establishment and sustainable cultivation of SG in marginal soils is in part enabled by beneficial plant-microbial interactions. We are investigating the succession of rhizosphere microbial communities, and ecosystem-scale effects of high- and low-performing SG plants grown in nutrient-limited soils in southern Oklahoma. The outcome of this research will provide a better genomic basis for SG cultivation in marginal soils, expand our knowledge of the interactions between soil microbiomes, plants and ecosystems, and ultimately guide efforts for translation into agronomic row crops.

To study plant-soil microbiome characteristics of switchgrass growing in 'marginal' nutrient or water-limited soils, we established field plots at two research farms, both remnants of the Dust Bowl Era in Oklahoma. Seeds of the 'Alamo' switchgrass cultivar were germinated in petri dishes and seedlings were transferred into 25-cm tall cones. 500 randomly selected 58-day old seedlings were then transplanted into a well-prepared field plot at the Red River farm of the Noble Research Institute (NRI) on May 13, 2016. Plants were arranged in a Honey-comb design with 1 m spacing between plants. Another 500 seedlings were planted into the NRI 3rd Street farm on June 23, 2016 following the same protocol. No chemical fertilizer or pesticides were applied in either location, and only hand weeding was used to keep the plots weed free. Data on different morphological characteristics were collected throughout 2016, 2017 and 2018. Biomass was harvested after senescence each year.

Spring green-up started as early as Feb. 17 and continued up to March 3, 2017, which is quite early compared to other years. In 2018, plants start regrowing between March 7 and April 3. All the plants flowered between June 28 and July 23, 2017. In 2018, flowering started a week early (June 21) but continued until July 29. Plants started regrowing and flowered almost at the same time in the two locations. Continuous height increment was observed throughout the growing season. Wide variation was observed among the plants for plant height. In 2016, the plants were as short as 45 cm, but grew up to 182 cm. A significant increase in plant height was observed in the second year of growth (2017). The height of the plants ranged between 48-301 cm with an average of 244 cm. Average plant height in 2018 was 249 cm. While plants were taller at the Red River than the 3rd Street farm in the first two years, plants were 12 cm shorter in the Red River farm compare to those in 3rd Street farm in 2018. Significant variation was observed for biomass production among the Alamo genotypes at both the locations. In 2016, the establishment year, biomass weight of the plants varied from 0.03-1.08 kg $plant^{-1}$ with an average of 0.46 kg $plant^{-1}$. However, in 2017, plant

biomass yield varied from 0.18-4.07 kg plant^{-1} with an average of 1.42 kg plant^{-1}. Compared to the establishment year (2016), a threefold increase in biomass yield was recorded in the following year. It has not yet been possible to harvest the plants after the 2018 growing season due to excessive wet conditions.

On the basis of biomass yield in 2016, high (n = 12) and low (n = 8) performing genotypes were selected from both locations. Clonal ramets from each selected high- and low-performing genotype were prepared at the NRI greenhouse. Four replicates of each of the 40 genotypes were established in field experiments at both the 3rd Street and Red River locations in May, 2017 following the same experimental design. Morphological characteristics of the selections were collected throughout the 2017 and 2018 growing seasons. Average plant height of high- and low-selections were 179 vs. 161 cm in June 27, 192 vs. 174 cm in July 27, and 217 vs. 205 cm in October 17, 2018. Some of the low-performing genotypes selected on the basis of establishment year performance did not continue to be low performing in the replicated trial. Biomass was harvested after a killing frost in December, 2017. Biomass yield of low selections varied between 0.21-0.76 kg plant^{-1}, whereas those of high selections varied between 0.38 to 1.12 kg plant^{-1}.

Biomass yield from the establishment year was used to identify high- and low-performing genotypes at both locations. To measure total belowground biomass, three high and three low performing individuals from Red River and two from each group at 3rd Street were excavated to ~1.5 meter depth. Physio-morphological characteristics were measured from each plant. Significant variations between high- and low-performing genotypes for root, shoot, and crown weight were observed across locations. Better shoot and crown growth were recorded in 3rd Street but root growth was higher in the Red River site. Variation for morphological traits between the high and low biomass genotypes was much more pronounced in 3rd Street genotypes compared to those from the Red River farm.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Number DE-SC0014079 to the UC Berkeley, Noble Research Institute, the University of Oklahoma, Lawrence Livermore National Laboratory and Lawrence Berkeley National Laboratory. Part of this work was performed at Lawrence Berkeley National Lab under contract DE-AC02-05CH11231 and at Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

Effects of Switchgrass (*Panicum virgatum* L.) on Deep Carbon Pools in Marginal Lands

Jialiang Kuang[1]* (kjialiang@ou.edu), Colin T. Bates[1], Amrita Bhattacharyya[2,3], Erin Nuccio[2], Arthur Escalas[4], Lauren Hale[5], Yuan Wang[6], Don Herman[3], Karis McFarlane[2], **Liyou Wu**[1], **Peter S. Nico**[3], **Malay Saha**[6], **Kelly Craven**[6], **Jennifer Pett-Ridge**[2], **Jizhong Zhou**[1], and **Mary Firestone**[7]

[1]The University of Oklahoma, Norman; [2]Lawrence Livermore, National Laboratory, Livermore, CA; [3]Lawrence Berkeley National Laboratory, Berkeley, CA; [4]University of Montpellier, Montpellier, France; [5]USDA-ARS, Parlier, CA; [6]Noble Research Institute, Ardmore, OK; [7]University of California, Berkeley, CA;

Project Goals: Switchgrass (SG; *Panicum virgatum* L.) is a perennial C4 grass native to the tallgrass prairies and a promising feedstock in the U.S. for bioenergy production. Capable of abundant biomass yield with minimal fertilizer or water, SG can survive on marginal soils, and even thrive once established. We hypothesize that successful establishment and sustainable cultivation of SG in marginal soils is in part enabled by beneficial plant-microbial interactions. We are investigating the succession of rhizosphere microbial communities, and ecosystem-scale effects of high- and low-performing SG plants grown in nutrient-limited soils in southern Oklahoma. The outcome of this research will provide a better genomic basis for SG cultivation in marginal soils, expand our knowledge of the interactions between soil microbiomes, plants and ecosystems, and ultimately guide efforts for translation into agronomic row crops.

This project aims to improve our understanding of the effects of long-term establishment of SG on soil carbon storage and nutrient availability. Specifically, we focus on how the accumulation of root biomass (i.e., the continuous input of slow C via root litter) in deep soil layers influences heterotrophic respiration by mediating the structure and functionality of microbial communities, which will consequently regulate the soil carbon balance. To this end, we attempt to address the differential responses (i.e., the dynamics of their decomposition/respiration along a depth profile) to plant residual carbon inputs (slow C) between long-term selected microbial communities from deep and shallow root systems.

To understand the impact of long-term cultivation of SG on carbon storage in marginal soils, we collected deep core soils (0-150cm) from two N/P poor sites (RR and PDF sites) with long-term (10yr) cultivation of perennial SG (deep-rooted system) and neighboring plots cultivated with annual row crops for more than 30 years (shallow-rooted system). We measured both soil C content and residence time (^{14}C) in these systems. Our initial results indicate that soils under 10 year SG cultivation have higher %C at all depths than soils that have been cultivated with annuals. Radiocarbon suggest that the marginal soils cultivated with SG have substantially younger C at all depths, indicating measurably higher soil carbon input with more modern ^{14}C. Ongoing measurements of soil chemistry and density fractionation will further illustrate the potential positive benefits of SG establishment on soil carbon and nutrient stocks.

To further explore how the accumulation of root biomass in deep soil layers influences the respiration of soil heterotrophic microorganisms and consequently regulates the soil carbon balance, we conducted a priming experiment using these deep core soils. We found a 60-day incubation with ^{13}C-labeled plant (oat) material significantly increased the microbial biomass (estimated by PLFA) in deeper soil layers. Adding

^{13}C plant material also significantly increased the carbon substrate (labile-C) utilization (measured by BioLog). Significant shifts in community structure were observed along a depth profile (16S rRNA gene high-throughput sequencing) yielding an increase of relative abundance of fast-growing bacteria and potential lignin degraders, especially in the shallow-rooted system, suggesting differential community level C degradation dynamics between the two root systems. Further measurements of soil microbial respiration (total CO_2 and ^{13}C-CO_2) are underway to elucidate differential response to external C input (slow C from root biomass) and priming effects between the two root systems. Our research will provide new insights into the prediction of deep soil carbon accrual after SG cultivation by comprehensively considering the dynamics of soil physiochemical properties and different carbon pools that are potentially mediated by microbial communities.

Funding statement.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Number DE-SC0014079 to the UC Berkeley, Noble Research Institute, the University of Oklahoma, the Lawrence Livermore National Laboratory and the Lawrence Berkeley National Laboratory. Part of this work was performed at Lawrence Berkeley National Lab under contract DE-AC02-05CH11231 and at Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

Diversity of Viruses in Soil: Do Interactions with Soil Organisms Impact Carbon Flow?

Alexa Nicolas[1]* (alexa_nicolas@berkeley.edu), Evan Starr[1], Erin Nuccio[2], Donald Herman[1, 3], Steve Blazewicz[2], Peter Weber[2], **Jennifer Pett-Ridge[2]**, Jillian Banfield[1,3],and **Mary Firestone[1, 3]**

[1]University of California, Berkeley, California; [2]Lawrence Livermore National Laboratory, Livermore, California; and [3]Lawrence Berkeley National Laboratory, Berkeley, California

Project goals: Understanding the flow and fate of carbon from roots into soil is critical to unravelling soil carbon sequestration. Our project explores the impact of phages on soil microbial communities and ultimately on the fate of carbon in soil. To realize the full extent of the impact of viruses on soil microbial communities and soil carbon flow we have synthesized meta-'omic techniques. These approaches enable capturing a broad diversity of viruses: active viruses revealed through stable isotope genome-resolved metagenomes, RNA viruses mined from soil metatranscriptomes, and filtered double and single-stranded DNA viruses from soil viromes. Additionally, we are optimizing imaging of viral-like particles to pair dynamics quantified from fluorescence imaging with genomic abundances, and to quantify viral contribution to carbon turnover using nanoscale secondary ion mass spectrometry (NanoSIMS). Taken together, analysis of soil viral diversity and dynamics with stable isotope probing will allow us to connect viral populations to carbon flow from roots into soil, microbial community dynamics and the soil food web, with important implications for soil health and agricultural productivity.

Disentangling the complex network of inter-organismal interactions is required to fully understand and manage carbon flow through soil. While viruses outnumber microbial cells in soil, recent analysis of CRISPR spacers suggests upwards of 90% of viruses infecting bacteria (phage) have not been identified. We identified novel viral populations in soil which may impact soil carbon flow, through interactions with and lysis of hosts. Little is known about environmental RNA viruses, especially in soils. By assembling 48 metatranscriptomes from different soil zones – root-influenced, litter-influenced, and bulk soil (no addition) – we resolved highly diverse RNA viral communities from multiple soil habitats. Viruses changed over time and by environmental condition; Narnaviridae and Leviviridae were the most diverse and dominant viral families. Much of the observed viral diversity seemed to parasitize either fungi or Proteobacteria. A mechanistic understanding of how RNA viruses affect host metabolism and impact carbon flow remains to be studied.

Simultaneously, we identified DNA viruses produced through host infections, and recently active DNA viruses. We hypothesize that one role of phage in soil may be to lyse host cells allowing for host material to become part of the dissolved organic carbon pool, which can be respired back to the atmosphere or stabilized on mineral surfaces – similar to the marine "phage shunt." Toward understanding a soil phage shunt we have isolated viral-like particles from soil, which represent viruses resulting from host infections and potential lysis events and thus may serve as a proxy for microbial death. We have also resolved active soil phage genomes through stable isotope probing. From the resulting genomes of these efforts we have begun to understand which viruses are present in soil, what are their genomic contents and how this may impact soil carbon. From viral metagenomes (viromes) derived from isolated viral-like particles we have tapped a vast

unexplored, unannotated domain of soil microbial communities. Using recently optimized methods, we were able to isolate a "viral fraction," smaller than 0.22μm and larger than 100kDa, from field soil growing *Avena barbata* (wild oat plant), and extract nucleotides from this fraction for Illumina HiSeq sequencing. Based on analysis of terminase genes, these soil virome sequences are more related to each other than to previously identified viral genomes from NCBI. We are in the process of ground-truthing these viromes with relevant soil metagenomes.

To trace atmospheric carbon into soil microbes and phages, we grew *Avena spp.* plants under $^{13}CO_2$ and collected rhizosphere soil, and bulk soil (soil not associated with roots) at weeks 0, 6, and 9. Following extraction of DNA from the samples, we performed density-gradient centrifugation, yielding DNA samples with a range of ^{13}C label: unlabeled, partially labeled, and heavily labeled DNA. DNA from the separated fractions were then sequenced, and the sequence data were assembled, binned for host genomes, and phage contigs were identified. This enabled us to identify active phages in the rhizosphere—i.e. phage incorporating newly added ^{13}C label. Of the viral genomes resolved from this stable isotope probing dataset, we identified 13 complete phage genomes. For two of the phages we resolved near complete genomes of their hosts, which allowed us to examine specific anti-phage defense systems in context, including CRISPR and restriction modification systems. We observed different populations of phages that incorporated the label into their genomes as compared to non-labelled phages, indicating that the development of rhizosphere-competent bacterial consortia enabled the production of new (different) phage populations. These combined approaches to discovering and understanding soil viruses allow us to track phage population dynamics through time in relation to identified hosts and uncover a previously unseen wealth of viral genomes in soil.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Numbers DE-SC0010570 and DE-SC0016247 to MF at UC Berkeley and DE-SC10010566 to JB at UC Berkeley. Work at LLNL was performed under the auspices of OBER award SCW1589 and the U.S. Department of Energy under Contract DE-AC52-07NA27344.

Plant-microbe and microbe-microbe interactions mediate switchgrass sustainability: following rhizosphere microbial communities during switchgrass establishment

Yuan Wang[1], Javier Ceja Navarro[2], Erin Nuccio[3*], Renmao Tian[4], Travis Simmons[1], Katerina Estera-Molina[5], Christina Fossum[5], Jialiang Kuang[4], Colin Bates[4], Lauren Hale[4], Na Ding[1], Josh Barbour[1], Nameer Baker[5], Abelardo Arellano[2], Kateryna Zhalnina[6], Eoin Brodie[2], Trent Northen[6], Wolf Scheible[1], Michael Udvardi[1], Jennifer Pett-Ridge[3], Malay Saha[1], Liyou Wu[4], Jizhong Zhou[4], **Mary Firestone[5]** and Kelly Craven[1] (kdcraven@noble.org)

[1]Noble Research Institute, Ardmore, Oklahoma; [2]Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, California; [3]Lawrence Livermore National Laboratory, Livermore, California; [4]University of Oklahoma, Norman, Oklahoma; [5]University of California, Berkeley, California; [6]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California.

Project Goals: Switchgrass (SG; *Panicum virgatum* L.) is a perennial C_4 grass native to the tallgrass prairies and a most promising feedstock in the U.S. for bioenergy production. Capable of abundant biomass yield with minimal fertilizer or water, SG can survive on marginal soils, and even thrive once established. We hypothesize that successful establishment and sustainable cultivation of SG in marginal soils is in part enabled by beneficial plant-microbial interactions. We are investigating the succession of rhizosphere microbial communities, and ecosystem-scale effects of high- and low-performing SG plants grown in nutrient-limited soils in southern Oklahoma. The outcome of this research will provide a better genomic basis for SG cultivation in marginal soils, expand our knowledge of the interactions between soil microbiomes, plants and ecosystems, and ultimately guide efforts for translation into agronomic row crops.

In the soils surrounding roots (rhizosphere), biotic, chemical and physical drivers enrich for specific bacterial and fungal communities. These organisms can play multiple roles, and some may benefit plant productivity through enhanced nutrient acquisition, water uptake and/or pathogen suppression. We are investigating the composition, succession and function of rhizosphere microbial communities during SG cultivation in an effort to better understand the plant-microbiome interactions that enable plant survival and adaptation in marginal soils.

To study plant-soil microbiome characteristics of SG growing in 'marginal' nutrient or water-limited soils, we selected two research farms, both remnants of the Dust Bowl Era in Oklahoma. Red River Farm lies near the border of Oklahoma and Texas, and has a silt loam soil low in NO_3-N and organic matter. Third Street Farm, has a clay loam soil and relatively low phosphorus availability. Five hundred genetically distinct Alamo seedlings were planted into each field in May-June of 2016. Other than hand weeding during the summer, no management, water, or nutrients were supplied to the fields. Thirty plants were randomly selected for monitoring of rhizosphere community succession, plant performance and soil physio-chemical characteristics including: gravimetric moisture, pH, and NO_3/NH_4. For each plant, rhizosphere and bulk soil were sampled over the first growing season (Year 1) at 5 time points (T1-T5): early and late vegetative growth, reproductive growth, maximal growth, and senescence. Due to plant genetic diversity, plant biomass was highly variable within the plots, and the selected thirty plants are representative of this diversity. This provided the basis for assessing correlations between plant biomass production and plant associated microbial communities—assessed by amplicon sequencing of marker genes specific to bacteria (16S), fungi (ITS), and soil protozoa (18S).

Overall, microbial and protozoan communities of the silt loam soils (Red River) exhibited higher alpha-diversity (species richness) relative to clay loam soils (Third Street). In contrast, clay loam soils (Third Street) had more variation in microbial community composition within plots, consistent with greater variability observed in soil physical characteristics. In both sites, beta-proteobacteria were significantly enriched in the rhizosphere soil relative to bulk soil. Both microbial and protozoan communities were less diverse in the rhizosphere soil than in corresponding bulk soil, indicating a selective plant effect at multiple trophic levels. Site, habitat type (rhizosphere vs. bulk) and plant developmental stage had significant effects on soil microbial and protozoan community composition. The site effect was greater than the rhizosphere effect, followed by effects of plant developmental stage. Soil nutrient availability (P, K, pH) was the most significant environmental driver for microbial community assembly, and its influence increased in magnitude over the growing season particularly in rhizosphere communities. The rhizosphere communities from different trophic levels dynamically changed over the growing season, where the early successional phase (first 3 months after planting) was most distinct from bulk soil. A few microbial OTUs were found to be significantly correlated with plant biomass production at the early phase. Although plant biomass was correlated with both rhizosphere and bulk communities, plant height was more significantly correlated with rhizosphere communities. In conclusion, switchgrass rhizosphere communities are highly correlated with plant biomass production, and early establishment phase (1-3 months) may be an important time frame for the microbial stimulation of plant biomass production.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Number DE-SC0014079 to UC Berkeley, Noble Research Institute, the University of Oklahoma, Lawrence Livermore National Laboratory and Lawrence Berkeley National Laboratory. Part of this work was performed at Lawrence Berkeley National Lab under contract DE-AC02-05CH11231 and at Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

Exploring Potential Interactions of Bacteria and Arbuscular Mycorrhizal Fungi Using Network Analyses of an Annual Grassland Soils

Mengting Yuan[1]*(maggieyuan@berkeley.edu), Anne Kakouridis[1], Shengjing Shi[2], Erin Nuccio[3], Jennifer Pett-Ridge[3], Jizhong Zhou[4] and **Mary Firestone**[1]

[1]University of California Berkeley, Berkeley CA; [2]AgResearch Ltd., New Zealand; [3]Lawrence Livermore National Laboratory; [4]University of Oklahoma, Norman OK

Project Goals: Soil surrounding plant roots is home to diverse biological activity and carbon (C) cycling. A substantial quantity of C moves from roots to the surrounding soil via mycorrhizal hyphae. A primary goal of our project is to understand and quantify how multi-trophic interactions mediate and control soil C dynamics; we focus here on one of the key components of belowground food webs, the interaction of arbuscular mycorrhizal fungi (AMF), bacteria, and roots. We applied co-occurrence network analyses to the bacterial and fungal communities sampled from a California annual grassland soil growing *Avena fatua* in a greenhouse experiment. Our goal was to identify covarying AMF and bacteria taxa adjacent to and distant from roots at different growth stages of the plant. We used patterns of covariation and coexclusion to identify potential AMF-bacterial interactions as a first step in elucidating interactions and their mechanisms.

AMF form associations with 80% of land plant species, facilitating the transfer of plant-derived C into surrounding soils, and increasing the volume of soil from which plants can access nutrients. The presence of AMF is known to correlate with changes in soil bacteria, and the interactions between plant, AMF, and bacteria likely play a significant role in mediating soil C dynamics. Yet less is known about the identities of associated AMF and bacterial species, the timing and location of interactions, and the mechanisms involved. Here we explore the co-occurrence patterns between soil fungi and bacteria using Random Matrix network analyses, expecting network analyses to provide novel information towards understanding multi-trophic interactions in soil food webs. In a greenhouse experiment with a highly replicated design, we studied both rhizosphere and bulk soils associated with different growth stages of the annual grass *Avena fatua*, under conditions mimicking two growing seasons in the relevant Mediterranean climate with a summer dry down between growing seasons. Sequences of ITS and 16S rRNA genes were used to investigate fungal, including AMF, and bacterial communities over the stages of plant growth.

Roots impacted both bacterial and fungal communities in the rhizosphere. Overall, rhizosphere communities had lower diversities and compositions distinct from those in bulk soils. As plants grew through both seasons, rhizosphere bacterial and fungal communities underwent dramatic succession and a decrease in diversity, while communities remained much less changed in bulk soil. Dry down (simulating summer) caused significant shifts in both rhizosphere and bulk community compositions, with a stronger impact on the diversity of bacteria than fungi. AMF detected in these soils also differed in composition between season and sampling locations. Interestingly, when compared to rhizosphere soil, the diversity of detected AMF was higher in bulk soil, where diversity increased across the two seasons, despite the dry down period.

Co-occurrence networks were constructed for both rhizosphere and bulk soils at different plant growth stages, including all pairs of fungi and bacteria whose abundances covaried. Using a random matrix theory-based detection method, levels of covariation significance were determined for each network. Twenty-two AMF OTUs from five genera, including *Glomus*, *Funneliformis*, *Claroideoglomus*, *Acaulospora*, and *Paraglomus*, were captured by the network. These AMF and their covarying bacteria formed larger networks that increased in size over time in bulk soil but not rhizosphere soil, and in the second season compared with the first. The increase in covarying AMF-bacteria pairs over time in bulk soil was consistent with the increase of AMF diversity in these soils.

We found more covarying AMF-bacteria pairs in bulk soil than rhizosphere soil. In the same system, covarying bacteria-bacteria and bacteria-fungi pairs were generally more frequent in rhizosphere soil where root exudates provided a source of labile C. There are several possible, complementary, explanations for the pattern of increased co-occurrence of AMF-bacteria pairs in bulk soil. As plants take up nutrients from the rhizosphere, a depletion zone is generally established directly around the root, leading AMF to extend hyphae beyond that zone to reach required nutrients. In fact, both AMF mycelium and spores may be more abundant in bulk soil than rhizosphere soil, paving the way for more interactions with bacteria in bulk soil. We found AMF taxa co-occurring with bacteria of the orders Myxococcales, Spinngobacteriales, and Xanthomonadales, including groups known to degrade biopolymers. It has been proposed that AMF exudates provide C to bacteria in bulk soil, possibly promoting decomposition by bacteria. AMF may then take up nutrients freed up by the decomposition. AMF spores have been reported to contain intracellular bacterial microbiomes which include bacteria from the order Burkholderiales; this bacterial order was found associated with AMF in our networks.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Awards Numbers DE-SC0010570 and DE-SC0016247 to UC Berkeley and University of Oklahoma. Work at LLNL was performed under the auspices of the U.S. Department of Energy under Contract DE-AC52-07NA27344 and work at Lawrence Berkeley National Laboratory under U.S. Department of Energy Contract No. DE-AC02-05CH11231.

Molecular basis of plant-microbe-soil interactions shaping microbial communities and ecosystem function in a Mediterranean grassland

Kateryna Zhalnina[1]* (kzhalnina@lbl.gov), Erin Nuccio,[2] Richard Allen White III,[3] Heejung Cho,[1,4] Benjamin P. Bowen,[1] Markus de Raad,[1] Kai Deng,[5] Ulas Karaoz,[1] Mary S. Lipton,[3] **Mary Firestone,[1,4] Jennifer Pett-Ridge,[2] Eoin Brodie,[1,4] Trent Northen[1]** (trnorthen@lbl.gov)

[1]Lawrence Berkeley National Laboratory, Berkeley, California; [2]Lawrence Livermore National Laboratory, Livermore, California; [3]Pacific Northwest National Laboratory, Richland, Washington; [4]University of California, Berkeley; [5]Joint BioEnergy Institute, Emeryville, California.

Project Goals: Our project addresses both a fundamental understanding of carbon (C) cycling as mediated by multi-trophic interactions in the rhizosphere and the potential impacts of altered precipitation regimes on these interactions. Specifically, our work employs the use of stable isotopes to identify and quantify pathways of C-flow through multi-trophic interactions in the rhizosphere. Primary goals are to broaden knowledge of the roles played by multi-trophic interactions in terrestrial C cycling and to discover if drought alters the interactions and/or C cycling. This research will substantially expand our knowledge of soil microbial ecology, belowground food web, and terrestrial C cycling under a changing climate.

Soil is the one of the most heterogeneous systems on earth where complex interactions exist between minerals, hydrology, chemistry and soil organisms. This heterogeneity creates diverse environments that often contrast in the type and availability of nutrients. Plants release substantial quantities of simple C molecules into the soil surrounding their roots, attracting bacteria with readily available resources. In contrast, the surrounding bulk soil represents an environment with significantly lower availability of small organic molecules. How microbes adapt to each of these environments and how these adaptations influence soil C transformations and microbial community assembly is complex and poorly understood.

To assess microbial adaptations and mechanisms of rhizosphere and bulk soil community assembly, we are using a common oat grass, *Avena spp.,* an abundant Mediterranean grassland species as a model system. We employ **(i)** exometabolomics to analyze microbial uptake of simple C molecules enriched in the rhizosphere; **(ii)** secretome analysis to identify mechanisms by which bacteria access root polymer carbon characteristic of the detritusphere; **(iii)** metatranscriptomic gene expression in the rhizosphere and bulk soil surrounding individual roots to determine how simple C compounds in root exudates influence complex soil organic matter degradation potential and community assembly.

(i) To investigate the ability of rhizosphere and bulk soil bacteria to metabolize simple carbon sources, we designed a medium to simulate previously identified plant exudates. We used LC-MS/MS based metabolomics to measure microbial consumption and release of metabolites and examine the substrate preferences of a diversity of rhizosphere and bulk soil bacterial isolates. We found that amino acids, sugars and nucleosides were consumed by all analyzed isolates. However, isolates that were preferentially stimulated by plant growth had substrate utilization preferences for aromatic organic acids, while bulk soil bacteria did not utilize these compounds. We used untargeted metabolomics to identify secondary metabolites released by rhizosphere

microorganisms that may be used to interact with the plant, influence its metabolism, and compete and cooperate with other organisms in rhizosphere.

(ii) Next we investigated the ability of soil bacteria to use complex polymeric carbon sources representative of the detritusphere. Carbon sources including cellulose, xylan, and a complex polymers contained within ground roots were provided as substrates and proteomics was used to identify extracellular proteins produced by rhizosphere and bulk soil isolates. We analyzed the activities of polymer degrading enzymes (β-glucosidase and β-xylosidase) with Matrix Assisted Laser Desorption/Ionization (MALDI) and labeled substrates. We determined that bulk soil adapted bacteria secrete a range of glycoside hydrolases (GH) and associated carbohydrate binding modules (CBMs) in the presence of plant polymers. We identified that bulk soil bacteria appear to have a preference for complex plant polymers, while rhizosphere bacteria appear metabolically adapted to use simple carbon sources that are commonly enriched in the rhizosphere. We propose that this substrate partitioning between rhizosphere and bulk soil organisms and release of secondary metabolites by bacteria are key mechanisms underlying rhizosphere community assembly and successional trajectories in observed soil.

(iii) To evaluate the hypothesized mechanisms of community assembly in the rhizosphere, we analyzed metatranscriptome gene expression of soil surrounding individual roots in microcosms experiments over 3 weeks of rhizosphere development. Half of the microcosms were also amended with root detritus to determine how simple and complex plant derived carbon influences community assembly and associated carbohydrate degradation potential. Population transcripts were binned using a unique reference database generated from soil metagenome-assembled genomes (MAGs), rhizosphere stable isotope probing (SIP)-MAGs, single amplified genomes (SAGs), and a panel of genomes from soil isolates. Both taxonomic and functional community composition rapidly diverged by 3 days, but functional composition underwent at major shift at 22 days that was undetectable by taxonomic community analyses. Bacterial niche differentiation during rhizosphere establishment was spatially and temporally regulated, where guilds of decomposers emerged that had distinct habitat preferences and timescales for decomposition. In particular, we identified three populations belonging to the Verrucomicrobiota, Fibrobacterota, and Burkholderiales that were synergistically upregulated in the rhizosphere amended with bulk litter. Roots stimulated the upregulation of genes coding for carbohydrate depolymerization, with the highest number of genes expressed in rhizospheres amended with root detritus. Together this work leads to a clear conceptual model where substrate preference and strategies for substrate utilization drive niche differentiation of bacteria in soil and provides a path to predictive models of rhizosphere microbial succession and its consequences for the fate of soil carbon.

References

1. Zhalnina, K., Louie, K., Mansoori, N., Hao, Z., Cho, H., Karaoz, U., Bowen, B., Firestone, M., Northen, T., Brodie, E. (2018) Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly, *Nature Microbiology*, 3: 470

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Numbers DE-SC0010570 and DE-SC0016247 to UC Berkeley. Part of this work was performed at Lawrence Berkeley National Lab under contract DE-AC02-05CH11231. Work at LLNL was performed under the auspices of OBER award SCW1589 and the U.S. Department of Energy under Contract DE-AC52-07NA27344.

Title: Resistance to Stalk Pathogens for Bioenergy Sorghum

Deanna Funnell-Harris* (Deanna.Funnell-Harris@ars.usda.gov), Maya Khasin and Scott Sattler

United States Department of Agriculture, Agricultural Research Service, Wheat, Sorghum and Forage Research Unit, Lincoln, NE; and University of Nebraska, Lincoln

Project Goals: This research is focused on discovering molecular and metabolic networks that drive sorghum resistance or tolerance to stalk rot pathogens. We are using advanced molecular techniques with the goal of identifying key factors associated with resistance and tolerance to stalk pathogens in sorghum lines optimized for key bioenergy traits (modified phenylpropanoid metabolism), with enhanced drought tolerance (nonsenescence) or with increased tolerance to stalk pathogens. Some lignin altered lines and drought tolerant lines already have demonstrated increased tolerance to these pathogens but specific mechanisms that mitigate pathogenic growth have yet to be identified. Our long-term goal is development of sorghum lines that withstand increased pathogen loads under reduced water conditions based on knowledge gained through this research.

Sorghum is a promising bioenergy crop with high yield potentials and significant tolerance to both drought and heat. However, under water or heat stress, sorghum is prone to stalk rots, which can significantly limit sorghum biomass yield through growth reductions and lodging. Stalk rot-causing fungi normally grow endophytically (asymptomatically) within sorghum plants. When sorghum plants experience water stress, host changes often trigger a developmental switch causing the fungi to become pathogenic, resulting in decayed stalk tissue. The underlying plant molecular circuits that either limit or exacerbate this fungal transition from endophytic to pathogenic growth are not known and are the focus of this proposal. Several publicly available lines have previously demonstrated resistance or tolerance to sorghum stalk pathogens, including lines with post-flowering drought tolerance (nonsenscence), which appears to suppress pathogenic growth, or a variety of lines that have exhibited increased resistance under field conditions. We have developed several near-isogenic *brown midrib* (*bmr*) *6* and *12* lines with altered lignin content and composition, which were previously demonstrated to have increased resistance or tolerance to stalk pathogens (2,3,5). Lignin, a component of plant cell walls, has been a focus for development of bioenergy sorghums because its presence increases recalcitrance of biomass to cellulosic ethanol conversion, but its presence also increases total energy content of biomass, which is important for thermal conversion technologies. The *bmr* lines have reduced lignin and increased ethanol conversion efficiency (1) due to single mutations in enzymes in the monolignol (lignin subunits) biosynthesis pathway, cinnamyl alchohol dehydrogenase (*bmr6*) or caffeic acid *O*-methyltransferase (*bmr12*). To increase energy content, we have engineered sorghum plants overexpressing a Myb transcription factor that induces synthesis of monolignols and a gene encoding caffeoyl-CoA *O*-methyltransferase, a monolignol pathway enzyme. Both transgenic and *bmr* plants accumulate phenolic intermediates from monolignol biosynthesis that inhibit stalk pathogens *in vitro* (5). We have developed an assay in a controlled environment, with applied water-stress, which reliably induces the developmental switch from endophytic to pathogenic growth of sorghum stalk rot fungi (2).

Our research may have identified sources of resistance in *bmr6* and *bmr12* lines, relative to wild-type, to the stalk rot pathogens, *Fusarium thapsinum* and *Macrophomina phaseolina.* We

have previously shown that following inoculation of peduncles with each of these fungi, a visible lesion is first apparent at 3 days post inoculation (dpi) and lesion expansion is first apparent at 13 dpi (4). In the current research, mean lesion lengths at 3 dpi were not significantly different between near-isogenic wild-type, *bmr6* and *bmr12* lines under either well-watered (100% field capacity) or reduced-water (25% field capacity) conditions. However, at 13 dpi, *bmr12* had significantly reduced lesion lengths, but only under reduced water, as compared with the wild-type; reductions in mean lesion lengths resulting on *bmr6* plants under this condition were not significant. Global gene expression of *bmr6, bmr12* and wild-type plants at 3 dpi under both watering conditions was conducted to identify early abiotic and biotic stress response genes.

The relative expression profile of infected tissues from the three plant genotypes under both watering conditions suggested common and unique host genetic responses influenced by genotype and watering condition. Gene expression analysis suggested that the reduced water condition primed *bmr12* for resistance to infection. Inoculated well-watered *bmr12* plants exhibited similar expression profiles to each fungus and to control reduced-water *bmr12,* but not control well-watered *bmr12*. Expression profiles of reduced water-treated wild-type and *bmr12* plants, and well-watered *bmr12* were mostly similar, suggesting that *bmr12* is stressed under well-watered conditions. Expression modules from *bmr12* well-watered control, but not wild-type, overlaps with reduced-water modules also suggesting priming for resistance to drought and infection. Nonetheless, there does not appear to be a unique expression pattern associated with the reduced lesion length observed in pathogen-infected *bmr12* plants under reduced water. Expression modules associated with reduced water included genes and pathways involved in photosynthesis, protein processing, carbon metabolism, and ubiquitin-mediated proteolysis. Expression modules associated with fungus infection under reduced water included amino acid biosynthesis; phenylpropanoid, glutathione and flavonoid biosynthesis; and, pyrimidine metabolism, DNA replication, homologous recombination and mismatch repair.

These results strongly indicated that *bmr6* and *bmr12* lines, and near-isogenic wild-type, are promising sources to identify genomic and metabolic markers that can be used to develop lines with increased resistance to stalk pathogens under water deficit conditions.

References:

1. Dien, B. S., Sarath, G., Pedersen, J. F., Sattler, S. E., Chen, H., Funnell-Harris, D. L., et al. 2009. Improved sugar conversion and ethanol yield for forage sorghum (*Sorghum bicolor* L. Moench) lines with reduced lignin content. *Bioenergy Research, 2,* 153-164, doi:10.1007/s12155-009-9041-2.
2. Funnell-Harris, D. L., O'Neill, P. M. and Sattler, S. E. 2018. Field response of near-isogenic brown midrib sorghum lines to fusarium stalk rot, and response of wildtype lines to controlled water deficit. *Plant Pathology, 67,* 1474-1482m doi: 10.1111/ppa.12863
3. Funnell-Harris, D. L., O'Neill, P. M., Sattler, S. E., Gries, T., Berhow, M. A. and Pedersen, J. F. (2017). Response of sorghum stalk pathogens to *brown midrib* plants and soluble phenolic extracts from near isogenic lines. *European Journal of Plant Pathology, 148,* 941–953, doi:10.1007/s10658-017-1148-2
4. Funnell-Harris, D. L., O'Neill, P. M., Sattler, S. E., & Yerka, M. K. (2016). Response of sweet sorghum lines to stalk pathogens *Fusarium thapinum* and *Macrophomina phaseolina*. *Plant Disease, 100,* 896-903, doi:10.1094/PDIS-09-15-1050-RE.
5. Funnell-Harris, D. L., Sattler, S. E., & Pedersen, J. F. (2014). Response of *Fusarium thapsinum* to sorghum *brown midrib* lines and to phenolic metabolites. *Plant Disease, 98*, 1300-1308, doi:10.1094/PDIS-09-13-0980.

This material is based upon work that is supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, under award number 2016-67009-25420.

Saprophytic Bacterium *Cellvibrio japonicus* Has Selective CAZyme Requirements During Physiologically Relevant Hemicellulose Degradation

Andrew D. Blake, Nina R. Beri, Hadassa S. Guttman, Raymond Cheng, and **Jeffrey G. Gardner*** (jgardner@umbc.edu)

University of Maryland - Baltimore County, Baltimore, MD

Project Goals: Completion of the project will identify and characterize the physiologically relevant carbohydrate active enzymes required to consume the polysaccharides found in lignocellulose by the saprophytic soil bacterium *Cellvibrio japonicus*. Additionally, over the course of the project the utility of these enzymes, including assessment of novel functions, will be evaluated for biotechnology applications.

Despite substantial published data on the structural and enzymatic parameters of recalcitrant polysaccharide deconstruction, there is much less work on understanding how carbohydrate active enzymes (CAZymes) exert their effects *in vivo*. We have used the soil bacterium *Cellvibrio japonicus* to generate a systems-level understanding of lignocellulose utilization, and here we report our progress on determining the essential factors for hemicellulose degradation. Specifically, via RNAseq seven CAZyme genes were significantly up-regulated during growth on xylan. A comprehensive gene deletion strategy determined that only one of these genes was essential for xylan utilization. Expression of this gene in *E. coli* generated an engineered strain that was proficient utilizing xylo-oligosaccharides. Other significant findings were that *C. japonicus* uses both secreted and membrane-bound CAZymes for hemicellulose degradation degradation. This strategy deviates from how the bacterium degrades cellulose. Finally, we found that a single critical CAZyme is necessary for soluble hemicellulose oligosaccharide utilization.

Publications

1. Blake AD, Beri NR, Guttman HS, Cheng R, and Gardner JG**.** 2018. The complex physiology of *Cellvibrio japonicus* xylan degradation relies on a single cytoplasmic β-xylosidase for xylo-oligosaccharide utilization. *Molecular Microbiology*. 107:610-22.

2. Nelson CE, Attia MA, Rogowski A, Morland C, Brumer H, and Gardner JG**.** 2017. Comprehensive functional characterization of the Glycoside Hydrolase Family 3 enzymes from *Cellvibrio japonicus* reveals unique metabolic roles in biomass saccharification. *Environmental Microbiology*. 19:5025-39.

3. Nelson CE, Rogowski A, Morland C, Wilhide JA, Gilbert HJ, and Gardner JG**.** 2017. Systems analysis in *Cellvibrio japonicus* resolves predicted redundancy of β-glucosidases and determines essential physiological functions. *Molecular Microbiology*. 104:294-05.

4. Nelson CE, Beri NR, and Gardner JG**.** 2016. Custom fabrication of biomass containment devices using 3-D printing enables bacterial growth analyses with complex insoluble substrates. *Journal of Microbiological Methods*. 130:136-43.

This work is supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0014183.

Automated Circuit Design in Saccharomyces cerevisiae

Ye Chen[1]*(yechen@mit.edu), Shuyi Zhang[1], Eric M. Young[1], and **Christopher A. Voigt**[1]

[1] Massachusetts Institute of Technology, Cambridge, MA

Project Goals: Our objective is to develop a new standard for the engineering of microbial systems based on rational design, engineering, and optimization of hybrid regulatory networks. We envision a future biorefinery that is based on the development of designer organisms that have exquisite and predictable control architectures governing the expression of a range of valuable traits. Computer aided design platforms will guide the assembly of synthetic constructs containing orthogonal heterologous circuits to recode native regulatory networks. Together, these will enable predictable and dynamic control of multiple designer phenotypes such as: i) growth on various feedstocks in consolidated bioprocesses, ii) feedback control to mitigate accumulation of toxic metabolites, iii) production of target molecules (C3-C4 alcohols), and/or iv) robustness to process upsets (e.g. temp., phage). The focus of this proposal is to develop the technical and computational infrastructure to enable this vision. We will develop this platform first in the model organism *E. coli* and then in DOE relevant non-model organisms.

Genetic circuits are widely used in scientific, industrial and therapeutic approaches. In prokaryotic systems, genetic circuits are successfully designed and built by Cello, the genetic circuit design automation platform. However, building genetic circuits in eukaryotic systems are challenging due to the limited sensors and regulatory units. Here, we developed a method to systematically import bacterial transcription factors into Saccharomyces cerevisiae to create a bunch of candidate sensors and NOT gates. We next determined strategies to link these elements together with minimum cross interactions. Finally, we incorporated these novel components and construction strategies with Cello to design DNA sequence. By using this platform, we successfully created several large circuit constructions (6-8 gates, 9–11 regulators, up to 20 regulatory operons) in yeast. We further built an ODE model to investigate the dynamics when input states changed. Our model predicted the intermediate faults during several transferences, which were confirmed in the following experiments. In this study, we demonstrate a general approach of circuit design automation in a novel eukaryotic system. This will highly expand our ability to build complicated circuits across different organisms.

This research is supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Design and Engineering of Synthetic Control Architectures for Complex Phenotypes in *E. coli* and Yeast

Emily Freed[1*] (emily.freed@colorado.edu), Rongming Liu[1], Liya Liang[1], Adam Arkin[2], Christopher Voigt[3], Carrie Eckert[1,4], and **Ryan Gill**[1,5]

[1]Renewable and Sustainable Energy Institute, University of Colorado, Boulder CO; [2]Lawrence Berkeley National Laboratory, Berkeley, CA; [3]Massachusetts Institute of Technology, Cambridge, MA; [4]National Renewable Energy Laboratory, Golden, CO; [5]Danish Technical Institute, Copenhagen, Denmark.

http://www.gillgroup.org/research/

Project Goals: Our objective is to develop a new standard for the engineering of microbial systems based on rational design, engineering, and optimization of hybrid regulatory networks. We envision a future biorefinery that is based on the development of designer organisms that have exquisite and predictable control architectures governing the expression of a range of valuable traits. Computer aided design platforms will guide the assembly of synthetic constructs containing orthogonal heterologous circuits to recode native regulatory networks. Together, these will enable predictable and dynamic control of multiple designer phenotypes such as: i) growth on various feedstocks in consolidated bioprocesses, ii) feedback control to mitigate accumulation of toxic metabolites, iii) production of target molecules (C3-C4 alcohols), and/or iv) robustness to process upsets (e.g. temp., phage). The focus of this proposal is to develop the technical and computational infrastructure to enable this vision. We will develop this platform first in the model organism *E. coli* and then in DOE relevant non-model organisms.

The sustainable production of biofuels is of continued importance in light of increasing concerns about climate change and energy security. Advances in metabolic engineering, synthetic biology, and systems biology have provided a number of strategies for the more rapid design, construction, and testing of model systems for the production of next generation fuel-grade compounds (e.g. C3–C5 alcohols). However, the titer and productivity of engineered strains that produce C3-C5 alcohols are still below those required for economic production. The rate limiting step is no longer our ability to construct designer strains, but rather how to design and engineer increasingly complex networks of combinatorial phenotypes required for the economic and sustainable production of these biofuels and other bioproducts.

To address these issues, we developed a foundation for forward engineering of regulatory control architectures, which combines CRISPR Enabled Trackable Genome Engineering (CREATE; developed in the Gill lab) and forward engineering of *E. coli*/yeast regulatory networks to access complex targeted phenotypes. We designed, constructed, and mapped libraries including ~300,000 mutations from 200 regulatory genes in *E. coli* and yeast for C2-C8 alcohols tolerance and production. In all cases, we identified mutants that significantly improve the cell growth, titer, and productivity.

This research is supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Design and engineering of native regulatory networks in non-model microbes

Emily Freed[1], Evan Johnson[1], Margaret Habib[1*](maha0174@colorado.edu), Jeff Cameron[1], Ian Wheeldon[2], Min Zhang[3], Christopher Voigt[4], Adam Arkin[5], Carrie Eckert[1,3], and **Ryan Gill**[1,6]

[1]Renewable and Sustainable Energy Institute, University of Colorado, Boulder CO; [2]University of California, Riverside, CA; [3]National Renewable Energy Laboratory, Golden, CO; [4]Massachussetts Institute of Technology, Boston, MA; [5]Lawrence Berkeley National Laboratory, Berkeley, CA; [6]Danish Technical Institute, Copenhagen, Denmark.

Project Goals: Our objective is to develop a new standard for the engineering of microbial systems based on rational design, engineering, and optimization of hybrid regulatory networks. We envision a future biorefinery that is based on the development of designer organisms that have exquisite and predictable control architectures governing the expression of a range of valuable traits. Computer aided design platforms will guide the assembly of synthetic constructs containing orthogonal heterologous circuits to recode native regulatory networks. Together, these will enable predictable and dynamic control of multiple designer phenotypes such as: i) growth on various feedstocks in consolidated bioprocesses, ii) feedback control to mitigate accumulation of toxic metabolites, iii) production of target molecules (C3-C4 alcohols), and/or iv) robustness to process upsets (e.g. temp., phage). The focus of this proposal is to develop the technical and computational infrastructure to enable this vision. We will develop this platform first in the model organism *E. coli* and then in DOE relevant non-model organisms.

The realization of a sustainable bioeconomy requires our ability to understand and engineer complex design principles for the development of platform organisms capable of efficient conversion of cheap and sustainable feedstocks (e.g. sunlight, CO_2, non-food biomass) to biofuels and bioproducts at sufficient titers and costs. Despite recent advances in DNA synthesis allowing for the construction of small, synthetic genomes and the development of high-throughput genome editing and metabolic engineering tools in model microbes, our ability to design genomes and predict design principles for intricate functions such as photosynthesis and tolerance are still limiting. In the proposed work, we will leverage our knowledge and expertise in these cutting-edge synthetic biology techniques currently only available for model microbes by partnering with expert collaborators in the adaptation of these methods for DOE-relevant microbial systems, accelerating and expanding genome editing capabilities for metabolic engineering in these organisms. Through doing so, we will be able to more quickly uncover genotype to phenotype relationships to better engineer these microbes for optimal production of chemicals and fuels from renewable feedstocks.

Synechococcus sp. PCC7002 is attractive as a photosynthetic chassis for sustainable biochemical production because of its fast growth rate and its ability to grow using only sunlight as an energy source to convert CO_2 into products. Transcriptional responses to different growth conditions have been measured, and the key players in RNA-turnover, a major process controlling cellular production, have been characterized by collaborator Jeff Cameron. In addition, a mass transport and kinetic modeling of the PCC7002 CO_2 concentrating mechanism identified potential ways to increase photosynthetic growth in industrial conditions, and metabolic flux models have been developed for PCC7002 to identify changes predicted to increase biofuel production. PCC7002 is naturally competent for DNA uptake and targeted gene deletions can be routinely performed at relatively high efficiencies using native homologous recombination machinery. A robust synthetic biology toolkit is available including promoter and ribosome binding site libraries, orthogonal/inducible promoter systems, counter-selectable markers, multiple neutral integration

sites in the genome, and an inducible CRISPR-interference system to enable titratable repression of endogenous genes.

Zymomonas mobilis ZM4 is a facultative anaerobe that is of interest for the fermentation of biomass sugars to biofuels and bioproducts, especially those branching from pyruvate. Collaborator Min Zhang demonstrated that carbon flux can be deviated from ethanol production at the pyruvate node into 2,3-butanediol as well as to farnescene (unpublished). They have generated a large volume of –omics data, providing insights into hydrolysate tolerance and feedstock utilization. Transformation efficiencies with a variety of vectors have been optimized, homologous recombination is performed routinely for genome engineering, and a number of functional promoters, both constitutive and inducible, have been characterized. In addition, heterologous expression of Cas9 and targeting gRNA was demonstrated to cure native plasmids in *Z. mobilis* (Cao et. al., 2017), representing initial proof of concept for CRISPR-based gene editing and regulation.

The yeast *Kluveromyces marxianus* is an interesting biomass sugar utilizing organism because of its tolerance to low pH, higher temperatures, and high flux to pyruvate. Numerous tools have been developed to enable directed engineering of *K. marxianus* by collaborator Ian Wheeldon and others. Centromeric episomal plasmid systems have been well established for this organism, and CRISPR-Cas9 has been developed for genome editing and gene knockdown (via CRISPR interference) technologies have been developed to enable rapid genome-wide screening efforts.

We are currently working to adapt genome engineering tools for each of the above target microbes to enable technologies such as CRISPR interference (CRISPRi) for gene knockdown and CRISPR-Enabled Trackable Geneome Engineering (CREATE) for multiplexed editing. The investigators on this project have all developed an array of vectors for Cas9/dCas9 and gRNA expression in a range of organisms, and a large number of additional vectors are available through Addgene (http://www.addgene.org/). These vectors will be adapted, as necessary, for stable replication and predictable expression in each of the testbed organisms. Targeting, cutting, and recombination efficiencies will be evaluated and optimized for Cas9/gRNA utilizing an appropriate screen/selection method (antibiotic resistance, auxotrophy, counterselective, colorimetric, etc.). In addition, we are evaluating and optimizing knockdown efficiencies using dCas9 in a similar fashion.

Once CRISPR-based genome editing is validated in each of the non-model systems, we will adapt the CREATE method for targeting global regulators in the testbed strains to mirror the regulator libraries utilized by other teams on this project (Ryan Gill and Chris Voigt) to expand the search space for regulatory control switches. Utilizing existing pathway maps for each organism, we will build pathway prediction models for each to identify global regulators to target. Information gleaned from the selection experiments (improved growth, tolerance, etc.) will guide further improvements to the models for development of synthetic regulatory networks by another team member (Adam Arkin). Together, this work will provide a blueprint for the development of systems to accelerate the engineering of non-model microbial systems as bioproduction chassis organisms.

References

1. Qing-Hua Cao, Huan-Huan Shao, Hui Qiu, Tao Li, Yi-Zheng Zhang & Xue-Mei Tan (2017) Using the CRISPR/Cas9 system to eliminate native plasmids of Zymomonas mobilis ZM4, Bioscience, Biotechnology, and Biochemistry, 81:3, 453-459

This research is supported by the Office of Biological and Environmental Research in the DOE office of Science, grant DE-SC0018368.

CRISPR-based trackable protein engineering

Eun Joong Oh[1,2*] (eun.oh@colorado.edu), Rongming Liu[1,2], Liya Liang[1,2], Emily F. Freed[1,2], Carrie A. Eckert[2,3] and **Ryan T. Gill**[1,2]

[1]University of Colorado Boulder, Boulder; [2]Renewable and Sustainable Energy Institute, Boulder, Colorado; and [3]Bioscience Center, National Renewable Energy Laboratory, Golden, Colorado.

Project Goals: Our objective is to develop a new standard for the engineering of microbial systems based on rational design, engineering, and optimization of hybrid regulatory networks. We envision a future biorefinery that is based on the development of designer organisms that have exquisite and predictable control architectures governing the expression of a range of valuable traits. Computer aided design platforms will guide the assembly of synthetic constructs containing orthogonal heterologous circuits to recode native regulatory networks. Together, these will enable predictable and dynamic control of multiple designer phenotypes such as: i) growth on various feedstocks in consolidated bioprocesses, ii) feedback control to mitigate accumulation of toxic metabolites, iii) production of target molecules (C3-C4 alcohols), and/or iv) robustness to process upsets (e.g. temp., phage). The focus of this proposal is to develop the technical and computational infrastructure to enable this vision. We will develop this platform first in the model organism *E. coli* and then in DOE relevant non-model organisms.

Microbes have been harnessed to produce value-added products as cell factories. Many research groups are investigating strategies to use microbial cell factories for sustainable and economical protein production such as enzymes and antibodies. Because the microbes have extensive regulation between metabolic enzymes, protein engineering by error-prone PCR and rational design sometimes leads to disruptive cellular metabolism and regulatory mechanism. In addition, our knowledge is limited to uncover the underlying system controlling metabolic homeostasis[1]. Understanding native regulatory networks will require making libraries of regulatory proteins.

CRISPR EnAbled Trackable genome Engineering (CREATE) is a strategy that combines CRISPR/CAS9 editing with multiplexed oligo synthesis, enabling mapping of mutations to traits of interest[2]. In our proof-of-concept study we built editing libraries in *E. coli* following our prior CREATE protocols and then used such editing libraries to build plasmid-based protein libraries in *S. cerevisiae*. Currently, we are attempting to apply CREATE to antibody engineering. CREATE-based single chain antibody mutants provided high-affinity antibodies with trackable high-throughput mapping of desired phenotypes. We could evaluate the contribution of each mutation to the improved affinity.

Our long-term goal is to develop the optimized system for protein engineering including enzymes. CREATE-based enzyme evolution might provide a better understanding of protein fitness landscapes than random mutagenesis in optimizing protein function. We envision that the

CREATE technology will enhance understanding of complex biological networks in protein engineering and further improve biofuel production in microbes.

References

1. Nielsen, J. & Keasling, J. D. Engineering Cellular Metabolism. *Cell* 164, 1185–1197 (2016).

2. Garst, A. D. et al. Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nat. Biotechnol*. 35, 48–55 (2017).

This research is supported by the Office of Biological and Environmental Research in the DOE Office of Science.

m-CAFEs: The use of EcoFABs to study beneficial plant-fungus interactions

Jens Heller[1], Peter Kim[2], Trenton K. Owens[1], Vincent Kam[1], Peter F. Andeer[1], Lauren K. Jabusch[1], Nameera Baig[1], Dawn M. Chiniquy[1], Spencer Diamond[3], Benjamin M. Rubin[3], Claudio Hidalgo[4], Matthew Nethery[4], Joelle Schlapfer[1], Kateryna Zhalnina[1], Mary K. Firestone[3], Jennifer A. Doudna[3], Jill Banfield[3], Rodolphe Barrangou[4], Romy Chakraborty[1], Adam M. Deutschbauer[1], Anup K. Singh[2], Trent R. Northen[1], and **N. Louise Glass[1]**

[1]Lawrence Berkeley National Laboratory, Berkeley; [2]Sandia National Laboratory, Livermore; [3]University of California, Berkeley; [4]North Carolina State University, Raleigh

http://eco-fab.org

Project Goals: To derive mechanistic understanding of fungal contributions to mineral accessibility in marginal soils

The m-CAFEs program is a collaborative, integrated, and mission-driven program to interrogate the function of soil microbiomes with critical implications for carbon cycling and sequestration, nutrient availability and plant productivity in natural and managed ecosystems. We use precisely controlled ecosystem fabrications (EcoFABs) for creating and manipulating plant-microbe interactions under specific environmental conditions. The EcoFABs design consists of a sterile plant growth chambers attached to a microscope slide. This set-up considerably enhances studying the impact of microbes on plant growth phenotypes.

In plants, phosphorus drives biological reactions and is essential for growth. Phosphorus facilitates root formation, reproductive development and synthesis of proteins. Phosphorus is one of the most limited nutrients for plants in the environment because the form that is preferentially asimulated, inorganic phosphate, is unevenly distributed in soils and >80% is not readily available to roots. In the environment, phosphorus is primarily present as insoluble iron and aluminum phosphates in acidic soils or calcium phosphates in alkaline soils.

Numerous microorganisms - especially those associated with roots - have the ability to increase plant growth and productivity. Filamentous fungi of the genus *Trichoderma* are among those microbes and the most commonly studied natural bio-control agents. Some *Trichoderma* species have been shown to protect plants against biotic and abiotic stresses and to promote plant growth by increasing nutrient uptake. Their potential in solubilization of otherwise unavailable mineral nutrients is under investigation. Using EcoFABs, we are able to spatially separate phosphorus sources and plant roots. We demonstrate that *Trichoderma harzianum* promotes plant growth in the presence of insoluble $AlPO_4$. Importantly, this effect was only evident when *T. harzianum* was inoculated close to the $AlPO_4$ source, whereas no growth promotion was observed when the fungus was inoculated close to the seedling. These data suggest that *T. harzianum* helps to solubilize phosphate and makes it available for plant uptake and utilization.

Another efficient strategy of plants to overcome phosphate limitation is the association of roots with arbuscular mycorrhizal (AM) fungi. Even though these microorganisms form symbiotic associations with nearly 80% of terrestrial plants, research on AM fungi is limited due to the challenging nature of these systems. New methods that facilitate the handling of AM fungi are essential to enhance and enable the research on these beneficial organisms. Here we present a new EcoFAB design to study the mycorrhizal plant-fungal interactions. This EcoFAB design consists of a plant growth chamber and compartments for nutrient sources that are separated by a mesh sheet; this set-up effectively prevents roots from accessing nutrients if not associated with AM fungi. Using such a design will enable us to measure and image direct impact of AM fungi on nutrient access, plant growth and rhizosphere microbiome associations.

Funding statement.

This material by m-CAFEs Microbial Community Analysis & Functional Evaluation in Soils, (m-CAFEs@lbl.gov) a Project led by Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

m-CAFEs: EcoFABs: Investigating Soil Microbe-Plant Interactions in Regulated Ecosystems

Authors: **Lauren K. Jabusch**[1*] , Joelle Schlapfer Sasse[1], Peter F. Andeer[1], Peter Kim[2], Jens Heller[1], Nameera Baig[1], Dawn M. Chiniquy[1], Trenton K. Owens[1], Andrew Klein[5], Borjana Arsova[6], Josefine Kant[6], Benjamin Cole[1], Michelle Watt[6], John Vogel[1], Rodolphe Barrangou[4], Romy Chakraborty, Jeff Dangl[5], Jennifer A. Doudna[3], Jill Banfield[3], Mary K. Firestone[3], Peter S. Nico[1], Anup K. Singh[2], Adam M. Deutschbauer[1], Trent Northen[1]* (TRNorthen@LBL.gov) and, N. Louise Glass[1]

[1]Lawrence Berkeley National Laboratory, Berkeley CA; [2]Sandia National Laboratory, Livermore CA; [3]University of California, Berkeley CA; [4]North Carolina State University, Raleigh NC; [5]University of North Carolina at Chapel Hill, NC, [6]Institut für Bio-& Geowissenschaften IBG-2, Forschungszentrum Jülich, Germany

http://eco-fab.org

Project goals: **To derive mechanistic understandings of plant-microbe-soil interactions using reproducible, simplified ecosystems.** The m-CAFEs program is a collaborative, coordinated and integrated, mission-driven program to interrogate the function of soil microbiomes with critical implications for nutrient availability and plant productivity in natural and managed ecosystems.

Ecological Fabrications (EcoFABs) are microfluidic devices of molded polydimethylsiloxane (PDMS) bonded or clamped to glass slides to create highly controlled plant-growth platform that address the research needs for microbial-plant studies. These systems are adaptable to variable levels of soil complexity, including hydroponic conditions or filled with soil or other solid substrates such as synthetic soil, perlite, or sand. EcoFABs can be designed to contain separations or pillars, creating niches for slower growing microbes; barriers for hyphae-mediated nutrient transport; or flat roots for improved microscopy. Imagining of plant roots is easily achieved through 0-4 weeks of plant growth in EcoFABs bonded to glass cover slips for magnification between 5-40x. We have found that EcoFABs support the growth of DOE-relevant grasses, including switchgrass (*Panicum virgatum*), *Panicum hallii*, sorghum, maize, and *Brachypodium distachyon.* Importantly, a recent intra-lab ring-trial has shown a high degree of reproducibility between labs for diverse plant traits. Root imaging in an ecoFAB device was recently improved through the incorporation of novel flow paths. The device regulates the root development path, allowing the root to stay in focus for microscopy. This modified device has been used to localize the distributions of rhizopsphere bacteria. The results of experiments using these devices will transform our understanding of soil metabolism and microbiome science, contributing to DOE missions in energy and environment.

Funding statement.
This material by m-CAFEs Microbial Community Analysis & Functional Evaluation in Soils, (m-CAFEs@lbl.gov) a Project led by Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

m-CAFEs Applications of Targeted Killing and Editing Methods in Microbial Networks

Matthew A. Nethery[3*](manether@ncsu.edu), Claudio Hidalgo[3], Benjamin M. Rubin[2], Spencer Diamond[2], Trenton K. Owens[1], Jens Heller[1], Dawn M. Chiniquy[1], Jill Banfield[2], Jennifer A. Doudna[2], Trent R. Northen[1], Adam M. Deutschbauer[1], **Rodolphe Barrangou[3]** and N. Louise Glass[1]

[1]Lawrence Berkeley National Laboratory, Berkeley; [2]University of California, Berkeley; [3]North Carolina State University, Raleigh

https://genomicscience.energy.gov/research/sfas/lbnlmcafes.shtml

Project Goals: To elucidate the complex dynamics of microbial networks using phage and CRISPR-based tools for targeted killing and editing of individual microbes within a mixed community.

Abstract

The m-CAFEs collaborative strives to understand the function and biological relevance of individual microbes within the complex ecosystem of the soil microbiome, facilitating reproducibility through use of fabricated ecosystems (EcoFABs). Current technologies for manipulation of an individual species within a consortium are extremely limited and face two primary challenges: construct delivery and target specificity. Therefore, a major goal of the m-CAFEs program is to harness and augment current phage-based targeted killing methods, as well as to develop and enhance delivery of CRISPR-Cas tools for targeted killing and editing in order to elucidate the functions and roles of both cultivatable and uncultivatable microbes within the soil microbiome. Here, we present preliminary data on phage delivery to a mixed population for selective ablation, as well as chemical-based delivery of CRISPR-Cas constructs for targeted killing and editing. Enhancing these delivery and targeting methods will provide essential insights into the functional genomics and biological importance of individuals within the mixed microbial community.

This material by m-CAFEs Microbial Community Analysis & Functional Evaluation in Soils, (m-CAFEs@lbl.gov) a Project led by Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231

Molecular Composition of Field Derived Microbial Necromass

Kirsten Hofmockel[1]* (kirsten.hofmockel@pnnl.gov), Sheryl Bell, Allison Thompson, Rob Starke, Lindsey Anderson, Andy Lipton

[1]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA

Project Goals: The overall goal of this project is to test if plant-microbe interactions are limited to influencing the rate of C accrual, while mineralogy regulates the sink capacity of biofuel cropping systems. To accomplish this goal, we are (1) identifying the microbial functions and biopolymers of microbial necromass that contribute to soil C accumulation under controlled conditions, (2) characterizing microbial necromass accumulation in response to crop selection and edaphic factors in situ and (3) generating long-term, cross-site data that can be used to model C cycling in bioenergy cropping systems under different soil conditions.

Abstract text. Crop selection and soil texture influence the physicochemical attributes of the soil, which structures microbial communities and influences soil organic matter formation, cycling and long-term storage. At the molecular scale, microbial metabolites and necromass alter the soil environment, which creates feedbacks that influence ecosystem functions, including soil organic matter accumulation. Yet the generalizable mechanisms regulating the accrual and long term stabilization of soil organic matter are still unclear. By integrating lab to field studies we aim to identify the molecules, organisms and metabolic pathways that control the formation of molecules that contribute to long term organic matter stabilization in bioenergy soils.

Microbial residues accumulate in soil and thus are an important contributor to the formation of soil organic matter. Plant-derived inputs undergo microbial decomposition, some of the resulting organic residues are incorporated into microbial biomass, and a significant proportion of soil organic matter is attributed to the resulting microbially derived residues, or necromass. This includes biomass residues (lipids, proteins, amino sugars) and microbial exudates (enzymes, exopolysaccharides, lipids, glycoproteins). Yet little empirical evidence is available to support this conceptual model and inform management decisions that aim to amplify biological processes that enhance soil organic matter formation and persistence. To address this knowledge gap, we investigated the chemical signatures of microbial biomass separated into intracellular and extracellular aqueous, and water insoluble fraction from seven taxonomically different cultures. The cultures were enriched from switchgrass (*Panicum virgatum*) plots from the DOE the Great Lakes Bioenergy Research Center (GLBRC) Intensive Biofuel Cropping System Experiments in MI and WI, USA. Microbial communities were selectively cultured from sandy loams of the Kellogg Biological Station (KBS) in MI, and silty loams from the Arlington Agricultural Research Station (AARS) in WI. Microbial communities cultivated from the MI and WI field trials were used to test the hypothesis that microbial communities differ in morphology and biochemistry and therefore in their potential contributions microbial necromass.

Media was chosen to represent a diversity of microbial morphology and taxonomy by broadly enriching for the following groups: (1) Gram+ bacteria, (2) G+ *Bacillus*, (3) G+ *Firmicutes*, (4) *Rhizobium* spp., (5) slow growers, (6) saprophytic fungi including *Aspergillus* spp and *Penicillium* spp and (7) diverse mycelial fungi. Soil diversity and complexity was captured in our experiment with average richness of 34 bacterial, 1 archaeal and 47 fungal genera across the selective media. Using identical culturing conditions, different communities were cultivated from the sandy and silty loams, allowing us to capture the phylogenetic diversity of the field within constrained enrichment cultures.

Results from 16S and ITS revealed unique variation in taxon representation across the seven nutrient-dependent enrichment techniques for both soils. Some enrichments differed in dominant genera depending on the originating soil. For example the media selecting for *Rhizobium* spp. was dominated by *Acinetobacter* in the sandy loam and by *Planococcaceae* in the silty loam. The two fungal enrichment media highlighted the most unique chemical signatures, driven by differing fungal composition between the sites, though the bacterial communities were similar. Within these enrichment cultures 16S bacterial composition for both sandy and silty loams were dominated primarily by *Firmicutes* (specifically *Burkholdriales*). However, ITS results reveal distinct fungal variation despite all taxa belonging to the same phylogenetic division. The saprophytic fungi enrichment in the sandy loam was highly dominated by *Pleosporales*, while in the silt loam, the mycelial fungi enrichment was mainly representative of *Hypocreales*. Although the gram positive selective media also had distinctive insoluble chemical signatures, this difference was not driven by differing fungal communities, as this media showed nearly no fungal growth.

Community differences were reflected in biomarkers specific to the extracellular and intracellular aqueous fraction of microbial biomass, identified using GC-MS and liquid state 1D ^{1}H-NMR. Metabolic profiles for both intracellular and extracellular fractions reveal site specific differences among the selective media. Intracellular metabolic pathways identified consistently revealed elevated levels in the sandy loam compared to the silt loam. Highly represented intracellular pathways were in secondary metabolite biosynthesis (antibiotic biosynthesis, nitrogen-containing secondary compound biosynthesis, and siderophore biosynthesis), amino acid degradation (Arg, Phe, and Thr degradation were of higher levels than others), and fermentation (acetate and glycerol being main metabolites of contribution). Intracellular metabolism was representative of metabolites with high turnover value, including substrates essential for core metabolic pathways. The opposite was true for extracellular metabolic profiles in which primary activity was elevated in silty loams over sandy loam. Pathways represented were secondary metabolite biosynthesis, carbohydrate degradation (carbohydrate, carboxylate, and hormone degradation), and fermentation and CO_2 fixation (formate being a main metabolite of contribution). These results show how differences in community composition among sites can induce intracellular and extracellular shifts that influence the production of microbial derived residues. Even under the same culturing conditions, differences in community composition revealed greater levels of intracellular metabolites in the sandy loam communities and greater extracellular metabolites in silty loam communities. These metabolic profiles influence the

microbial residues available for SOM formation and accumulation, resulting in distinct necromass signatures as evidenced in cell wall chemistry.

Cell walls are presumed to be the most persistent portion of microbial necromass. Our analysis of cell wall residues via solid state ^{13}C-NMR revealed unique lipid and terminal methyl chemical profiles between microbes cultivated from the two soils. Specifically the two fungal and the gram positive enrichments from the sandy loam revealed increased profiles of lipids and terminal methyl groups. By contrast, the silty loam communities increased in a distinctly different profile of lipids and terminal methyl groups that was evident for communities grown on the bacterial selective media. Together these insoluble chemical signatures suggest that fungal and bacterial communities both contribute to necromass accumulation via lipid and fatty acid production and the relative contribution appears to vary with community membership and metabolism. Lipid signals were greatest in bacterial communities enriched from the silt loams and fungal terminal methyl and fatty acid groups were greatest in fungal cultures derived from the sand loam. The hydrophobicity and complexity of these molecules make them good candidates for persistent necromass. Our results suggest that biochemical differences in fungal and bacterial communities may be important for the quantity and quality of necromass production. If lab results translate to field conditions, we expect that differences in community membership between sites will strongly influence the microbial interactions that lead to the formation and persistence of SOM.

This research was supported by an Early Career Research Program award to KS Hofmockel, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research Genomic Science program under FWP 68292. A portion of this work was performed in the William R. Wiley Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility sponsored by OBER and located at PNNL. PNNL is a multi-program national laboratory operated by Battelle for the DOE under Contract DE-AC05-76RLO 1830.

Can we measure taxon-specific biochemical efficiency in natural microbial communities?

Bruce A. Hungate[1,2] (Bruce.Hungate@nau.edu), Paul Dijkstra[1,2], Egbert Schwartz[1,2], Ember M. Morrissey[3], Bram W. Stone[1], Michaela Hayer[1], Steven J. Blazewicz[4], Xavier Mayali[4], Peter Weber[4], Jennifer Pett-Ridge[4], Kirsten Hofmockel[5]

[1]Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ, 86011, USA; [2]Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, 86011, USA; [3]Department of Biology, West Virginia University, Morgantown, WV, 26506, USA; [4]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, 94550, USA; [5]Environmental and Molecular Sciences Division, Earth Scientist Pacific Northwest National Laboratory, Richland, WA 99352

Project Goals: Short statement of goals. The objectives of the proposed work are to 1) develop a quantitative SIP targeted –omics-based framework to infer taxon-specific rates of growth, mortality, and associated carbon (C) fluxes in soil microbial communities, a framework that will scale from individual microbial taxa to the integrated soil system; 2) develop a new means to quantify taxon-specific C-use and growth efficiency in order to quantify taxon-specific contributions to soil C loss; and 3) apply newly developed quantitative tools to test key hypotheses about responses of the soil microbial community to experimental warming and to latitudinal variation in temperature. This work will push the frontier of –omics enabled techniques by demonstrating their applicability at the ecosystem scale, and by relating taxon-specific isotope assimilation to dissimilatory processes in the C cycle, thereby enabling the identification of organisms especially responsible for soil C loss.

Carbon use efficiency is a key parameter in global biogeochemical models of the carbon cycle, and is usually described as the partitioning of carbon within a cell between allocation to biosynthesis and growth versus allocation to energy production and respiration. Yet, as commonly measured (e.g., by measuring the distribution of ^{13}C-tracers between biomass and respiration), carbon use efficiency includes biochemical efficiency, growth efficiency, and biomass turnover (predation, viruses, etc.), creating confusion in the literature about exactly what it is we are measuring (Geyer et al., 2016). Furthermore, carbon use efficiency is almost always considered in aggregate, as a property of the entire microbial community, ignoring potential differences among taxa, despite evidence from pure culture that components of carbon use efficiency vary strongly among taxa, variation sometimes invoked to explain apparent changes in carbon use efficiency observed in nature.

Here, we focus on the use of metabolic flux analysis as a means to measure biochemical efficiency, the instantaneous partitioning between biosynthesis and respiration within the central metabolic network common to the vast majority of life. We chose this focus for two reasons: 1) it provides clarity about the component of carbon use efficiency under investigation, and 2) features of its measurement create the possibility of resolving biochemical efficiency on a taxon-specific basis.

Given the tremendous diversity of microorganisms in evolutionary history and physiological traits, variation in biochemical efficiency among microbial taxa seems very likely. While the evidence is scarce, microbial community composition has frequently been posited as

a potential driver of carbon use efficiency. Biochemical efficiency captures processes that influence growth, survival, and thus fitness, so it should be subject to selective pressure. Ecological strategies may underlie differences in biochemical efficiency. Slow growing organisms, typically present in low nutrient environments, are expected to exhibit high biochemical efficiency, whereas fast growing organisms that are top competitors in high nutrient environments will have lower biochemical efficiency, potentially due to physiological tradeoffs between efficiency and maximum growth rate.

Still, even if it is important, how could we possibly measure it for individual taxa in complex assemblages? The idea is challenging. On the one hand, isotopic tags on the elements assimilated and recovered in nucleic acids provide a way to discern which organisms use what resources at what rates for biomass growth. On the other hand, biochemical efficiency is about the balance of growth and respiration, or dissimilation, where the product, carbon dioxide, cannot be traced back to the organism that produced it.

Metabolic flux analysis can, in principal, be used to develop taxon-specific estimates of biochemical efficiency: using position-specific ^{13}C-labeled substrates such as glucose and pyruvate, monitoring the isotope incorporation into nucleotides and DNA and RNA combines the measurement of metabolic efficiency while retaining information about the organisms that produced them. The atom mapping logic for this approach can be understood from the biochemistry of glucose anabolism and catabolism, in glycolysis, the Krebs cycle, the pentose phosphate pathway, and gluconeogenesis, pathways most important for energy production and biosynthesis, including the biosynthesis of nucleic acids.

Here, we present the model for determining biochemical efficiency of individual bacterial taxa in complex communities, and evaluate the sensitivity with respect to two proposed methods of measuring taxon-specific isotope composition: ChipSIP/NanoSIMS, and quantitative stable isotope probing (qSIP). In this work, the target molecules for the recovery of these isotope tracers are nucleic acids, the very molecules that convey the identities of the organisms. This is critical because it overcomes the problem of assigning responsible taxa to dissimilatory processes. Our analysis indicates that sensitivity with quantitative stable isotope probing is too coarse for this approach, but that NanoSIMS measurements of taxonomically resolved nucleic acid sequences provides sufficient resolution to detect ecologically meaningful changes in biochemical efficiency, suggesting that our proposed technique is a viable way to measure biochemical efficiency for individual taxa growing in a community, a dream of microbial ecologists young and old from around the world.

Reference: Geyer, K.M., Kyker-Snowman, E., Grandy, A.S., Frey, S.D., 2016. Microbial carbon use efficiency: accounting for population, community, and ecosystem-scale controls over the fate of metabolized organic matter. Biogeochemistry 127, 173-188.

Funding Statement This work was supported by the US Department of Energy, Program in Genomic Sciences, Award Number: DE-SC0016207.

Genomic traits predict microbial growth in culture but fail in soils, except during resource pulses

Junhui Li[1], Benjamin J. Koch[1,2*] (Benjamin.Koch@nau.edu), Rebecca L. Mau[1], Paul Dijkstra[1,2], Egbert Schwartz[1,2], Xiao-Jun Allen Liu[1,2], Ember M. Morrissey[3], Steven J. Blazewicz[4], Jennifer Pett-Ridge[4], Bram W. Stone[1], Michaela Hayer[1], **Bruce A. Hungate**[1,2]

[1]Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ, 86011, USA; [2]Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, 86011, USA; [3]Department of Biology, West Virginia University, Morgantown, WV, 26506, USA; [4]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, 94550, USA

Project Goals: Short statement of goals. (Limit to 1000 characters)
The objectives of the proposed work are to 1) develop a quantitative SIP targeted –omics-based framework to infer taxon-specific rates of growth, mortality, and associated carbon (C) fluxes in soil microbial communities, a framework that will scale from individual microbial taxa to the integrated soil system; 2) develop a new means to quantify taxon-specific C-use and growth efficiency in order to quantify taxon-specific contributions to soil C loss; and 3) apply newly developed quantitative tools to test key hypotheses about responses of the soil microbial community to experimental warming and to latitudinal variation in temperature. This work will push the frontier of –omics enabled techniques by demonstrating their applicability at the ecosystem scale, and by relating taxon-specific isotope assimilation to dissimilatory processes in the C cycle, thereby enabling the identification of organisms especially responsible for soil C loss.

The genomic revolution in microbial ecology has encouraged optimism that genomic traits of microorganisms will coalesce into coherent ecological strategies which enable predictions of the ecosystem processes that soil microorganisms perform in nature. Relationships between microbial genes and performance are evaluated in the laboratory in pure cultures, with little validation in nature. Here, we show that genomic traits related to maximum growth potential failed to predict the growth rates of bacteria in unamended soil, but successfully predicted growth responses to resource pulses: growth increased with 16S rRNA gene copy number and declined with genome size after substrate addition to soils, responses that were repeated in four different ecosystems. Genome size best predicted growth rate in response to addition of glucose alone; adding ammonium with glucose weakened the relationship, and the relationship was absent in nutrient-replete pure cultures, consistent with the idea that reduced genome size is a mechanism of nutrient conservation. Our findings demonstrate that genomic traits of soil bacteria can map to their ecological performance in nature when resources are abundant, but not when scarce, when genomic traits related to stress tolerance may prove more predictive. These results remind that the phenotype depends on environmental context, underscoring the importance of verifying proposed schemes of trait-based strategies through direct

measurement of performance in nature, an important and currently missing foundation for translating microbial processes from genes to ecosystems.

Microbiology has advanced by studying microorganisms in culture, and improved culturing strategies will undoubtedly support new insights about quantitative variation in important traits of microorganisms. Ecology has advanced by studying organisms in nature, accepting environmental heterogeneity and community interactions as essential features of the world organisms inhabit. Microbial ecology should bridge these approaches, but the frequency of speculation in the field based on principles derived from pure culture laboratory studies is belied by the paucity of direct tests of those principles in nature, where community interactions are intact, resource availability is native, and environmental stressors apply. Results presented here show how microbial ecology can advance beyond the pure culture paradigm by measuring quantitative trait variation of microorganisms in the habitats where they naturally occur. Techniques like quantitative stable isotope probing can evaluate where principles derived from the laboratory apply to microorganisms in nature, and where they fail. It is not surprising that growth responses to resource pulses corresponded with traits derived from studies of resource-rich laboratory cultures, and that the translation of genomic traits related to high resource abundance failed in unamended soils. At the same time, the high variation in growth rates observed without resource amendment points to important phenotypic variation in growth, even under resource-limited conditions, and the need to explore the genomic traits and environmental conditions that drive that variation.

Assigning ecological strategies based on taxonomy is a common approach for interpreting microbiome data, but this effort to date is largely divorced from measurements testing whether the organisms actually utilize their assigned strategies in nature. Growth is a useful metric for evaluating ecological strategies, because it integrates ecological and evolutionary processes, from metabolism, to resource uptake and use and thus the imprint of biology on element cycles, to fitness, the ultimate result of variation in genomic traits. Assessing growth in natural microbial communities, combined with molecular tools, provides access to a richer suite of ecological mechanisms influencing organismal performance than can be assessed in laboratory cultures. Findings presented here show that it is now possible to pair genomic traits of microorganisms with their growth rates in nature. Such efforts hold promise for a refined and evidence-based foundation for proposed ecological strategies, whether best defined by a single axis of copiotrophic to oligotrophic, by a triangle of competitive, ruderal, and tolerance, or by multi-dimensional spectra of traits and tradeoffs.

Notes

1. This work is currently in review in the journal, ISME.

Funding Statement

This work was supported by the US Department of Energy, Program in Genomic Sciences, Award Number: DE-SC0016207. We thank Jingrun Sun for assistance in the laboratory and Alicia Purcell for intellectual contributions.

Evolutionary history constrains microbial traits across environmental variation

Ember M. Morrissey[1]*(ember.morrissey@mail.wvu.edu), Rebecca L. Mau[2,3], Michaela Hayer[2], Xiao-Jun Allen Liu[2,4,5], Egbert Schwartz[2,4], Paul Dijkstra[2,4], Benjamin J. Koch[2,4], Kara Allen[6], Steven J. Blazewicz[7], Kirsten Hofmockel[8], Jennifer Pett-Ridge[7], **Bruce A. Hungate**[2,4]

[1]Division of Plant and Soil Sciences, West Virginia University, Morgantown, WV, 26506 [2]Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ, 86011, USA [3]Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, AZ 86004, USA [4]Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, 86011, USA [5]Department of Microbiology, University of Massachusetts, Amherst, MA, 01003, USA[6]Department of Biology, West Virginia University, Morgantown, WV, 26506, USA [7]Physical and Life Sciences Directorate, Lawrence Livermore National Lab, Livermore, CA 94551 [8]Environmental Molecular Science Laboratory, Pacific Northwest National Laboratory, Richland, WA, 99354

Project Goals: Connecting the composition of microbial communities with biogeochemical process rates has the potential to improve our understanding of, and ability to model, ecosystem function. This project aimed to determine if the growth of prokaryotic taxa is consistent across ecosystems. This work dovetails with our our larger goal to characterize *in situ* rates of biogeochemically significant microbial activity at the community scale and for specific taxa. This information will be leveraged to establish whether there is a "phylogenetic imprint" on soil carbon and nitrogen cycling processes that can facilitate better incorporation of microbial data into process-scale modeling efforts.

Organisms influence ecosystems, from element cycling to disturbance regimes, to trophic interactions, to energy partitioning. Microorganisms are part of this influence, but the field of microbiology has been historically limited to inferring functional potential from traditional pure culture investigations or genome-based, qualitative analyses. Understanding the ecology of microorganisms in nature requires studying traits of these organisms quantitatively in their natural habitats, a challenging task, but one which new approaches now make possible. Here, we show that growth rate and carbon assimilation rate of soil microorganisms are more influenced by taxonomy than by climate, even across a broad climatic gradient spanning major temperate life zones, from mixed conifer forest to high desert grassland. Taxonomy accounted for ~50% to ~90% of the explained variation in growth rate and carbon assimilation rate, and was more predictive for organisms responding to resource addition. With added carbon and nitrogen substrates, taxonomy explained ~8 times more variance in growth rate than did the ecosystem type. Taxon-specific growth and carbon assimilation rates were highly intercorrelated across the four ecosystems, constrained by the taxonomic identity of the organisms, such that plasticity driven by climate was limited across ecosystems varying as much as 4.5°C in mean annual temperature and 560 mm in mean annual precipitation. Taken together our results suggest that, similar to multicellular life, the traits of prokaryotes in their natural habitats are constrained by evolutionary history to a greater degree than environmental variation.

This research was supported by the Office of Biological and Environmental Research in the DOE Office of Science

Conserved Genetic Mechanisms for Biotic Stress in Sorghum

Tiffany Jamann[1*](tjamann@illinois.edu), Bill Rooney, [2] and Santiago Mideros,[1]

[1]University of Illinois, Urbana; [2]Texas A&M University, College Station

Project Goals: Short statement of goals. (Limit to 1000 characters)

Developing durable disease resistance for biofuel crops is crucial, particularly as the range of biofuel crop production expands and pathogens of other plant species evolve to cause diseases of bioenergy feedstocks. *Setosphaeria* species are significant fungal pathogens of the Andropogonae, and *S. turcica* can infect both maize and sorghum, making it a strong candidate to be an increasing problem for biofuel sorghum. Our overall objective is to gain a systems-level understanding of the pathosystem by leveraging natural genetic variation, host specificity of the pathogen, and transcriptome analysis to improve biotic stress resistance in sorghum. The proactive strategy of paired identification of fungal effectors and plant resistance genes in a pathosystem with a high likelihood of producing a host jump is a paradigm shift in disease management through host resistance.

This work is funded by DOE award number DE-SC0019189 (Plant Feedstocks Genomics).

Soil Moisture Modulates Inter-kingdom Interactions as Visualized Using a Simulated Soil Core

Christopher R. Anderton[1*] (christopher.anderton@pnnl.gov), Arunima Bhattacharjee, Sheryl L. Bell[1], Thomas W. Wietsma[1], Dušan Veličković[1], Rosalie K. Chu[1], **Kirsten S. Hofmockel[1], Janet K. Jansson[2]**

[1]Environmental Molecular Sciences Division; [2]Biological Sciences Division, Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA

https://www.pnnl.gov/biology/programs/

Project Goals: PNNL's Soil Microbiome SFA aims to achieve a systems-level understanding of the soil microbiome's phenotypic response to changing moisture through spatially explicit examination of the molecular and ecological interactions occurring within and between members of microbial consortia. Integrated experiments were designed to confront both the scaling challenges and inter-kingdom interactions that regulate networks of biochemical reactions. Individual- and population-based models for predicting interspecies and inter-kingdom interactions were parameterized using experimental data, and predictions were tested in soil to reveal spatially explicit microbial interactions. Discoveries from controlled experiments are planned to be cross validated in the field, using moisture gradient experiments at a new local field site. Data was captured and shared through an optimized data management pipeline. Knowledge gained will provide fundamental understanding of how soil microbes interact to decompose organic carbon and enable prediction of how biochemical reaction networks shift in response to changing moisture regimes.

Abstract: Understanding the basic biology that underpins soil microbiome interactions is required in order to predict the metaphenomic response to environmental shifts, such as changing moisture content. A significant knowledge gap is how such changes will affect microbial community structure and its metabolic landscape. Here, we visualized the metabolome of interacting organisms within the soil habitat by obtaining high resolution multidimensional maps of the compositional and functional state of soil microbial communities. This entails mapping the metabolic exchanges that occur within soil microbiomes using the mass spectrometry imaging (MSI) methodologies we recently developed.[1,2] Using a custom-built Soil Box system, we demonstrate changes in microbial community growth and composition in different soil environments, contingent upon access to reservoirs of nutrient sources. The Soil Box, designed to press functionalized-glass slides against the soil surface at different depths from the top soil surface, emulates the probing depth of a common soil core. This enabled determining both the spatial organization of the microbial communities that form on the slides and their metabolites by using confocal microscopy in combination with matrix-assisted laser desorption ionization (MALDI)-MSI.

We found that increased adhesion of soil microbial biomass occurred on slides functionalized with chitin islands, with the seeding attachment event occurring as early as 2 h. The MALDI-MSI data showed a high abundance of bacterial-related lipid families on the chitin islands and low abundance on areas without chitin. Confocal microscopy measurements of these samples confirmed the increased growth of microbial biomass, and consumption of chitin during growth. The microbial growth and community dynamics were also sampled at different moisture regimes (i.e., 14%, 24%, and 34% of total soil weight). Fungal hyphal networks bridging different chitin islands over distances of 17 mm were observed only in the driest of conditions, indicating that such bridges may act as fungal highways during drought conditions. In all, through the use of multiple correlative imaging platforms, these results illustrate a system that provides unprecedented spatial information about interactions within soil microbial communities as a function of changing environments. We anticipate that further use of our Soil Box approach will be invaluable in probing specific intra- and inter-kingdom metabolic networks arising from a gradient of environmental stresses.

References

1. C.R. Anderton, R.K. Chu, N. Tolic, A.V. Creissen, L Pasa-Tolic, "Utilizing a Robotic Sprayer for High Lateral and Mass Resolution MALDI FT-ICR MSI of Microbial Cultures," *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 556
2. D. Veličković, R. K. Chu, A. A. Carrell, M. Thomas, L. Paša-Tolić, D. J. Weston, C. R. Anderton, "Multimodal MSI in conjunction with broad coverage spatially resolved MS^2 increases confidence in both molecular identification and localization," *Anal. Chem.* **2018**, *90*, 702

Funding statement: *This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (OBER), as part of BER's Genomic Science Program (GSP), and is a contribution of the Pacific Northwest National Laboratory (PNNL) Soil Microbiome Scientific Focus Area "Phenotypic Response of the Soil Microbiome to Environmental Perturbations." A portion of this work was performed in the William R. Wiley Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility sponsored by OBER and located at PNNL. PNNL is a multi-program national laboratory operated by Battelle for the DOE under Contract DE-AC05-76RLO 1830.*

Determining the Soil Microbiome Community Phenotype "Metaphenome" in Response to Changing Soil Moisture and Carbon Content

Janet K. Jansson[1*] (janet.jansson@pnnl.gov) and **Kirsten S. Hofmockel**[2]

[1]Biological Sciences Division; [2]Environmental Molecular Sciences Division, Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA

https://www.pnnl.gov/biology/programs/

Project Goals: PNNL's Soil Microbiome SFA aims to achieve a systems-level understanding of the soil microbiome's phenotypic response to changing moisture through spatially explicit examination of the molecular and ecological interactions occurring within and between members of microbial consortia. Integrated experiments were designed to confront both the scaling challenges and inter-kingdom interactions that regulate networks of biochemical reactions. Individual- and population-based models for predicting interspecies and inter-kingdom interactions were parameterized using experimental data, and predictions were tested in soil to reveal spatially explicit microbial interactions. Discoveries from controlled experiments are planned to be cross validated in the field, using moisture gradient experiments at a new local field site. Data was captured and shared through an optimized data management pipeline. Knowledge gained will provide fundamental understanding of how soil microbes interact to decompose organic carbon and enable prediction of how biochemical reaction networks shift in response to changing moisture regimes.

Abstract: As climate changes, the mesic grassland ecosystems of the continental United States are predicted to experience increasing periods of drought. The influence of extended drought on functions carried out by interacting members of microbial communities across trophic scales is largely unknown but vital for understanding and predicting outcomes of future climate regimes on soil health and biofuel feedstock sustainability. PNNL's soil microbiome SFA has tackled this challenge using a cross-scale approach to decipher the molecular mechanisms underpinning interactions between soil microorganisms. Within this context, our team is working towards defining how water stress affects community functions, interkingdom interactions, and the decomposition rates of chitin, a highly abundant ubiquitous polysaccharide found in soil.

Starting at the field scale, we screened existing soil metagenomes to determine the global distribution of soil viruses and used this data to link viruses with their hosts. We identified thousands of novel soil viruses, including giant viruses, as well as novel auxiliary metabolic genes (AMGs), thus illuminating the potential roles of soil viruses. In addition, we deeply sequenced metagenomes (~6 Tb

total) from disparate grasslands (Kansas, Iowa, and Washington) to determine the functional potential of the different soils. To understand molecular mechanisms underlying signaling in the field, we identified novel siderophores[1,2]. Concurrently, using novel metabolic modeling approaches we revealed that wet and dry grassland soils differ in key metabolic pathways. These findings will provide targets for future investigations at our local field site, where we established an irrigation field trial. Also, at our local field site, we planted a new bioenergy field stock (Tall Wheatgrass) to determine the influence of soil moisture and drought on the soil microbiome and plant-microbe interactions.

At the other end of the scale, we dissected the complex soil microbiome into tractable, stable, low complexity model consortia. Using our SFA consortia, we demonstrated the influence of species richness on community convergence and stability. These consortia provide a valuable resource for determining specific metabolic and species interactions which can be validated at our local field trial. Experiments with the SFA consortia also revealed the importance of spatial structuring. We visualized spatial interactions between specific soil populations and metabolic interactions during chitin decomposition, using newly developed soil microfluidic chips (Soil Chips) and Soil Boxes. We found that specific soil fungi help to bridge soil pores as soil dries down and that soil chemistry plays a vital role in this process. To determine which of the species synthesize and secrete chitinolytic enzymes we developed a suite of Activity-Based Probes (ABPs) to selectively label the active subset of chitinolytic enzymes. These data will be used to parameterize individual- and population-based models currently in development.

At the mesoscale, we examined the soil metaphenomic response of native soil communities in response to varying moisture regimes. Using real-time CO_2 measurements as the response variable, we investigated chitin degradation using multi-omics to determine the functional response of specific members of the soil microbiome. Together these studies are moving us towards gaining an understanding of the microbial community's phenotypic response "metaphenome"[3] to soil moisture and drought.

References

1. *Boiteau et al. 2018. Soil Biol. Biochem. 120:283-291*
2. *Boiteau et al. 2018. Metallomics. in press. DOI: 10.1039/c8mt00252e*
3. *Jansson and Hofmockel. 2018. Curr Op Microbiol. 43:162-168.*

Funding statement: *This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (OBER), as part of BER's Genomic Science Program (GSP), and is a contribution of the Pacific Northwest National Laboratory (PNNL) Soil Microbiome Scientific Focus Area "Phenotypic Response of the Soil Microbiome to Environmental Perturbations." A portion of this work was performed in the William R. Wiley Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility sponsored by OBER and located at PNNL. PNNL is a multi-program national laboratory operated by Battelle for the DOE under Contract DE-AC05-76RLO 1830.*

Probing the Relationship Between Osmolytes and Respiration in Soils - in Real Time

Mary S. Lipton[1*] (mary.lipton@pnnl.gov), Karl K. Weitz[2], Sneha P. Couvillion[2], Montana L. Smith[1], Thomas O. Metz[2], Emily B. Graham[2], Sheryl L. Bell[1], Vanessa L. Paurus[2], **Kirsten S. Hofmockel[1], Janet K. Jansson[2]**

[1]Environmental Molecular Sciences Division [2]Biological Sciences Division; Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA

https://www.pnnl.gov/biology/programs/

Project Goals: PNNL's Soil Microbiome SFA aims to achieve a systems-level understanding of the soil microbiome's phenotypic response to changing moisture through spatially explicit examination of the molecular and ecological interactions occurring within and between members of microbial consortia. Integrated experiments were designed to confront both the scaling challenges and inter-kingdom interactions that regulate networks of biochemical reactions. Individual- and population-based models for predicting interspecies and inter-kingdom interactions were parameterized using experimental data, and predictions were tested in soil to reveal spatially explicit microbial interactions. Discoveries from controlled experiments are planned to be cross validated in the field, using moisture gradient experiments at a new local field site. Data was captured and shared through an optimized data management pipeline. Knowledge gained will provide fundamental understanding of how soil microbes interact to decompose organic carbon and enable prediction of how biochemical reaction networks shift in response to changing moisture regimes.

Soil is a highly diverse ecosystem with microbial and chemical dark matter that remains to be discovered. While this is an exciting frontier for discovery, it also represents an obstacle to deciphering predictable interactions that promote fertility and productivity of our nation's soils. Environmental stress from drought is increasing in frequencing with unknown outcomes for soil microbiomes. The mechanisms in which the microbial communities respond to dehydration is very important, particularly in arid and marginal lands. As soils dry, the water potential across the cell wall decreases leading to osmotic stress which is compounded by limited diffusion. To address environmental changes caused by differences in moisture level, microbes must must shift physiology and metabolic interactions to survive. One mechanism microbes employ to offset decreases in moisture levels is the production of osmolytes to reduce their internal water potential and maintain fluid balance. Mostly, microbes use simple, soluble organic molecules such as amino compounds (amino acids and glycine betaine) and simple sugars (trehalose). However, little is understood about the molecular response of the microbiome to the changes in soil moisture, and how that affects the phenotype.

Here, we aimed to understand how the physiology, metabolism, and interactions of soil microbes change in response to moisture, and to use this understanding as a basis for predicting the soil metaphenome. We are testing the hypothesis that during dessication in arid and marginal soil collected from our local Washington field site, microbes will initially accumulate osmolytes only to rapidly metabolize them upon rewetting. In our initial experiment, we incubated soils at 10% moisture and harvested after 2 weels. The baseline GC-MS based analysis of the metabolomic extract from the soils revealed the presence of potential osmolytes including trehalose, glycine and other sugars and amino acids. We have also used the same GC-MS approaches to monitor the changes in the abundance of these metabolites over time during a rewetting event where after rewetting dried soil to 19% moisture was harvested at particular timepoints over a 3 hour period. The changes in the osmolytes reveal potential microbial mechanisms active after the rewetting event.

One mechanism for the disposal of the osmolytes is the rapid release of CO_2, DOC and nutrients, arising from intracellular material and varies in magnitude based on the biomass. However, most analyses focused to date have studied this event by measuring respiration on the hour to day time frame, thus missing the immediate rewetting events. To provide information about the initial microbial response to wetting, we developed an atmospheric monitoring instrument that measures the levels of CO_2, O_2, N_2, H_2O and other relevant gases, simultaneously in real time, seconds, minutes and hours after a rewetting event (Figure 1). Using this apparatus, we have evaluated the effect of hydration level on the rate and amplitude of the initial respiration rate after rewetting revealing a linear relationship between the amount of water added and the rate and amount of CO_2 produced. Substrates (glucose, cellulose, chitin, plant exudates) can also be added to the reaction either by being dissolved in the H_2O that is added or directly mixed with the soil prior to the reaction initiation depending on the solubility of the substrate and experimental design. Additionally, O_2, H_2O, $^{12}CO_2$ and $^{13}CO_2$ production can be monitored simultaneously throughout the reaction for a more thorough understanding of the respiration effect.

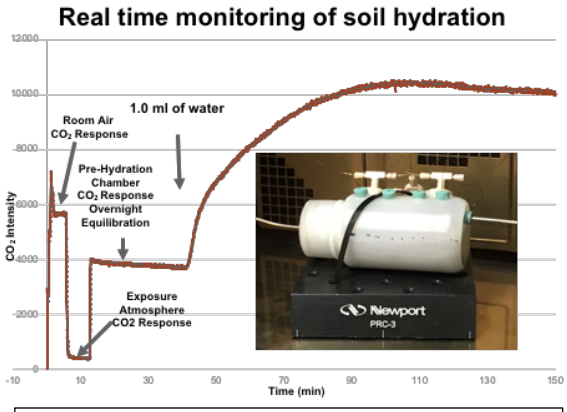


Figure 1: Picture of the gas monitoring reaction chamber (inset) and the rate and extent of CO_2 production over time as measuremed every 2 sec.

Combining data outputs from the standard baseline GC-MS approach to elucidate the metabolomic profile of the soils over time and the novel mass spectrometry based continuous gas monitoring system will further elucidate the relationship between osmolyte production, metabolism respiration rate in the soil after a rewetting event. These data were then used to better understand the metabolic pathways active in the soil microbiome and were linked with other omics data for the creation of more comprehensive models of microbial activity.

Funding statement: *This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (OBER), as part of BER's Genomic Science Program (GSP), and is a contribution of the Pacific Northwest National Laboratory (PNNL) Soil Microbiome Scientific Focus Area "Phenotypic Response of the Soil Microbiome to Environmental Perturbations." A portion of this work was performed in the William R. Wiley Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility sponsored by OBER and located at PNNL. PNNL is a multi-program national laboratory operated by Battelle for the DOE under Contract DE-AC05-76RLO 1830.*

Development and Analysis of Reduced Complexity Microbial Consortia Emerging from Native Grassland Soil Systems

Ryan McClure[1]* (Ryan.McClure@pnnl.gov), Dan Naylor[1], Colin Brislawn[1], Yuliya Farris[1], Sarah Fansler[1], **Kirsten S. Hofmockel[2] and Janet K. Jansson[1]**

[1]Biological Sciences Division; [2]Environmental Molecular Sciences Division, Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA

https://www.pnnl.gov/biology/programs/

Project Goals: PNNL's Soil Microbiome SFA aims to achieve a systems-level understanding of the soil microbiome's phenotypic response to changing moisture through spatially explicit examination of the molecular and ecological interactions occurring within and between members of microbial consortia. Integrated experiments were designed to confront both the scaling challenges and inter-kingdom interactions that regulate networks of biochemical reactions. Individual- and population-based models for predicting interspecies and inter-kingdom interactions were parameterized using experimental data, and predictions were tested in soil to reveal spatially explicit microbial interactions. Discoveries from controlled experiments are planned to be cross validated in the field, using moisture gradient experiments at a new local field site. Data was captured and shared through an optimized data management pipeline. Knowledge gained will provide fundamental understanding of how soil microbes interact to decompose organic carbon and enable prediction of how biochemical reaction networks shift in response to changing moisture regimes.

Abstract:

Soil microbial communities are critical to the overall carbon cycle and to the decomposition of complex biopolymers such as chitin and cellulose. Despite the critical nature of these microbiomes, a detailed understanding of how the interactions between members lead to emergence of community functions is lacking. This is due, in part, to the complex nature of the soil microbiome with thousands of species across several kingdoms contributing to the overall response of soil. In order to gain a more detailed view of the soil environment, we took an approach based on developing and analyzing reduced complexity microbial consortia that contain fewer species than the native soil but are still representative of this site and are more experimentally tractable (**Figure 1**). We hypothesize that predictable reaction modules exist and that analysis of these model consortia can identify these reaction modules, and other fundamental aspects of soil microbiome interactions. These interactions and aspects can then be tested in a specific, robust manner in the native soil sites, greatly expanding our knowledge of the functioning of soil microbiomes.

To generate reduced complexity soil microbial consortia, we collected samples from our native field site containing a grassland silt loam soil. This native soil was then diluted to various levels, ranging from 10^{-1} to 10^{-4} and cultured in irradiated sterile native soil. Cultivation was carried out with the monomer of chitin, N-acetyl-glucosamine (NAG), exogenously added at 100 ppm. We confirmed that initial dilutions decreased the richness and complexity of our native soil and found that culturing these dilutions in soil for 15 weeks lead to stable microbial consortia containing both bacteria and fungi. We also found that emerging communities were highly representative of the native soil site. These stable consortia contained significantly reduced membership with between 250-300 OTUs (compared to the native soil containing ~2000 OTUs), while still containing a number of phyla that are characteristic of soil such as

Verrucomicrobia and Actinobacteria. In addition to cultivation in soil, we also developed consortia on plates using chitin itself as a major source of carbon and nitrogen. These plate communities represented microbiomes that were even further reduced in complexity while again containing a diverse community of several different phyla. Plate communities contained between 20-70 OTUs, richness levels that are far easier to interrogate with –omics approaches, showed significant bioactivity when moved back into a soil environment as measured by CO_2 production and also showed stable levels of taxa across time.

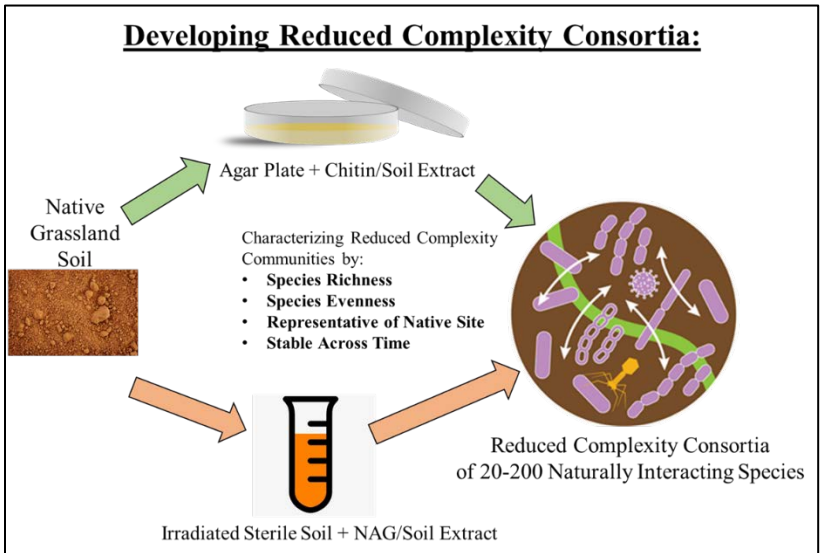


Figure 1. Developing reduced complexity Consortia. Soil was collected from native grasslands and diluted before being cultured on either plates or in native soil for several weeks with chitin or NAG (chitin monomer). Culturing lead to a stable consortia of 20-200 naturally interacting species

This series of reduced complexity consortia that are representative of native soil sites are powerful tools that can be used by the soil community at large to interrogate the response of soil microbiomes to a number of perturbations and to confirm critical interactions between microbial species, particularly inter-kingdom interactions that characterize the emergent behavior of soil microbiomes. Constructed reduced complexity consortia provide a means to more powerfully leverage high-throughput, multi-omic techniques to better characterize these interactions and the major constituent players that are a part of them. Further knowledge of these interactions will help us better understand the overall metaphenome of soil systems, especially as they respond to critical perturbations including drought.

Funding statement: *This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (OBER), as part of BER's Genomic Science Program (GSP), and is a contribution of the Pacific Northwest National Laboratory (PNNL) Soil Microbiome Scientific Focus Area "Phenotypic Response of the Soil Microbiome to Environmental Perturbations." A portion of this work was performed in the William R. Wiley Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility sponsored by OBER and located at PNNL. PNNL is a multi-program national laboratory operated by Battelle for the DOE under Contract DE-AC05-76RLO 1830.*

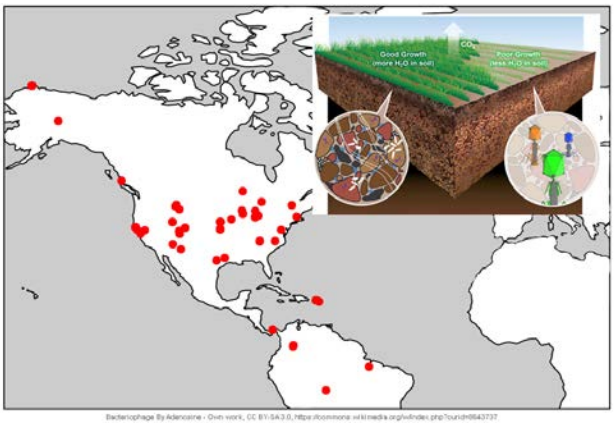
Untapped Viral Diversity in Global Soil Metagenomes

Jason E. McDermott[1,2]* (Jason.McDermott@pnnl.gov), Emily B. Graham[1], David Paez-Espino[3], Colin Brislawn[1], Russell Y. Neches[3], Ruonan Wu[1], Nikos C. Krypides[3], **Kirsten S. Hofmockel[1,4], and Janet Jansson[1]**

[1]Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA; [2]Department of Molecular Microbiology and Immunology, Oregon Health & Sciences University, Portland, OR; [3]Joint Genome Institute, Walnut Creek, CA; and [4]Environmental and Molecular Sciences Laboratory, Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA

https://www.pnnl.gov/biology/programs/

Project Goals: PNNL's Soil Microbiome SFA aims to achieve a systems-level understanding of the soil microbiome's phenotypic response to changing moisture through spatially explicit examination of the molecular and ecological interactions occurring within and between members of microbial consortia. Integrated experiments were designed to confront both the scaling challenges and inter-kingdom interactions that regulate networks of biochemical reactions. Individual- and population-based models for predicting interspecies and inter-kingdom interactions were parameterized using experimental data, and predictions were tested in soil to reveal spatially explicit microbial interactions. Discoveries from controlled experiments are planned to be cross validated in the field, using moisture gradient experiments at a new local field site. Data was captured and shared through an optimized data management pipeline. Knowledge gained will provide fundamental understanding of how soil microbes interact to decompose organic carbon and enable prediction of how biochemical reaction networks shift in response to changing moisture regimes.



Soil viruses were determined from existing soil metagenomes from >600 soil samples collected across 3 continents.

PNNL's soil microbiome SFA aims to understand the importance of interactions between different kingdoms (Bacteria, Archaea, Fungi and Viruses) on carbon decomposition and how changes in soil moisture, such as expected with drought, influence those interactions. Here, we specifically aimed to determine the role of soil viruses. Although soil viruses are particularly diverse in comparison to other habitats[1], we have limited understanding of their identities and distribution. The study of soil viruses is

particularly challenging because of the enormous range of soil ecosystems, and until recently, the lack of appropriate molecular screening tools. We aimed to overcome these challenges by screening hundreds of existing soil metagenomes that were deposited at the Joint Genome Institute (JGI) for the presence of viruses. We found that the majority of viral diversity in soil is completely novel, including new giant viruses. Additionally, we screened the viruses for auxiliary metabolic genes (AMGs) and found several AMGs that corresponded to enzymes involved in carbon decomposition. We also linked viruses to their hosts using co-occurrence networks and CRISPR spacer sequence similarity.

To identify patterns of viral biogeographic distribution we manually assigned biomes to each study using metadata deposited into Integrated Microbial Genomes with Microbiomes (IMG/M)[2]. Biomes considered included arctic soil, cultivated and uncultivated prairie, rhizosphere, and Canadian, Mediterranean, or tropical forests. We found that estimated viral abundances varied dramatically between these biomes, and that inferred biogeographic relationships varied as well. For example, comparisons between cultivated and native grasslands revealed that, in contrast to bacterial diversity[3], diversity of viral-encoded AMGs decreased after cultivation.

We also found that most metagenomics datasets displayed a linearly increasing relationship between overall nucleotides sequenced and viral nucleotides sequenced indicating that viral diversity is significantly under-sampled. However, our study greatly advances existing knowledge of soil viral diversity and abundance and highlights the vast, uncharacterized viral diversity in soil microbiomes and their potentially important roles in carbon cycling.

References

1. Paez-Espino, D. et al. Uncovering Earth's Virome. Nature 536, 425 (2016).
2. Markowitz, V. M. et al. IMG/M: a data management and analysis system for metagenomes. Nucleic acids research 36, D534-D538 (2007)
3. Mackelprang, R. et al. Response of the soil microbiome to cultivation in native tallgrass prairie soils of the Midwestern United States. Frontiers in microbiology 9, 1775 (2018).

Funding statement: *This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (OBER), as part of BER's Genomic Science Program (GSP), and is a contribution of the Pacific Northwest National Laboratory (PNNL) Soil Microbiome Scientific Focus Area "Phenotypic Response of the Soil Microbiome to Environmental Perturbations." A portion of this work was performed in the William R. Wiley Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility sponsored by OBER and located at PNNL. PNNL is a multi-program national laboratory operated by Battelle for the DOE under Contract DE-AC05-76RLO 1830.*

Integrative Network Modeling for Predicting Biochemical Signatures in Complex Soil Systems

Hyun-Seob Song,[1]* (hyunseob.song@pnnl.gov), Joon-Yong Lee,[1] Ryan McClure,[1] **Kirsten S. Hofmockel**[2], and **Janet K. Jansson**[1]

[1]Biological Sciences Division; [2]Environmental Molecular Sciences Division, Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA

https://www.pnnl.gov/biology/programs/

Project Goals: PNNL's Soil Microbiome SFA aims to achieve a systems-level understanding of the soil microbiome's phenotypic response to changing moisture through spatially explicit examination of the molecular and ecological interactions occurring within and between members of microbial consortia. Integrated experiments were designed to confront both the scaling challenges and inter-kingdom interactions that regulate networks of biochemical reactions. Individual- and population-based models for predicting interspecies and inter-kingdom interactions were parameterized using experimental data, and predictions were tested in soil to reveal spatially explicit microbial interactions. Discoveries from controlled experiments are planned to be cross validated in the field, using moisture gradient experiments at a new local field site. Data was captured and shared through an optimized data management pipeline. Knowledge gained will provide fundamental understanding of how soil microbes interact to decompose organic carbon and enable prediction of how biochemical reaction networks shift in response to changing moisture regimes.

Abstract: Soil microbiomes provide diverse ecosystem services that are vital for life on our planet. A greater understanding of how microbe-regulated soil biochemistry shifts in varying environments is critical for ensuring stable nutrient cycling and healthy ecosystem functioning. In an effort to achieve this goal, scientists have been keen to identify molecular signatures that may characterize soil biochemistry uniquely associated with specific environmental conditions. This requires a systems approach that integrates advanced experimental and modeling technologies[1], but generalizable protocols have not been fully established yet. Central challenges include the lack of coordinated collection of multi-omics data and the difficulty in building predictive models that can successfully translate molecular-level data to soil metaphenomes[2].

Leveraging PNNL's multi-omics profiling technology, here we present a general data-model integration approach that enables predicting biochemical signatures of complex soil systems. Our approach combines metabolic and molecular networks that are constructed from multi-omics data. Complementary predictions from these two networks make their integration synergistic: 1) the metabolic network predicts condition-specific reaction pathways and modules;

2) the molecular network identifies molecular clusters that show coordinated response to perturbations.

We demonstrated the proposed approach through a case study of native prairie soils sampled from multiple field locations in Kansas, which were subjected to extreme moisture perturbations. First, we predicted reaction pathways/modules specifically associated with dry and wet conditions by incorporating gene expression and metabolite profiles into a master metabolic network constructed by assembling all chemical reactions available in public databases such as KEGG and ModelSEED. Using the same omics data, we also constructed a molecular interaction network to identify clusters of strongly correlated genes and metabolites. Finally, we combined these outputs to determine molecular signatures, i.e., sets of condition-specific genes/metabolites that show similar response patterns to moisture perturbations (**Figure 1**). Those signatures were predicted particularly for dry soils.

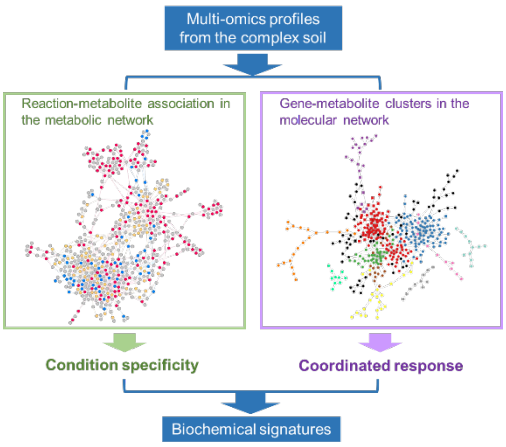


Figure 1. Integration of multi-omics data (top), a metabolic network (left), and a molecular network (right) to define biochemical signatures

The general protocol of model building and integration proposed this work should be applicable to many other environmental systems. The metabolic network construction framework is currently being extended to incorporate high-resolution organic carbon profiles generated from high-accuracy mass spectrometry. This unprecedented development can significantly improve our ability to predict condition-specific biochemical footprints of complex environmental systems.

References

1. Song H-S, Cannon W, Beliaev A, Konopka A. Mathematical modeling of microbial community dynamics: a methodological review. *Processes* **2**, 711-752 (2014).
2. Jansson JK, Hofmockel KS. The soil microbiome - from metagenomics to metaphenomics. *Curr Opin Microbiol* **43**, 162-168 (2018).

Funding statement: *This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (OBER), as part of BER's Genomic Science Program (GSP), and is a contribution of the Pacific Northwest National Laboratory (PNNL) Soil Microbiome Scientific Focus Area "Phenotypic Response of the Soil Microbiome to Environmental Perturbations." A portion of this work was performed in the William R. Wiley Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility sponsored by OBER and located at PNNL. PNNL is a multi-program national laboratory operated by Battelle for the DOE under Contract DE-AC05-76RLO 1830.*

Poplar Esterified Cell Wall Transformations and Metabolic Integration (PECTIN) Study

Jenny Mortimer,[1] Nancy Washton,[2] Mary S. Lipton,[2] Wellington Muchero,[3] Breeanna Urbanowicz,[4] Nate McDowell[5], Alex Guenther[6], and **Kolby Jardine** (kjjardine@lbl.gov)[1*]

[1]Lawrence Berkeley National laboratory, Berkeley, CA; [2]Environmental Molecular Sciences Laboratory (EMSL), Richland, WA; [3]Oak Ridge National Laboratory, Oak Ridge, TN; [4]Athens, GA; [5]Pacific Northwest National Laboratory, Richland, WA; and [6]University of California Irvine, Irvine, CA

http://cellwallesters.pbworks.com/w/page/127623629/FrontPage

Project Goals: The main goal of **PECTIN phase 1** is to quantify O-acetyl and methyl ester cell wall metabolism in poplar trees and its interactions with central metabolism and volatile metabolite exchange with the atmosphere. Critical to this goal is the development of novel methodology for xylem injections of ^{13}C-metabolite solutions involved in cell wall ester metabolism with injection flow rates scaled to sap flow. Using detached leaves to simulate drought and initiate senescence, a secondary goal is to determine quantitative relationships between the emissions of leaf methanol and acetic acid and changes in the degree of cell wall acetylation and methylation.

Abstract: Polysaccharides are major components of plant cell walls that can be converted into fuels by microbial fermentation, making plant biomass an important bioenergy resource. However, a substantial fraction of plant cell wall polysaccharides are chemically modified with methyl and acetyl groups that reduce yield of microbial fermentation. Although little is known about the biochemical and physiological functions of those cell wall modifications, it has been shown that their volatile intermediates (methanol and acetic acid) are tightly associated with plant growth, stress, and senescence processes but are not captured by traditional metabolomics analysis, representing an important gap in our knowledge of cell wall metabolism. This project studies the metabolism of those cell wall modifications and volatile intermediates as well as their role in central physiological processes in the emerging biofuel tree species California poplar (*Populus trichocarpa*) using field settings and controlled environmental conditions. The main goal of this research is to modify the expression of key genes involved in cell wall metabolism in order to reduce the amount of methyl and acetyl groups present on cell walls. These genetic modifications will be evaluated for potential impacts on important plant hydraulic and physiological processes including proper functioning of vascular tissues to support transpiration, leaf water potential and stomatal regulation, net photosynthesis, and high temperature/drought stress responses. Understanding and manipulating the metabolism of cell wall modifications will not only provide important knowledge on the physiology and ecology of plants but will also allow the generation of engineered bioenergy crops such as poplar for sustainable production of biofuels and bioproducts, addressing BER's goal of developing renewable bioenergy resources.

This research is supported by the Office of Science Early Career Research Program (FY18 DOE National Laboratory Announcement Number: LAB 17-1761), Topic: Plant Systems for the Production of Biofuels and Bioproducts (Program Manager: Pablo Rabinowicz Pablo.Rabinowicz@science.doe.gov)

Mining the largest collection industrially-deployed *Clostridium* strains for highly-evolved gene variants related to improved productivity and industrial robustness

Rasmus O. Jensen,[1] Steven D. Brown,[1] Eric (FungMin) Liew,[1] Marie Dawn Klingman,[2] Zamin K. Yang,[2] Nicole Shapiro,[3] Ashty Karim,[4] Sashini DiTissera,[1] Vinicio Reynoso,[1] Shilpa Nagaraju,[1] Asela Dassanayake,[1] Robert L. Hettich[2], Yasuo Yoshikuni,[3] Tanja Woyke,[3] Sean D. Simpson,[1] Timothy Tschaplinski,[2] Michael C. Jewett,[4] and **Michael Köpke**[1]*

[1]LanzaTech Inc, Skokie, IL; [2]Oak Ridge National Laboratory, Oak Ridge, TN; [3]Joint Genome Institute, Walnut Creek, CA; [4]Northwestern University, Evanston, IL

http://www.lanzatech.com; http://jewettlab.northwestern.edu

Project Goals: Even with the advent of next generation sequencing, there are limited options for many key metabolic genes considered for next-generation biofuel and bioproduct synthesis, and most genes found in public repositories are derived from type-strain or environmental sequences with unproven performance. To expand the pool of available sequences that are likely to result in high performance, we have sequenced and mined the largest collection of industrially-deployed *Clostridium* strains, evolved over several decades of intense development.

To rapidly prototype the performance of identified genes and to develop improved, industrial-robust production strains for conversion of lignocellulosic biomass to next-generation biofuels and bioproducts, we are establishing a new interdisciplinary venture, the clostridia Foundry for Biosystems Design (cBioFAB) that combines advancements in cell-free and *Clostridium* engineering metabolic engineering.

Clostridia have been commercially deployed in large-scale industrial fermentations for more than 100 years. The acetone-butanol-ethanol (ABE) fermentation has historically been the second largest industrial fermentation process only behind ethanol fermentation[1]. More recently, LanzaTech has commercialized a *Clostridium* based gas fermentation process for ethanol production, that utilizes CO/CO_2 containing gases such as syngas derived from lignocellulosic biomass as feedstock rather than sugar or starch that the ABE process requires as substrates[2].

High substrate costs eventually led to a decline of the ABE process in most Western countries in favor of petrochemical production, but development and commercial operation continued in politically isolated countries. LanzaTech owns the largest collection of industrially-deployed ABE strains. The collection dates back to 1944 and contains hundreds of highly evolved strains spanning several decades of development at commercial plants in South Africa and Taiwan as well as a number of research strains, including immunized strains[1].

In order to leverage the genetic diversity in this collection we have sequenced hundreds of strains from the collection using single-molecule, real-time (SMRT) sequencing, which we have

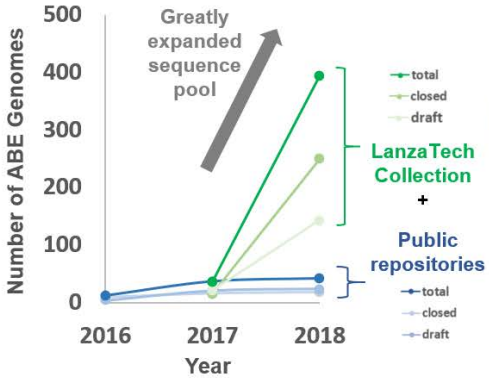
previously demonstrated to be effective in yielding high-quality genomes with a minimal number of contigs for complex Clostridium genomes.[2] To generate high-molecular weight, high-quality genomic DNA from all of the strains, we have developed a modified *Clostridium* growth medium and refined extraction protocols. Using this pipeline, we were able to generate more than 230 genomes with 1-6 contigs. To supplement this, we also sequenced over 100 genomes using Illumina chemistry.

All genomes were annotated using several different pipelines and mined for differences in key metabolic genes. On an amino acid level between 33 and 13 unique new sequences were mined for the genes in the core ABE pathway, with thiolase having the most and butyrate kinase the least diversity. The analysis resulted in discovery of new gene variants for the genome editing machinery and other genetic elements.

Of the identified genes, we have selected over 200 genes for synthesis. To rapidly prototype the performance of identified genes and benchmark them against wild-type genes, we will leverage a cell-free protein synthesis approach.[3] ABE products have successfully been engineered into gas-fermenting Clostridia and preliminary data shows a significant improvement in production from lignocellulosic biomass syngas using genes from the collection over wild-type genes.

NCP commercial ABE plant
1935-1982

Jones DT and Woods DR, Microbiol Rev (1986).



References

1. Jones DT and Woods DR (1986). Acetone-butanol fermentation revisited. *Microbiol Rev.* 50: 484–524.
2. Brown SD et al (2014). Comparison of single-molecule sequencing and hybrid approaches for finishing the genome of *Clostridium autoethanogenum* and analysis of CRISPR systems in industrial relevant clostridia. *Biotechnol Biofuels* 7:40
3. Karim AS and Jewett MC (2018) "Cell-free synthetic biology for pathway prototyping" *Methods Enzymology* 608: 31-57.

We acknowledge the U.S. Department of Energy Office of Science, Biological and Environmental Research Division (BER), Genomic Science Program for funding of this project under Contract No. DE-SC0018249. The work conducted by the U.S. Department of Energy Joint Genome Institute (JGI) is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 and the Community Science Program Award No. CSP-503280.

Development of a Clostridia Cell-free Platform Facilitating Accelerated Engineering of Clostridia Strains

Antje Krüger[1,2,3]* (antje.kruger@northwestern.edu), Alex Mueller[4], Michael Köpke[4], and **Michael C. Jewett**[1,2,3,5,6]

[1]Department of Chemical and Biological Engineering, [2]Chemistry of Life Processes Institute, [3]Center for Synthetic Biology, Northwestern University, Evanston, IL; [4]LanzaTech Inc., Skokie, IL; [5]Robert H. Lurie Comprehensive Cancer Center, [6]Simpson Querrey Institute, Northwestern University, Chicago, IL

http://jewettlab.northwestern.edu; http://www.lanzatech.com

Project Goals: We are addressing the complex challenge of designing, building, and optimizing biosynthetic pathways in cells in a new interdisciplinary clostridia Foundry for Biosystems Design (cBioFAB) with government and industry partners. Working both in vitro and in vivo, the goal is to expand the set of platform organisms that meet DOE bioenergy goals by utilizing and advancing state-of-the-art cell-free technologies, omics measurements, systems-biology analyses, computational modeling, and genome editing. cBioFAB will (i) reconceive how we engineer complex biological systems by linking pathway design, prospecting, validation, and production in an integrated framework, (ii) enable systems-level analysis of the David T. Jones collection, one of the largest collections of clostridia strains in the world, to uncover novel metabolic pathways, regulatory networks, and genome editing machinery, and (iii) realize more efficient clostridia strain engineering for the synthesis of biofuels and bioproducts.

Modern world challenges like rapid population growth, rising global living standards and its accompanying increase in energy demand and waste generation necessitate the production of low-cost fuels and high-value compounds from sustainable resources. Microbes can be engineered to produce biofuels, chemicals, materials, and therapeutics. Particularly attractive engineering targets are gas- and food waste-fermenting anaerobes like clostridia strains. Unfortunately, designing, building, and optimizing biosynthetic pathways in clostridia for manufacturing applications remain complex challenges. Cell-free protein synthesis (CFPS) platforms have the potential to accelerate biological design by speeding up design-build-test cycle. Here, we present the development of a clostridia CFPS system. We demonstrate its potential to facilitate rapid studying of metabolic pathway performance and prototyping in vitro. We expect that our new cell-free platform will accelerate engineering clostridia strains that efficiently convert wastes into high-value products.

We acknowledge the Department of Energy grant DE-SC0018249 for funding of this project.

Enhancing Control of Cell-free Metabolism Through pH Modulation

Blake Rasor[1,2,3*] (b.rasor@u.northwestern.edu), Ashty Karim[1,2,3], **Michael C. Jewett**[1,2,3,4,5]

[1]Department of Chemical and Biological Engineering, [2]Chemistry of Life Processes Institute, [3]Center for Synthetic Biology, Northwestern University, Evanston, IL;

[4]Robert H. Lurie Comprehensive Cancer Center, [5]Simpson Querrey Institute, Northwestern University, Chicago, IL

http://jewettlab.northwestern.edu/research/metabolic-engineering

Project Goals: We are establishing an interdisciplinary clostridia Foundry for Biosystems Design (cBioFAB) with government and industry partners to accelerate engineering efforts in non-model organisms through *in vitro* metabolic pathway prototyping, computational modeling, and integrated omics analysis. Through these diverse approaches, we seek to utilize genomic information from the David T. Jones clostridia collection to expand our knowledge of clostridia species and realize more efficient strain engineering for the synthesis of biofuels and bioproducts.

Cell-free systems allow the biosynthesis of enzymes and chemical products without the cell viability and growth constraints that hamper traditional metabolic engineering. Utilizing these systems to rapidly prototype metabolic pathways with many homologs *in vitro* enables the informed design of high-titer combinations of enzymes without constructing hundreds of unique bacterial strains. Such prototyping in *E. coli* extracts has proven useful for butanol biosynthesis in clostridia, a slow-growing, non-model organism with limited genetic tools available. However, prototyping pathways in the cytoplasmic milieu of *E. coli* is inherently different from the cytoplasm of clostridia. Such differences include the acidic environment preference of clostridia. Modifying the *in vitro* platform to better reflect clostridia could enhance its predictive ability, and the open environment of cell-free reactions allows for extensive control over the chemical composition. In this work, we altered the pH of cell-free reactions to provide a chemical environment more similar to extremophiles, including acidophilic clostridia. We found that cell-free reactions containing *E. coli* glycolytic enzymes and clostridial butanol synthesis enzymes consume glucose most rapidly at pH 8 and produce butanol most rapidly at pH 6, which is similar to the cytosolic pH of the respective bacteria. Additionally, this system can serve as a testbed for the pH tolerance of enzymes in a more biologically relevant context than purified protein in buffer. Alcohol dehydrogenases from other extremophiles convert butyraldehyde to butanol less efficiently than the canonical *C. acetobutylicum* homolog at acidic pH but show similar activity at alkaline pH. Overall, this enhances the utility of cell-free metabolic engineering to elucidate optimal chemical environments and sets of enzyme homologs to produce a desired product.

We acknowledge the Department of Energy grant DE-SC0018249 for funding of this project

Computational Tools to Predict Novel Chemical Reactions and Metabolic Network Behavior

Jake Martin, Joseph Ni, Linda Broadbelt, Keith E.J. Tyo* (k-tyo@northwestern.edu), and **Michael Jewett**

Northwestern University, Evanston, IL

https://pamspublic.science.energy.gov/WebPAMSExternal/Interface/Common/ViewPublicAbstract.aspx?rv=ec602522-ba87-436b-a2be-f70c5e95ef80&rtc=24&PRoleId=10

Project Goals: The *Clostridium* Foundry for Biosynthetic Design (cBioFab) goal is to provide tools and engineering strategies to enable high-level synthesis of next-generation biofuels and bioproducts from lignocellulosic biomass and expand the breadth of platform organisms that meet DOE bioenergy goals. The project combines *in vitro* (cell-free) and *in vivo* work to interweave and advance state-of-the-art pathway design, computational modeling, genome editing, and systems-biology analyses. The cBioFAB goal is to engineer complex biological systems by linking pathway design, prospecting, validation, and production in an integrated framework with system-level omics data.

Advances in high throughput biology have enabled the routine collection of large, comprehensive datasets, which cannot be easily interpreted manually. In the design-build-test-learn cycle, there is an urgent need to extract actionable conclusions from these datasets (i.e. learn) to enable more sophisticated and successful designs. The *Clostridium* Foundry for Biosystems Design (cBioFab) is reconceptualizing the way we engineer complex biological systems by linking pathway design, prospecting, validation, and production in an integrated framework that relies on computational modeling, cell-free technologies, and system-level omics data. Our computational efforts are focused on two important questions: (1) how to utilize high throughput cell-free metabolic engineering data to inform the design of metabolic pathways in microorganisms that are challenging to engineer? and (2) how to utilize enzyme activity and promiscuity data to predict novel biochemical reactions that have utility in biosynthesis of chemicals and fuels?

To inform microorganism engineering from cell-free measurements, we are exploiting metabolic ensemble modeling (MEM) as a multiparameter sampling approach toward estimating enzymatic parameters. As the cell-free environment has several distinct differences from the cytoplasmic milieu, strong metabolic performance in a cell-free reaction may not translate to strong performance in the cell. By estimating enzyme parameters that are independent of environment, we can replace cell-free environmental parameters with appropriate *in vivo* parameters. We have developed a number of unique features in the MEM framework that will enable us to model cell-free systems, a first step toward our larger goal.

Toward the discovery of useful, novel biochemical reactions, we are using the Biochemical Network Integrated Computational Explorer (BNICE) to predict possible products from known metabolites and retrosynthetically predict precursors that lead toward target compounds. We are currently improving prediction accuracy by developing algorithms to better 'learn' from many available metabolic databases. The pathways and candidate enzymes predicted by BNICE can be rapidly tested in cell-free systems to identify lead enzymes for implementing in microbes. Altogether, these computational tools should enable faster and more sophisticated engineering of metabolic pathways to produce fuels and chemicals.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018249.

Climate adaptation and sustainability in switchgrass: exploring plant-microbe-soil interactions across continental scale environmental gradients

Thomas E. Juenger,[1,*] (tjuenger@austin.utexas.edu), Katherine Behrman,[1] Jeremy Schmutz,[2] Tanja Wolke,[3] Felix Fritschi,[4] Laura Bartley,[5] Julie Jastrow,[6] Roser Matamala,[6] Sean Watson,[7] Denise Costich,[8] Alina Zare,[9]

[1]University of Texas, Austin;[2] HudsonAlpha Institute for Biotechnology, Huntsville, Alabama; [3]DOE Joint Genome Institute, Walnut Creek, California; [4]University of Missouri, Columbia;[5] University of Oklahoma, Norman; [6]Argonne National Lab, Argonne, Illinois; [7]Ladybird Johnson Wildflower Center, Austin, Texas; [8]International Maize and Wheat Improvement Center; [9]University of Florida

https://sites.cns.utexas.edu/juenger_lab/switchgrass

Project Goals: Our collaborative project is focused on understanding switchgrass genetic diversity and adaptation across continental scale environmental gradients. Our goal is to improve the sustainability of switchgrass as a biofuel by gaining insight into the interaction of switchgrass diversity with its associated microbiome and environmental conditions. Our approach involves 1) the collection and characterization of new switchgrass germplasm from across the species range, 2) the development of a genetic association mapping panel and extensive common gardens to evaluate switchgrass performance, 3) a detailed characterization of the switchgrass microbiome, 4) studies of the impact of switchgrass stands on ecosystem processes, and 5) extensive multiscale modeling to define conditions of sustainability and identify key tradeoffs between genetic diversity, productivity, and ecosystem services. Ultimately, these studies will identify critical plant-microbe-soil traits that may be manipulated, through breeding or agronomic management, to improve the sustainability of biofuel feedstocks.

Less carbon-intensive energy sources are needed to reduce greenhouse gas emissions and their predicted role in climate change. There is growing interest in the potential of biofuels for meeting this need. A critical question is whether large-scale biofuel production can be sustainable over the time scales needed to mitigate our carbon debt from fossil fuel consumption. The carbon balance and ultimately the sustainability of biofuel feedstock production is the result of complex climate-coupled interactions between carbon fixation, sequestration, and release through combustion. The long-term productivity of biofuels depends on the genetic and environmental factors limiting plant growth. These factors are often related to soil resources which involve complex interactions at the plant-microbe-soil interface impacting their availability and cycling.

Our collaborative project addresses sustainable switchgrass (*Panicum virgatum*) production through a detailed characterization of plant growth and performance in both individual spaced and stand plantings. The project represents an unprecedented field-based experimental system for a bioenergy grass. We bring together diverse skill sets from plant and microbial genetics and genomics, physiology and ecosystem modeling. An underlying theme of the research is the use of locally adapted plant material to explore plant function, to understand the mechanistic basis of

Environmental interactions, and to discover the plant genes important for adaptation and sustainability in the face of climate change. To this end, we have been collecting new genetic diversity in switchgrass from natural populations across the species range in North America. These new genotypes are being characterized through genome resequencing and clonally propagated for inclusion in experimental gardens at 14 field locations. This material will provide a detailed population genomic characterization of switchgrass along with resources for association mapping and genomic selection for future breeding programs.

Our plant-microbiome project will fully characterize the microbial communities associated with switchgrass at our planting locations using genomic tools. Specifically, we have sampled field plantings of switchgrass for leaf and root microbial communities with 16S iTAG sequencing in collaboration with the DOE Joint Genome Center. Analyses will center on quantifying the relative importance of switchgrass host genotype, the planting environment, or their interaction on microbial community composition. Ultimately, these data will be linked with plant trait information to evaluate drivers of plant-microbiome interactions and their impacts on ecosystem processes.

Our ecosystem processes research focuses on carbon cycle responses at the ecosystem level using stand plantings of switchgrass diversity. We will couple plant growth and physiology measurements with measurements of key carbon pools and fluxes that integrate carbon dynamics across different temporal and spatial scales. Finally, our modeling will define conditions of a sustainable biofuel system and identify key tradeoffs between genetic diversity, productivity, and ecosystem services. Here, we highlight the background resources for our project, our progress establishing our new experiments, and plans for the coming field season.

This research is supported by funding through the Office of Biological & Environmental Research within the Department of Energy Office of Science and through the DOE Joint Genome Institute Community Science Program.

Microbial successions in switchgrass roots

Esther Singer[1*] (esinger@lbl.gov), Jason Bonnette[2], Tanja Woyke[1] and **Thomas E. Juenger**[2]

[1]DOE Joint Genome Institute, Walnut Creek, CA; [2]University of Texas Austin, Austin, TX;

https://sites.cns.utexas.edu/juenger_lab/switchgrass

Project Goals: Short statement of goals.

Our project aimed at establishing a high-throughput pipeline for characterizing diversity and community composition of the switchgrass microbiome. The pipeline we aimed to establish includes the optimization of 1) sampling techniques of various plant tissue types, 2) sample processing protocols as well as 3) the streamlined analysis of large 16S rRNA and ITS amplicon sequence datasets. The results of this project contribute essential information for the exploration of plant-microbe-soil interactions across continental scale environmental gradients.

Part of the DOE's strategy to ensure American energy independence is to produce biofuels from dedicated biomass crops. Achieving DOE's ambitious goal of displacing 30% of 2004 gasoline demand with biofuels by 2030 will require major increases in plant productivity. In an effort to providing sustainable solutions for enhanced plant growth in marginal lands, plant microbiomes play a vital role. Especially root colonizing bacteria have a large impact on plant traits, including biomass yield and overall plant health, which has been demonstrated for a number of plants. Switchgrass (*Panicum virgatum*) has been championed as a promising bioenergy species, however, few switchgrass microbiome studies have been conducted to date. We here present 3 complementary studies, in which 16S rRNA gene and ITS amplicon surveys were used to obtain insights into root microbial community dynamics across 3 different field sites and over 3 years. We show that plant genotype/ecotype, site and seasonal changes impact microbiomes differentially across different plant compartments[1]. Root endosphere showed most sensitivity towards plant genetic variance. Seasonality affected rhizosphere microbial communities more than site and showed a transition of dominant bacterial classes from *Gammaproteobacteria* and *Actinobacteria* in the spring to *Alphaproteobacteria* in phyllosphere and rhizosphere in the fall. We identified re-occurring microbial players *Streptomyces* and *Sphingobium* conserved across rhizosphere and root endosphere from various switchgrass genotypes and even across various *Panicum* species. These results of these studies have allowed unprecedented insight into switchgrass microbiomes and will direct future field microbiome studies QTL mapping and GWAS populations.

References

1. Singer, E., Bonnette, J., Kenaley, S., Woyke, T. & Juenger, T. E. Plant compartment and genetic variation drive microbiome composition in switchgrass roots. *Environmental Microbiology Reports* doi:10.1111/1758-2229.12727

This research is supported by funding through the Office of Biological & Environmental Research within the Department of Energy Office of Science and through the DOE Joint Genome Institute Community Science Program.

Direct Biomass Conversion through the One-Pot Ionic Liquid Technology

Ezinne C. Achinivu[1,3]* (eachinivu@lbl.gov), Gabriella Papa,[2] Eric Sundstrom,[2] Fabrice Masson,[2] Junko Yaegashi,[1] Jipeng Yan,[2] Ning Sun,[2] Blake Simmons[1,4] and **John Gladden**[1,3]

[1]Deconstruction Division, Joint Bio Energy Institute, Emeryville, California; [2]Advanced Biofuels Process Development Unit, Lawrence Berkeley National Laboratory, Berkeley, California; [3]Biomass Science and Conversion Technology, Sandia National Laboratories, Livermore, California; and [4]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California

https://www.jbei.org/

Project Goals: Pretreatment using ionic liquids (ILs) is one of the most effective methods for producing high yields of biomass-derived fermentable sugars. We are also exploring the utility of a related class of compounds, deep eutectic solvents, for their ability to pretreat lignocellulose. Our current research focuses on developing IL pretreatment technologies that enable the efficient depolymerization of both polysaccharides and lignin that can be tightly integrated with downstream processes, including saccharification and fermentation.

Successfully attaining a low-cost, high efficiency, lignocellulose deconstruction is a critical step towards the widespread adoption of lignocellulosic biofuels. Renewable ionic liquids (ILs) are a novel alternative solvent for biomass pretreatment and conversion. The use of bio-derived ILs such as choline lysinate [Ch][Lys] in a ***one-pot process*** reduces the need for solid-liquid separations and water washing of pretreated solids. This unconventional approach results in an increased techno-economic and environmentally benign process and will enable the direct commercialization of lignocellulosic biofuels. To assess the potential of the "one pot" ionic liquid-based configuration, different feedstocks including grasses and woody biomass (sorghum and poplar) were pretreated with the IL choline lysinate. Sorghum and poplar were pretreated at high solid loading of (15-30 wt%) using IL/water solution of [Ch][Lys] (7 wt%) at [120-160] °C for 1-3 h. After extensive process optimization, a glucose titer of ~50 g/L and 60 g/L was achieved from sorghum and poplar, respectively. These hydrolysates were then converted to advanced biofuels using the *Rhodosporidium toruloides* conversion host, demonstrating a complete feedstocks-to-fuels pipeline. Therefore, this approach has been demonstrated to be an efficient means to deconstruct and convert biomass to biofuels.[1,2,3] The fully consolidated process eliminates the requirement for IL separation prior to saccharification and fermentation thereby improving the potential for commercialization of lignocellulosic biofuels.

References

1. Sun *et al*., *Green Chem.* **2014**,*16*, 2546.
2. Sun *et al*., *Energy Environ. Sci*. **2016**, *9*, 2822-2834.
3. Sundstrom *et al.*, *Green Chem.* **2018**, *20*, 2870-2879.

This work was carried out as part of the DOE Joint BioEnergy Institute (http:// www.jbei.org) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy

Biomass Sorghum Supply Systems for Advanced Biofuel Production

Nawa Raj Baral[1*] (nrbaral@lbl.gov), Daniel Putnam,[2] Jeff Dahlberg,[3] and **Corinne D. Scown**[1]

[1]Life-cycle, Economics, & Agronomy Division, Joint BioEnergy Institute, Emeryville, CA; [2]University of California, Davis; and [3]Kearney Agricultural Research & Extension, Parlier, CA

Project Goals: Enhancing bioenergy sorghum productivity over a wide range of field environments while minimizing nutrient inputs, identifying the most cost-competitive and scalable biomass sorghum-to-fuel and bioproducts conversion technologies that will reduce GHG and water footprints, reduce the generation of pollutants and toxins, minimize demands and impacts on natural resources, and create more jobs.

Availability, cost, and quality (total carbohydrates) of biomass feedstock are some of the key drivers for the commercial success of the advanced biofuels production (such as aviation fuels) in the future. Biomass sorghum could be an ideal feedstock to address these concerns due to a high biomass production rate of 28 metric ton (dry)/ha, natural drought-tolerance (can grow across the United States), and a high total carbohydrates of ~70 wt% (photoperiod-sensitive or brown midrib (BMR) sorghum varieties). However, biomass sorghum feedstock supply system for cellulosic biorefinery have not been well studied and could be very challenging due to a high moisture content at harvest of 60-70 wt%. This study identifies six potential supply systems and quantifies corresponding supply costs considering two different production methods (rainfed and irrigated) and four different feedstock forms, including chopped (ensiled)-biomass, module, bale, and pellet. The results indicate that the chopped (ensiled)-biomass supply system with or without preprocessing depot between the field and the biorefinery is the most cost effective option when compared to the bale- or module-based biomass supply system. We find that if the biorefinery is located within the supply radius of 94.8 km (58.9 miles) the direct transportation of the chopped (ensiled)-biomass from the field to the biorefinery is more cost effective than a two-stage transportation system (a combination of the chopped-biomass and pellet system where pellets are produced at the preprocessing depot). This economic cut-off supply distances for the module and bale systems are 119.6 km (74.3 miles) and 107.9 km (67.1 miles), respectively. These promising feedstock supply routes require system-wide improvements including a biomass yield, sustainable farming with a low nutrient and water uses, and best management practices to reduce harvesting and handling hours, material losses, and supply distances. These future improvements enhance the availability of biomass sorghum feedstock for cellulosic biorefineries at a reasonable price.

Funding statement.

This work was part of the DOE Joint BioEnergy Institute (http:// www.jbei.org) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

Structural and Biochemical Investigation of Terminal Alkene Formation in Allylmalonyl-CoA Biosynthesis

Jacquelyn Blake-Hedges (jblakehedges@lbl.gov),[1,2,3] Jose Henrique Pereira,[2,3] Jeffrey Chen,[1] Rohith Krishna,[1] Leanne Chan,[2,3] Yan Chen,[2,3] Jennifer Gin,[2,3] Leonard Katz,[2] Christopher Petzold,[2,3] Paul Adams,[2,3] **Jay Keasling**[2,3,4]

[1]Department of Chemistry, University of California, Berkeley, Berkeley, CA; [2]Joint BioEnergy Institute, Emeryville, CA; [3]Biological and Systems Engineering, Lawrence Berkeley National Laboratory, Berkeley, CA; [4]Department of Chemical Engineering, University of California, Berkeley, Berkeley, CA

https://www.jbei.org/

Polyketide natural products are replete with diverse and complex chemical structures. This vast chemical diversity is generated through the condensation and subsequent reduction of small malonyl-CoA-derived extender units in an assembly line-like fashion. Of these extender units, allylmalonyl-CoA is of particular interest for incorporation into bioproducts because of the reactivity of the alkene handle it contains. This alkene handle is formed by TcsD, an acyl-CoA dehydrogenase (ACAD. ACADs are a large family of flavoenzymes that catalyze the oxidation of fully saturated acyl-CoA thioesters to their α,β-unsaturated counterparts. However, TcsD catalyzes the formation of a γ,δ-olefin on an α,β-unsaturated substrate, 2-pentenoyl-CoA. Herein, we interrogate the unusual regiochemistry of this transformation. We present biochemical data from experiments using natural substrates and substrate analogs to probe the mechanism of TcsD. In addition, we report high resolution (1.5 Angstroms) structural data from the first solved crystal structure of this enzyme. We use this combination of biochemical and structural data to propose a logical mechanism by which TcsD controls this novel ACAD activity. By understanding the activity of TcsD, we can better engineer strains to heterologously produce allylmalonyl-CoA for introduction into novel bioproducts such as polymer precursors.

This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Additional funding sources include an NSF GRFP DGE 1106400 to J. Blake-Hedges.

Title: Development Of An Automated "Cells-To-Peptides" Sample Preparation Workflow For High-Throughput Quantitative Proteomic Applications

Yan Chen[1]* (yanchen1998@lbl.gov), Joel M. Guenther[1], Leanne Jade G. Chan[1], Jennifer Gin[1], Paul D. Adams[1], **Christopher J. Petzold[1]**

[1] Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, CA, USA;

https://www.jbei.org/

Project Goals:

Recently, novel synthetic biology and metabolic engineering methods have enabled production of many valuable chemicals and advanced biofuel compounds. Mass spectrometry (MS)–based omics analysis has proven valuable for diagnosing the engineered organisms and aid metabolic pathway optimization. Innovative computational tools and recombinant DNA technology greatly speed up strain design and construction efforts, which places heavy demand on high throughput analytical platform. Several efforts have demonstrated that automated omics sample preparation could increase sample throughput and consistency. However, current effort mainly focused on automating procedures from proteins digestion to peptides cleaning up prior to LCMS analysis. Here, we developed a head-to-toe automated platform that prep are samples for omics analysis from biomaterial to LCMS analysis. Our goal of this complete automated platform is for high reproducible and less labor intensive omics sample preparation. This automation method could handle up to 384 samples in cell pellet form, and takes up to 5 hours prior to trypsin digestion. This method brings down the per sample cost to around $1.25. In addition, our automation method is the first-time report to prepare both proteomics and metabolomics samples from the same biomaterial. Our initial evaluation of the reproducibility of the automation platform were performed on quantitative proteomic analysis of a single 96 well plate samples. Our results showed that 88 % of 50 peptides from a total of 25 proteins have less than 20 % CV. The CV median was 10 %. We then extended our targets to 700 peptides out of 360 proteins, and our current result showed that the CV of 42.6% targets were below 20%, and the median CV of all targets was 21.8%. On the metabolomics side, we were able to consistently detect several common central carbon metabolites.

References:

1. Petzold, C.J. et al. Analytics for metabolic engineering. *Front. Bioeng. Biotechnol*. 3:135. doi:10.3389/fbioe.2015.00135.

2. Batth, T.S. et al. A targeted proteomics toolkit for high-throughput absolute quantification of Escherichia coli proteins. *Metab. Eng*. 2014, 26, 48-56.

This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Heterologous Production of Lignin Modifying Enzymes in *Aspergillus niger*

Comyn SA[1,2]* (scomyn@lbl.gov), Baker SE[1,3], Simmons BA[1,4], Gladden JM[1,5], **Magnuson JK**[1,2]

[1] Joint BioEnergy Institute, Deconstruction Division, Lawrence Berkeley National Laboratory, Emeryville, California
[2] Chemical and Biological Process Development Group, Pacific Northwest National Laboratory, Richland, Washington
[3] Environmental Molecular Sciences Division, Pacific Northwest National Laboratory, Richland, Washington
[4] Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California
[5] Biomass Science and Conversion Technologies Department, Sandia National Laboratories, Livermore, California

https://www.jbei.org/research/divisions/deconstruction/
https://www.jbei.org/research/divisions/deconstruction/fungal-biotechnology/

Project Goals: The fungal biotechnology group at JBEI is focused on developing fungal expression systems for industrial scale production of enzymes related to biofuel production. Specifically, we are expanding the capacity of heterologous expression hosts to produce complex metalloproteins involved in lignin depolymerization.

A desire to move away from fossil fuels toward more sustainable, renewable, or carbon-neutral sources has motivated the search for alternative feedstocks. Lignocellulosic biomass is an attractive option due to its abundance, low cost, and high sugar content. However, the high cost and limited availability of the enzyme cocktails required for efficient biomass conversion are significant barriers to the adoption of lignocellulosic biomass as a feedstock to produce biofuels and other value-added products. Filamentous fungi have an astounding capacity to secrete digestive enzymes into their extracellular environment through the presence of efficient systems for the transcription, translation, folding, and secretion of polypeptides. These features have led filamentous fungi to be exploited for the manufacturing of homologous and heterologous proteins for the chemical, pharmaceutical, and biofuel industries. Efficient heterologous protein production often requires multilevel optimization strategies for both the protein and the expression host as factors limiting production are often protein specific and bottlenecks can exist at any stage from transcription and translation to folding and secretion. We are developing fungal host systems for the expression and characterization of lignin modifying enzymes (LME) in the filamentous fungus *Aspergillus niger*. More specifically, we are focusing on the post-translational bottlenecks that are known to limit, in terms of quantity and quality, heterologous protein production. We will present the engineering strategies used to build a fungal expression system for LME, and the plan for genetically engineering key components of the secretory pathway with the aim of enhancing heterologous protein production. The development of efficient expression systems and hosts to identify, characterize, and enhance the expression levels, yields, and activities of heterologous LME is critical to the successful adoption of lignocellulosic biomass into the new bio-economy.

This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Lignin Manipulation in Feedstocks Towards Lignin Valorization and Bioproducts

Lin Chien-Yuan[1,2], Khahn Vuu[1,2], Patrick Shih[1,2], Henrik V. Scheller[1,2], Aymerick Eudes[1,2]*(ageudes@lbl.gov), and **Jay D. Keasling**[1,3]

[1]Joint BioEnergy Institute, Emeryville, California
[2]Lawrence Berkeley National Laboratory, Environmental Genomics and Systems Biology Division, Berkeley, California
[3]Lawrence Berkeley National Laboratory, Biological Systems and Engineering Division, Berkeley, California

https://www.jbei.org/

Project Goals: The most abundant organic material on earth is lignocellulosic biomass or non-food plant material. JBEI's mission is to convert biomass to biofuels and bioproducts. The goal is to provide the nation with clean, renewable transportation fuels and chemicals. Building a successful lignocellulosic biofuels industry depends, in part, on developing specialized bioenergy crops or feedstocks that are optimized for deconstruction and conversion.

Lignin, along with cellulose and hemicellulose, comprises the majority of plant cell wall. To improve the quality of biomass for conversion into biofuels or bioproducts, genetic manipulation mostly focused on lignin engineering. At JBEI, we propose novel approaches for lignin valorization by manipulating lignin to reduce biomass recalcitrance and produce bioproducts without growth penalties. To achieve our goal, we apply JBEI's synthetic biology tools to introduce new genes/pathways utilizing the shikimate and phenylpropanoid pathways. Four strategies were chosen for tailoring feedstocks including: 1) Reduction of the biomass recalcitrance and increasing the yield of biomass; 2) Reduction of the degree of polymerization (DP) of lignin; 3) Manipulation of lignification by overproducing a native monolignol precursors, *p*-coumarate (p-CA), and incorporate it into lignin polymer; 4) Lignin Valorization for accumulation of bioproducts such as 4-hydroxybenzoic acid (HBA), protocatechuic acid (PCA), and muconic acid (MA). All the possible pathways for reducing biomass recalcitrance, manipulating lignification and producing bioproducts are introduced to model and bioenergy feedstocks, including *Arabidopsis*, *Brachypodium*, sorghum, tobacco and poplar.

Funding statement.
This work was part of the DOE Joint BioEnergy Institute (http:// www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.

An Automated Microfluidic System for On-Chip Genetic Engineering Processes

Kosuke Iwai,[1,2] William R. Gaillard[1,2]* (rgaillard@lbl.gov), David Ando,[1,3] Peter W. Kim,[1,2] Kai Deng,[1,2] Trent R. Northen,[1,4,6] Hector Garcia Martin,[1,3,4] Nathan J. Hillson,[1,3,6] Paul D. Adams,[1,4,5,7] and **Anup K. Singh**[1,2]

[1]Joint BioEnergy Institute, Emeryville, CA; [2]Biological and Material Science, Sandia National Laboratories, Livermore, CA; [3]Biological Systems and Engineering, Lawrence Berkeley National Laboratory, Berkeley, CA; [4]Environmental Genomics and Systems Biology, Lawrence Berkeley National Laboratory, Berkeley, CA; [5]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA; [6]DOE Joint Genome Institute, Walnut Creek, CA; and [7]University of California, Berkeley, CA

https://www.jbei.org/research/divisions/technology/microfluidics/

Project Goals: The JBEI mission is to conduct basic and applied research to enable economically-viable conversion of lignocellulosic biomass into biofuels to provide the nation with clean, renewable transportation fuels identical to gasoline, diesel, and jet fuel. The goal of this project, performed in the Microfluidic Assays group in the Technology Division at JBEI, is to deliver robust and easy-to-use microfluidic platforms to automate the genetic engineering processes for advancing synthetic biology applications including biofuels development.

In recent years, synthetic biology has drawn significant interest for both scientific research and industrial applications, such as biofuel and pharmaceutical production. Synthetic biology processes require multiple molecular biology steps, making it a very time-consuming and labor-intensive effort. Therefore, automated and efficient processes to perform molecular biology assays are highly desired. Droplet based microfluidic technologies offer powerful approaches to improve synthetic biology processes due to faster reactions fostered by minute dimensions, reduced reagent consumption resulting from smaller working volumes, and increased control over experimental conditions.

We are involved in developing innovative microfluidic assays and integrated devices for many biofuel research applications including enzyme screening, enzyme evolution, and synthetic biology. Our hybrid microfluidic platforms utilize continuous-flow (analog) microfluidics that manipulate droplets by controlling the hydrodynamic force, and digital microfluidics (DMF) that utilize surface tension from electrowetting on dielectric with arrayed electrodes. The systems can handle large numbers of droplets at once and actively manipulate target droplets in a programmable manner. Hybrid and DMF devices are capable of multiple droplet manipulation steps including formation of aqueous droplets, encapsulation of reagents and cells, hydrodynamic capture, arraying of droplets, electric-field driven merging and splitting of droplets to achieve specific

volumes and concentrations of various reagents, on-chip electroporation, and incubation with localized temperature control.

Specifically, for electroporation devices, multiple pairs of electrodes are designed and placed at each chamber to apply voltages to arrayed droplets for on-chip electroporation. This configuration allows us to customize the electroporation conditions at each droplet for multiplexed DNA transformation processes, and it also enables us to easily scale-up the numbers of reactions for high-throughput transformation processes simply by scaling the array size. In addition, we integrate optical fibers in the microchannels to add on-chip fluorescence-based detection of encapsulated cells and enzymatic activities in the discrete droplets and for triggering sorting of droplets. We utilize our microfluidic methodologies for automating CRISPR/Cas9 based gene editing processes such as recently the established CRMAGE for *Escherichia coli* or the cloning-free tool kit for *Saccharomyces cerevisiae*.

Unlike conventional microtiter plate-based reactions, our analog-digital microfluidic platforms with on-chip electroporation and fluorescence detection allow completely automated genetic engineering steps using 10-100-fold lower amounts of reagents and can be useful for applications requiring high throughput screening and reactions.

References

1. P.C. Gach, K. Iwai, P.W. Kim, N.J. Hillson, and A.K. Singh, "Droplet Microfluidics for Synthetic Biology," *Lab Chip*, 2017, **17**, 3388-3400.
2. P.C. Gach, S.C.C. Shih, J. Sustarich, J.D. Keasling, N.J. Hillson, P.D. Adams and A.K. Singh, "A Droplet Microfluidic Platform for Automating Genetic Engineering," *ACS Synth. Biol.*, 2016, **5**, 426-433.
3. S.C.C. Sigh, G. Goyal, P.W. Kim, N. Koutsoubelis, J.D. Keasling, P.D. Adams, N.J. Hillson and A.K. Singh, "A Versatile Microfluidic Device for Automating Synthetic Biology," *ACS Synth. Biol.*, 2015, **4**, 1151-1164.
4. S.C.C. Shih, P.C. Gach, J. Sustarich, B.A. Simmons, P.D. Adams, S. Singh and A.K. Singh, "A Droplet-to-Digital (D2D) Microfluidic Device for Single Cell Assays," *Lab Chip*, 2015, **15**, 225-236.
5. P.C. Gach, S.C.C. Shih, J. Sustarich, J.D. Keasling, N.J. Hillson, P.D. Adams and A.K. Singh, "A Droplet Microfluidic Platform for Automating Genetic Engineering," *ACS Synth. Biol.*, 2016, **5**, 426-433.
6. C. Ronda, L. E. Pedersen, M. O. Sommer, and A. T. Nielsen, "CRMAGE: CRISPR Optimized MAGE Recombineering", *Sci. Rep.*, 2016, **6**, 19452.
7. A. Reider Apel, L. d'Espaux, M. Wehrs, D. Sachs, R. A. Li, G. J. Tong, M. Garber, O. Nnadi, W. Zhuang, N. J. Hillson, J. D. Keasling, and A. Mukhopadhyay, "A Cas9-Based Toolkit to Program Gene Expression in *Saccharomyces Cerevisiae*", *Nucleic Acids Res.*, 2017, **45**, 496-508.

This work was part of the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.

Quantitative metabolic modeling at the Joint BioEnergy Institute

Hector Garcia Martin[1,2,3,*](hgmartin@lbl.gov), Zak Costello[1,2,3], Tyler Backman[1,2,5], William Morrell[3,4], Mark Forrer[3,4], Nathan Hillson[1,2,3], Kosuke Iwaii[3,4], Anup Singh[3,4], Edward Baidoo[1,2], Paul Adams[1,2,5], **Jay Keasling[1,2,5]**

[1]Joint BioEnergy Institute, Emeryville, CA; [2]Lawrence Berkeley National Laboratory, Berkeley, CA; [3]DOE Agile BioFoundry; [4]Biological and Material Science, Sandia National Laboratories, Livermore, CA, [5]University of California at Berkeley, Berkeley, CA;

Project Goals: The Quantitative Metabolic Modeling (QMM) group is devoted to developing predictive models of metabolism that can leverage high-throughput 'omics data and systematically direct metabolic engineering efforts.

Our inability to predict the behavior of biological systems severely hampers progress in bioengineering and biomedical applications. We cannot predict the effect of genotype changes on phenotype, nor extrapolate the large-scale behavior from small-scale experiments. Machine learning techniques recently reached a new level of maturity, and are capable of providing the needed predictive power without a detailed mechanistic understanding. However, they require large amounts of data to be trained. The amount and quality of data required can only be produced through a combination of synthetic biology and automation, so as to generate a large diversity of biological systems with high reproducibility.

Here we show the variety of methods created in the QMM group to leverage 'omics data and guide metabolic engineering. We have used **machine learning** approaches to predict pathway dynamics directly from time-series data [1], **microfluidics chips** for automated synthetic biology [2] and created **online tools for –omics data visualization.**

References

1. Costello Z, Martin HG (2018) A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data. npj Syst Biol Appl 4: 19. doi:10.1038/s41540-018-0054-3.

2. Iwai K, Ando D, Kim PW, Gach PC, Raje M, et al. (2018) Automated flow-based/digital microfluidic platform integrated with onsite electroporation process for multiplex genetic engineering applications. 2018 IEEE Micro Electro Mechanical Systems (MEMS). IEEE. pp. 1229–1232. doi:10.1109/MEMSYS.2018.8346785.

This work was part of the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.

Collaboration with the Experiment Data Depot

William Morrell[2,3,4], Mark Forrer[2,3,4], Garrett Birkel[1,3,5], Traci Lopez[2,3,4], Hector Garcia Martin[1,3,5]*, and **Nathan J. Hillson**[2,3,5]* (njhillson@lbl.gov)

[1]Biofuels and Bioproducts and [2]Technology Divisions, DOE Joint BioEnergy Institute; Emeryville, CA 94608, USA; [3]DOE Agile BioFoundry; Emeryville, CA 94608, USA; [4]Sandia National Laboratory; Livermore, CA 94550, USA; [5]Lawrence Berkeley National Laboratory; Berkeley, CA 94720, USA.

https://public-edd.jbei.org

Project Goals:

Although recent advances in synthetic biology allow us to produce biological designs more efficiently than ever, our ability to predict the end result of these designs is still nascent. Predictive models require a large corpus of high-quality data to be usefully parametrized and tested. Suitable datasets for these models are often not generally available. Here, we present the Experiment Data Depot (EDD), an online platform designed to act as a repository of experimental data and metadata. EDD provides a convenient way to upload a variety of data types, visualize these data, and export them in a standardized fashion for use with predictive algorithms. In this poster, we describe EDD and showcase its utility for three different use cases: the characterization of promoters for synthetic biology parts, leveraging proteomics data to improve biofuel yield, and the use of extracellular metabolite concentrations to predict intracellular metabolic fluxes.

References

1. Morrell W, Birkel G, Forrer M, Lopez T, Backman T, Dussault M, Petzold C, Baidoo E, Costello Z, Ando D, Alonso Gutierrez J, George K, Mukhopadhyay A, Vaino I, Keasling J, Adams P, Hillson NJ*, Garcia Martin H*. (2017) "The Experiment Data Depot: a web-based software tool for biological experimental data storage, sharing, and visualization" *ACS Synthetic Biology* DOI: 10.1021/acssynbio.7b00204

This work was part of the DOE Joint BioEnergy Institute (http:// www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, and was part of the Agile BioFoundry (http://agilebiofoundry.org) supported by the U.S. Department of Energy, Energy Efficiency and Renewable Energy, Bioenergy Technologies Office, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.

Developing a Targeted Mass-Spectrometry Platform for High-Throughput Characterization of Lignocellulosic Biomass Deconstruction

Nicole Ing[1,2]* (ingn@lbl.gov), Kai Deng[1,2], Markus De Raad[3], Jijiao Zeng[1,2], Blake Simmons[1,3], Paul Adams[1,3], Anup Singh[1,2], Jay Keasling[1,3], and **Trent Northen[1,3]**

[1]Joint BioEnergy Institute, Technology Division, Emeryville, CA 94608, USA; [2] Biosystems Research, Sandia National Laboratories, Livermore, CA 94551; [3] Biological Systems and Engineering, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

https://www.jbei.org/

http://www.northenlab.org/research/mass-spectrometry-innovations

Project Goal: This work aims to develop an analytical platform for high-throughput analysis of lignocellulosic biomass deconstruction. This platform combines surface-based mass spectrometry with bioconjugation chemistry to screen enzymatic and microbial activity toward lignin, cellulose, and hemicellulose degradation.

Lignocellulosic biomass is the most abundant raw material on the planet and is composed primarily of the cellulose, hemicellulose and lignin. Understanding how to deconstruct this complex material is critical to achieving a cost-competitive replacement for petrochemicals. Enzymes and microbes are capable of degrading this biomass in nature; however, the species, mechanisms, and pathways involved in deconstructing this material are vast and complex. To this end, we are developing mass spectrometry-based assays as a high-throughput platform to support the rapid development of minimal cocktails for lignocellulosic deconstruction. We present complementary developments in our mass spectrometry surfaces and the chemical probes used for targeted analysis of biomass degradation.

We present a new surface assisted laser desorption ionization method, Insulator Nanostructure Desorption Ionization Mass Spectrometry (INDI-MS). INDI-MS utilizes a self-assembling perfluoroalkyl silsesquioxane coating to achieve femtomolar sensitivity comparable to the hydrofluoric-etched Nanostructure Initiator Mass Spectrometry (NIMS) chips previously developed by our group. In addition to removing the etching step, this new method can also be integrated with photolithography to achieve sample self-desalting and enhanced performance.

To enhance our nanostructure-based MS detection of lignocellulosic biomass degradation products, our assay approach utilizes two chemical probe techniques to target activity toward the separate polysaccharide and aromatic components. A previously-developed perfluorinated tagging technique targets glycan products for rapid and reliable characterization of cellulose and hemicellulose degradation on our nanostructured surfaces. This approach has since been extended to characterize ligninase activity towards for β-O-4 type compounds. We show that unlike colorimetric assays used to study lignin decomposition, this lignin assay can provide unique and important information about specific bond cleavage reactions.

References

(1) Duncombe, T. A.; Raad, M. D.; Bowen, B. P.; Singh, A. K.; Northen, T. R. Insulator Nanostructure Desorption Ionization Mass Spectrometry. *Anal. Chem.* **2018**, *90* (16), 9657–9661.

(2) Deng, K.; Zeng, J.; Cheng, G.; Gao, J.; Sale, K. L.; Simmons, B. A.; Singh, A. K.; Adams, P. D.; Northen, T. R. Rapid Characterization of the Activities of Lignin-Modifying Enzymes Based on Nanostructure-Initiator Mass Spectrometry (NIMS). *Biotechnol. Biofuels* **2018**, *11* (1), 266.

(3) Deng, K.; Guenther, J. M.; Gao, J.; Bowen, B. P.; Tran, H.; Reyes-Ortiz, V.; Cheng, X.; Sathitsuksanoh, N.; Heins, R.; Takasuka, T. E.; et al. Development of a High Throughput Platform for Screening Glycoside Hydrolases Based on Oxime-NIMS. *Front. Bioeng. Biotechnol.* **2015**, *3*.

(4) Deng, K.; Takasuka, T. E.; Heins, R.; Cheng, X.; Bergeman, L. F.; Shi, J.; Aschenbrener, R.; Deutsch, S.; Singh, S.; Sale, K. L.; et al. Rapid Kinetic Characterization of Glycosyl Hydrolases Based on Oxime Derivatization and Nanostructure-Initiator Mass Spectrometry (NIMS). *ACS Chem. Biol.* **2014**, *9* (7), 1470–1479.

(5) Northen, T. R.; Lee, J.-C.; Hoang, L.; Raymond, J.; Hwang, D.-R.; Yannone, S. M.; Wong, C.-H.; Siuzdak, G. A Nanostructure-Initiator Mass Spectrometry-Based Enzyme Activity Assay. *Proc. Natl. Acad. Sci.* **2008**, *105* (10), 3678–3683.

(6) Northen, T. R.; Yanes, O.; Northen, M. T.; Marrinucci, D.; Uritboonthai, W.; Apon, J.; Golledge, S. L.; Nordström, A.; Siuzdak, G. Clathrate Nanostructures for Mass Spectrometry. *Nature* **2007**, *449* (7165), 1033–1036.

This work conducted by the Joint BioEnergy Institute was funded by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Isopentenol Production Using the IPP-Bypass Mevalonate Pathway

Daniel Mendez-Perez[1,2], Xi Wang[1,2], Aram Kang[1,2], Aindrila Mukhopadhyay[1,2], and

Taek Soon Lee[1,2,*] (tslee@lbl.gov)

[1]Joint BioEnergy Institute, 5885 Hollis Street, Emeryville, CA 94608, USA.
[2]Biological Systems & Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

http://www.jbei.org/research/divisions/biofuels-and-bioproducts/pathway-metabolic-engineering/

Project Goals: Isopentenol is a drop-in biofuel and a precursor for commodity chemicals, and biosynthetic pathways to this product have been developed. But most production study has been limited to batch experiments in shake flasks. In this study, we explored isopentenol production via IPP-bypass pathway in a 2-L bioreactor using fed-batch fermentation as an initial scale-up demonstration of the technology. With a success of the initial scale up effort, we may attempt the production of isopentenol at a commercially more relevant scale in the future.

Isopentenol (3-methyl-3-buten-1-ol, isoprenol) is a valuable compound as a drop-in biofuel and a precursor of commodity chemical such as isoprene. Synthetic microbial system using heterologous mevalonate pathway has been developed for the production of isopentenol via extensive metabolic engineering efforts[1]. However, the toxicity of key intermediate isopentenyl diphosphate (IPP) has been a significant bottleneck of high titer production of isopentenol in microbial hosts[2]. In previous study, we have developed an IPP-bypass pathway for isopentenol production which bypasses the toxic intermediate formation as well as relieves energy demand of the original mevalonate pathway[3,4]. This alternative pathway has been developed and further engineered for the production of isopentenol in *E. coli*. However, production studies have been limited to batch experiments in shake flasks. In this study we tested isopentenol production via IPP-bypass pathway in a 2-L bioreactor using fed-batch fermentation. After several optimization strategies such as media optimization and the elimination of the acetate generating pathways, isopentenol production reached 10.8 g/L, which is the highest reported titer for this compound. To achieve this high titer, we found that it is required to use a two-phase fermentation process in which isopentenol is partially removed from the culture media to organic phase (with oleyl alcohol) to relieve the toxic effects by isopentenol. This IPP-bypass pathway was also engineered in *P. putida* for initial production of isoprenoids.

References

1. George, et al., *Biotech. Bioeng.* **2014**, *111*, 1648-1658 (doi:10.1002/bit.25226)

2. George, et al., *Metabol. Eng*. **2018**, *47*, 60-72 (doi: 10.1016/j.ymben.2018.03.004)
3. Kang, et al., *Metabol. Eng*. **2016**, *34*, 25-35 (doi: 10.1016/j.ymben.2015.12.002)
4. Kang, et al., *Metabol. Eng*. **2017**, *41*, 125-134 (doi: 10.1016/j.ymben.2017.03.010)

Funding statement

This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Machine Learning for Bioenergy Sorghum Yield Prediction under Future Climate Scenarios

Tyler Huntington[1], Umakant Mishra[1,2] (umishra@anl.gov), Xinguang Cui[1], Corinne Scown[1]

[1]Life-cycle, Economics, & Agronomy Division, Joint BioEnergy Institute, Emeryville, CA 94608.

[2]Environmental Science Division, Argonne National Laboratory, Lemont, IL, 60439.

Crop yield modeling has the potential to play an important role in the design of national strategies for agricultural production, particularly under a changing climate. The application of predictive techniques in agriculture has historically focused on food crops; however, in the emerging bioeconomy, forecasting yields of bioenergy crops at fine spatial resolutions has grown increasingly useful to the evaluation of near and long term pathways to scaling up biofuel production. Understanding how bioenergy crops targeted for cellulosic fuel production such as sorghum will perform under future climate scenario is especially crucial as the future emission scenarios indicate higher severity droughts and temperature extremes in the continental USA. In order to inform the genetic engineering of bioenergy crops to withstand these challenges, robust modeling techniques are needed to provide spatially-explicit predictions and insights on how crop performance will respond to changing environmental conditions. We present a comparative analysis of several modeling approaches for predicting yields of biomass sorghum at the county level in the US based on historical data. Validation of candidate models revealed that Extra Randomized Trees (ERT) regression, a variant of the Random Forest machine learning algorithm, outperformed a myriad of linear models which included stepwise, least absolute shrinkage and selection operator, Ridge and Geographically Weighted regression. Incorporating future climate projections from an ensemble of Global Circulation Models, we demonstrate the use of our ERT model to predict sorghum yields in 2020, 2050, 2080, and 2099 across four Representative Concentration Pathway (RCP) scenarios of the Intergovernmental Panel on Climate Change. Our predictions indicate a slight decreasing trend of sorghum yields over these future years, with greater decline in more severe RCP scenarios.

Applying Multi-dimensional Solid-state NMR to Explore the Nanoarchitecture of Native and Engineered Plant Cell Walls

Yu Gao[1,2], Andrew S. Lipton[3], Nancy Washton[3], Henrik V. Scheller[1,2], and **Jenny C. Mortimer[1,2]*** (jcmortimer@lbl.gov)

[1]Environmental and Systems Biology Division, Biosciences Area, Lawrence Berkeley National Laboratory, Berkeley, CA; [2]Feedstocks Division, Joint BioEnergy Institute, Emeryville, CA; [3]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA

http://jbei.org

Project Goals:

The secondary cell wall forms the majority of plant biomass, and therefore is the primary feedstock for generating sustainable biofuels and bio-based products. In the US, dedicated bioenergy crops will include monocots (e.g. sorghum, and switchgrass), whose cell walls are different from eudicots (e.g. Arabidopsis and poplar). Here, we aim to understand the monocot secondary cell wall nanoarchitecture, which will be used to predictively engineer future bioenergy crops and improve engineering models.

Most techniques are limited to characterizing extracted or digested cell walls. NMR allows the study of native cell walls in situ, as previously demonstrated for Arabidopsis[1]. Here, we developed a system for the high percentage incorporation of ^{13}C from $^{13}CO_2$ into plant material including larger plants such as sorghum. This is necessary to perform multi-dimensional NMR, which is required for the analysis of glycans. We have used this to produce both wild-type and engineered material which is being analyzed at EMSL. We will present some initial data from these experiments.

References

1. Simmons TJ*, Mortimer JC* et al. (2016) Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat Commun.*.

Funding statement.

This work was carried out as part of the DOE Joint BioEnergy Institute (http:// www.jbei.org) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy. The high-field NMR was performed at EMSL (grid.436923.9), a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research.

Importance of metabolic state for microbial bioproduction using nonribosomal peptide synthetases (NRPS)

Maren Wehrs[1,2], JP Prahl[1,3], Jadie Moon[1,2], Lukas Platz[1,2], Yuchen Li[1,2], Deepti Tanjore[1,3], Todd Pray[1,3], Blake A. Simmons[1,2], Jay D. Keasling[1,2], John M. Gladden[2,4], Aindrila Mukhopadhyay (amukhopadhyay@lbl.gov)[1,2]*

[1]Biofuels and Bioproducts Division, Joint BioEnergy Institute, Emeryville, CA 94608, [2]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, [3]Advanced Biofuels and Bioproducts Demonstration Unit, Emeryville, CA 94608, [4]Deconstruction Division, Joint BioEnergy Institute, Emeryville, CA 94608, [5]Biological and Engineering Sciences Center, Sandia National Laboratories, Livermore, CA 94551

Goal: For facultative aerobes, oxygen depletion controls the switch from respiratory to fermentative metabolism. Crabtree-positive yeasts are exceptions to this principle as they preferably use fermentation even in the presence of oxygen. While the Crabtree-effect makes model yeasts such as *S. cerevisiae* an ideal host for high production of ethanol, this metabolic quirk can be confounding for other bioproducts. Here, we show that the production of the blue pigment indigoidine (a nonribosomal peptide, NRP) can serve as an indicator for respiratory growth. We observed strong correlation of the production efficiency of this NRP and the metabolic state of the yeast cell, highlighting the importance of understanding the metabolic characteristics of a given production host. NRPS are an important class of biocatalysts that provide access to a wide array of secondary metabolites. The goal of this project is to develop NRPS expression and use it for microbial production of final targets from both pure sugars as well as plant biomass hydrolysates as part of the Joint BioEnergy Institute (URL: www.jbei.org).

Field testing of a low lignin engineering strategy in switchgrass

Kavitha Satish Kumar[1,2*] (kavithask@lbl.gov), Ramana Pidatala[1,2], Mi Yeon Lee[1,2], Brenda Perez[3], Christopher De Ben[3], Jasmine Ortega[1,2], Yi-Chun Chen[1,2], Daniel Putnam[2,3], Corinne Scown[1,2], Christa Pennachio[1,4], Katherine Louie[1,4], Benjamin P Bowen[1,4] Trent Northen[1,4], Henrik V Scheller[1,2], Aymerick Eudes[1,2], **Jenny Mortimer**[1,2]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA; [2]Joint BioEnergy Institute, Emeryville, California, USA; [3]Department of Plant Science, University of California, Davis, California, USA; and [4]US Department of Energy Joint Genome Institute, Walnut Creek, California, USA

https://www.jbei.org/

Project Goals: A multi-species data set will be generated to understand the changes in the molecular mechanism observed in plants as a result of the low lignin engineering strategy using transcriptomics and untargeted metabolomics.

Switchgrass *(Panicum virgatum* L.) is a promising perennial dedicated bioenergy feedstock. It can grow on marginal lands and produce abundant biomass. Modification of lignin composition for improved deconstruction is an important strategy for biomass improvement. Previously, the JBEI Feedstocks team developed a dominant strategy to reduce lignin content, while increasing the proportion of H-lignin, which relied on the expression of a bacterial enzyme - 3-dehydroshikimate dehydratase (*QsuB*) in Arabidopsis[1]. This strategy was first successfully tested in switchgrass in the greenhouse and is now being tested in a small-scale multi-year field trial. Previously, we had demonstrated that this strategy resulted in increased saccharification efficiency and low lignin content, with no change to growth in Arabidopsis or greenhouse-grown switchgrass. Field data collected from three-cuttings showed an increase in plant height and biomasss from the QsuB lines as compared to untransformed controls (Alamo). We are now using RNAseq and metabolomics to explore the reasons for these changes in collaboration with JGI. We are also testing the QsuB strategy in other key bioenergy feedstocks, including sorghum and poplar. Future work will include expanding the field trial to include multiple field sites, as well as a water-deficit trials.

References

1. Eudes A, et al. (2015) Plant Biotech. J. 13:1241.

The work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.
The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Drought and Bioenergy Grasses: Effects of Drought on Biomass Composition and Effect of Biomass Composition on Drought Tolerance

Tess Scavuzzo-Duggan[1,2], Devon Birdseye[1], Jingwei Yan[1,3], Siwen Deng[2,4], Ling Xu[2,4], Devin Coleman-Derr[2,4], Jeffrey Dahlberg[5], Aymerick Eudes[1], Corinne Scown[5], and **Henrik V. Scheller**[1,2*] (hscheller@lbl.gov).

[1]Feedstocks Division, Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Berkeley, CA; [2]Department of Plant & Microbial Biology, University of California, Berkeley; [3]Nanjing Agricultural University, Nanjing, China; [4]USDA-ARS, Albany, CA; [5]LEAD Division, Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Berkeley, CA

www.jbei.org

Project Goals: To determine the impact of drought on bioenergy crops and biofuel production.

Lignocellulosic biomass can be used to produce biofuels, bioproducts or biopower and presents a renewable energy resource that has the capacity to lower the global carbon footprint. However, shifting climates leading to an increase in arid lands and limits on water available for irrigation make it imperative to understand how the water deficit these bioenergy crops face may impact their biomass and bioenergy characteristics. In collaboration with the EpiCon project, we are reporting the results from the first trial comparing well-watered RT430 sorghum lines with RT430 lines that have been subjected to drought. Our results indicate that while there are differences in biomass characteristics of sorghum vegetative tissue in response to post-flowering drought stress, these changes are relatively small in comparison to total sugar composition. Notable changes appear in pectic monosaccharides (rhamnose, arabinose, galactose, galacturonic acid) of newly expanding tissues, in addition to hemicellulosic changes in older tissues (glucose, mannose, glucuronic acid). Consistent with the relatively small monosaccharide changes, the saccharification efficiencies of biomass differed only little between plants with and without post-flowering drought. In general, the results suggesting that post-flowering drought stress has little impact on biomass characteristics.

Another aspect of biomass and drought is the situation where biomass has been deliberately changed by breeding or engineering. How would such changes influence drought tolerance and water use efficiency? We have engineered Arabidopsis plants with low lignin content and/or low xylan content in fiber cells. Vessels were not impacted in these plants in order to ensure good growth properties, and all the engineered plants are indistinguishable from control plants during optimal growth conditions. Upon exposing these engineered Arabidopsis plants to severe drought, we observed better survival rates in those with low lignin and/or low xylan content compared to those in wild-type plants (1). In comparison, increased pectic galactan content had no effect on drought tolerance. The plants with low lignin were obtained by expressing QsuB, a bacterial dehydroshikimate dehydratase. The same gene has been expressed in switchgrass. Data will be reported on the drought response of the low lignin switchgrass plants. The mechanism for the

increased drought tolerance in the engineered plants is still under investigation. Nevertheless, the drought tolerance is an important finding because it demonstrates that modification of secondary cell walls does not necessarily render the plants less robust in the environment, and it shows that substantial changes in biomass composition can be achieved without compromising plant resilience.

References

1. Yan J, Aznar A, Chalvin C, Birdseye DS, Baidoo EEK, Eudes A, Shih PM, Loque D, Zhang A, Scheller HV: **Increased drought tolerance in plants engineered for low lignin and low xylan content**. *Biotechnol Biofuels* 2018, **11**:195.

Funding

This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. J.Y. was supported by a fellowship from China Scholarship Council.

Development and demonstration of CRISPR/Cas9 platform for *Aspergillus niger*

Laure M.C. Leynaud-Kieffer,[1,2,3] Blake A. Simmons[2,3]* (basimmons@lbl.gov), Samuel C. Curran,[2,3,4] Irene Kim,[5] Jon K. Magnuson,[2,6] John M. Gladden,[2,7] Scott E. Baker,[2,8] and **Jay D. Keasling**[2,3]

[1]Swiss Federal Institute of Technology Lausanne, Lausanne, Vaud, Switzerland, [2]Joint BioEnergy Institute, Emeryville, CA, [3]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, [4]Comparative Biochemistry Graduate Group, University of California Berkeley, Berkeley, CA, [5]Department of Chemistry, University of California, Berkeley, CA, United States of America, [6]Chemical and Biological Process Development Group, Pacific Northwest National Laboratory, Richland, WA, [7]Department of Biomass Science and Conversion Technology, Sandia National Laboratories, Livermore, CA, [8]Biosystems Design and Simulation Group, Environmental Molecular Sciences Division, Pacific Northwest National Laboratory, Richland, WA

www.jbei.org

Project Goals: JBEI's long-term vision is that bioenergy crops can be converted into economically-viable, carbon-neutral, specialty biofuels, all of the organic chemicals currently derived from petroleum, and many other bioproducts that cannot be efficiently produced from petroleum. This vision will only be possible when we have sustainable bioenergy crops, biorefinery technologies capable of converting as much carbon in biomass into biofuels as possible, and a vast array of bioproducts that will make biorefineries economically viable. JBEI's mission is to establish the scientific knowledge and new technologies in feedstock development, deconstruction and separation, and conversion needed to transform the maximum amount of carbon available in bioenergy crops into biofuels and bioproducts. When fully scaled, JBEI's technologies will enable the production of replacements for petroleum derived gasoline, diesel, jet fuel, and bioproducts. In doing so, JBEI will reduce the nation's dependence on fossil fuels, significantly reduce the amount of carbon added to the atmosphere, reduce contamination of the environment, and provide the scientific tools and knowledge required to transform the bioenergy marketplace.

Aspergillus niger and other filamentous fungi are widely used in industry, but efficient genetic engineering of these hosts remains nascent. For example, while molecular genetic tools have been developed, including CRISPR/Cas9, facile genome engineering of *A. niger* remains challenging. To address these challenges, we have developed a simple Cas9-based gene targeting method that provides selectable, iterative, and ultimately marker-free generation of genomic deletions and insertions.[1] This method leverages locus-specific "pop-out" recombination to suppress off-target integrations. We demonstrated the effectiveness of this method by targeting the phenotypic marker *albA* and validated it by targeting the *glaA* and *mstC*loci. After two

selection steps, we observed 100% gene editing efficiency across all three loci. This method greatly reduces the effort required to engineer the *A. niger* genome and overcomes low Cas9 transformations efficiency by eliminating the need for extensive screening. This method represents a significant addition to the *A. niger* genome engineering toolbox and could be adapted for use in other organisms. It is expected that this method will impact several areas of industrial biotechnology, such as the development of new industrially relevant fungal strains for the secretion of heterologous enzymes and the discovery and optimization of metabolic pathways.

References

1. Leynaud-Kieffer LMC, Curran SC, Kim I, Magnuson JK, Gladden JM, Baker SE, Simmons BA (2019) A new approach to Cas9-based genome editing in *Aspergillus niger* that is precise, efficient and selectable. PLoS ONE 14(1): e0210243. https://doi. org/10.1371/journal.pone.0210243

The Joint BioEnergy Institute is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

The response of *Pseudomonas putida* to a sorghum lignolysate

Mee-Rye Park[1,2], Yan Chen[1,2], Jian Sun[1,3], Veronica T. Benites[1,2], Edward Baidoo[1,2], John M. Gladden[1,3], Chris Petzold[1,2], Blake A. Simmons[1,2] and **Steven W. Singer** (SWSinger@lbl.gov)[1,2]

[1]Joint BioEnergy Institute, Emeryville, CA; [2]Lawrence Berkeley National Laboratory, Berkeley, CA; Sandia National Laboratories, Livermore, CA

http://jbei.org

Project Goals: The goal of the Joint BioEnergy Institute is to establish the scientific knowledge and new technologies in feedstock development, deconstruction and separation, and conversion needed to transform the maximum amount of carbon available in bioenergy crops into biofuels and bioproducts.

There is strong interest in lignin valorization to produce valuable products. Nevertheless, the bioprocessing of lignin into bioproducts is a major bottleneck because of the structural complexity of the biopolymer. Here, we employed cholinium-based ionic liquid pretreatment to obtain a soluble aromatic-rich fraction by base-catalyzed depolymerization (BCD) of sorghum that had been depleted of sugars by enzymatic hydrolysis. Growth of *Pseudomonas putida* on BCD liquor coincided with disappearance of aromatic peaks, consistent with the complete utilization of *p*-coumarate, ferulate, and the other aromatic monomers present in the depolymerized substrate. The growth of *P. putida* in the BCD liquor was higher than that observed on individual aromatic substrates, suggesting that the BCD liquor contained additional carbon sources beyond lignin-related aromatics. Aromatic-independent growth was confirmed in a *P. putida* mutant strain that was unable to grow on p-coumarate and ferulate. Metabolite analysis demonstrated that the sorghum BCD liquor was a mixture containing at least four distinct substrates for *P. putida* (aromatic monomers, amino acids, cholinium and fatty acids). Comparative proteomic analysis revealed the significant upregulation of aromatic, amino acid and cholinium catabolic pathways as well as genes for fatty acid β-oxidation and acetate assimilation. These results indicate that lignolysates obtained from plant biomass are complex substrates whose assimilation into microbial hosts requires multiple metabolic pathways responding in concert.

This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Title
Investigating Lignin Modifying Enzymes and their Synergistic Effect with Ionic Liquid Pretreatment

Paul Wolski[1,2], Carolina Barcelos[1,3], Michael Doane[1], Kai Deng[1,2], Gabriella Papa[3], Alberto Rodriguez[1,2], Blake Simmons[1,3], Kenneth Sale[1,2]
Deconstruction

1. Joint BioEnergy Institute, Emeryville, CA 94608, USA
2. Sandia National Laboratories, Livermore, CA, 94550, USA
3. Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

https://www.jbei.org/

Project Goals

Biomass recalcitrance is a function of the plant's state of polymerization of various polymers, including sugars and lignin, the interactions among these polymers and their crystallinity. Extracting the valuable sugars and aromatics from biomass requires hash chemical and thermal pretreatment. We hypothesized that by first breaking down lignin using laccases, the severity and thus the expense of the pretreatment process can be reduced and evaluated the synergistic effect of Lignin-Modifying Enzymes (LME's) and ionic liquid pretreatment on glucose yields and reducing the amount of other enzymes required.

Further LME engineering may enhance the rate of lignin breakdown to monomers for fuels or chemicals.

Depolymerization of lignin and subsequent valorization and enhancement of saccharification and reducing the overall cost of 2nd and 3rd generation biofuels by reducing required enzyme loading and cost.

Funding statement.

This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

NitFix: Global-Scale Phylogenomics of the Nitrogen-fixing Clade

Heather Rose Kates, [1] **Ryan Folk***[1] (rfolk@flmnh.ufl.edu), **Robert P. Guralnick**,[1] **Douglas E. Soltis**,[1] **Pamela S. Soltis**[1,] **Jean-Michel Ane**,[2] **Sushmita Roy**,[2] **Matias Kirst**[3]

[1]University of Florida, Florida Museum of Natural History, Gainesville, FL; [2]University of Wisconsin, Madison, WI; [3]University of Florida, Gainesville, FL

www.nitfix.org

Project Goals: The collaborative project "Phylogenomic discovery and engineering of nitrogen fixation into the bioenergy woody crop poplar" is focused on identifying the genomic novelties that enable the symbiotic relationship between nodulating plants and Nitrogen-fixing bacteria to support genetically engineering this capability into bioenergy crops. The first aim of this project is a comparative phylogenomic study of the nodulating clade to uncover the genomic novelties that were required for the evolution of these root nodule symbioses. A massively improved phylogenetic framework for the Nitrogen-fixing clade that includes genetic data for nodulation genes for all species in the phylogeny will provide a robust, improved framework for understanding the origin and evolution of nodulation. These data will also inform a series of experiments comparing close relatives that do or don't to identify the genes underlying nodulation. To achieve these goals, we have developed a cutting-edge phylogenetic approach with an unprecedented sampling effort which is illustrated by the project described below.

The robust inference of the origin of nodulation, and of how nodulation has been gained and/or lost, are key to understanding the evolutionary lability and thus the likelihood of successful transferability of Nitrogen-fixing symbioses to diverse angiosperm crops. A well-resolved and well-sampled Nitrogen-fixing clade phylogeny is therefore a prerequisite to identifying this evolutionary history of symbiosis and to jump-starting the discovery of a core set of genes shared across this diverse group that determine nodule development. Multiple phylogenetic analyses have been conducted on the Nitrogen-fixing clade with the aim of elucidating the origins of Nitrogen-fixing symbioses; however, these analyses have relied on trees estimated using a few genes and in which species sampling in the large Nitrogen-fixing clade (30,000 species) was limited. We present our first steps toward a revised phylogeny based on deliberate and extensive sampling, phylogenomic data, and rigorous statistical analysis. These will allow more accurate inference of precursors of Nitrogen-fixing symbioses, gain and/or loss events, and potential transferability of the capability to crop plants not in the Nitrogen-fixing clade. The final phylogeny will include 15,000 species in the Nitrogen-fixing clade.

This ambitious phylogenetic study requires novel strategies for rapid specimen sampling and DNA data generation. Novel elements of our strategy that enable geographically and taxonomically comprehensive sampling include (1) a protocol for rapid tissue sampling of 15,000 historical (herbarium) specimens, (2) a high-throughput, high-yield DNA extraction protocol specifically suited to degraded DNA, (3) a targeted-enrichment DNA sequencing kit that works across phylogenetic scales and includes key functional genes, and (4) a custom-developed, scalable laboratory information management system (LIMS) together with custom

analysis tools. The protocols and methods developed for this project can be broadly leveraged as a toolkit for phylogenetic researchers to generate very large phylogenetic datasets, and will facilitate greater utilization of historical specimens in phylogenetic research.

The results so far (13% of our ultimate effort) highlight the rapid progress we have made in sampling and data generation and demonstrate the utility of our phylogenetic products for meeting project goals. Our preliminary phylogeny is based on 100 nuclear loci sequenced for 1,900 species from across the Nitrogen-fixing clade (Fig. 1); this tree informs our interpretation of patterns of presence/absence of ~129 nodulation genes inferred from genome screening of all species in our phylogeny. We highlight a set of genes for which there is evidence of correlated evolution with nodulation phenotypes. We show that these evolutionary relationships differ among legume and non-legume nodulators, suggesting more limited genetic homology among these organs than suggested with fewer taxa. We will leverage the total phylogenetic and comparative genomic results to discover gene candidates that potentially underlie nodule development and test these for function in nodulating and non-nodulating model systems.

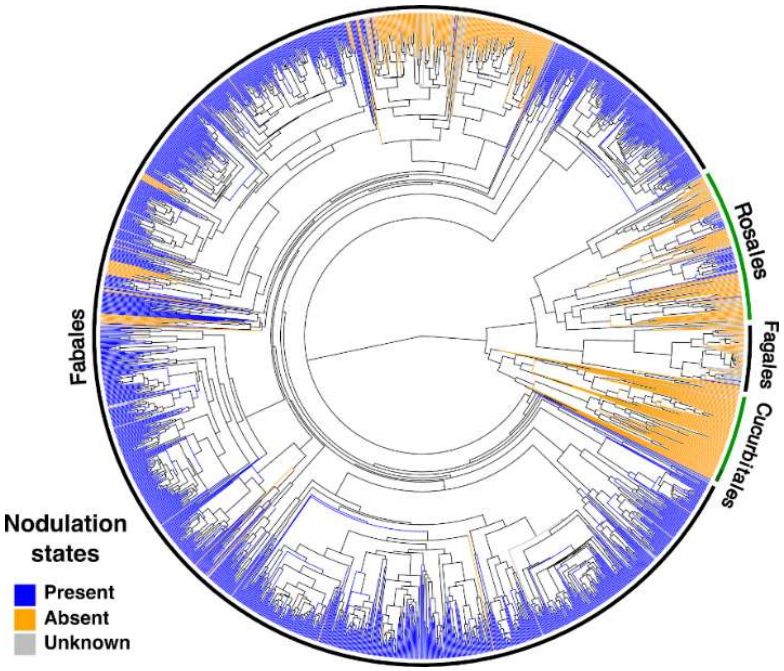


Fig. 1. Preliminary phylogeny of a subset of the Nitrogen-fixing clade based on our gene capture data available at the time of the abstract submission. This tree represents ~13% of the complete phylogeny that will be available for meeting project goals. Terminal branches are colored by presence or absence of nodulation.

This project is funded by the Biosystems Design Program from the Biological and Environmental Research (BER) Office of Science at the U.S. Department of Energy (grant #DE-SC0018247).

NitFix: Phylogenomic Discovery and Engineering of Nitrogen Fixation into the Bioenergy Woody Crop Poplar

Jean-Michel Ané,[1] Kelly Balmant,[2] Daniel Conde,[2] Matthew Crook,[3] Christopher Dervinis,[2] Ryan Folk,[4] Robert P. Guralnick,[4] Thomas Irving,[1] Heather Rose Kates,[4] Sara Knaack,[1] Lucas Maia,[1] Sushmita Roy,[1] Douglas E. Soltis,[4] Pamela S. Soltis,[4] and **Matias Kirst**[2]* (mkirst@ufl.edu)

[1]University of Wisconsin, Madison, WI; [2]University of Florida, Gainesville, FL; [3]Weber State University, Ogden, UT; [4]University of Florida, Florida Museum of Natural History, Gainesville, FL

http://nitfix.org; Twitter: @nit_fix

Project Goals: The collaborative project "Phylogenomic discovery and engineering of nitrogen fixation into the bioenergy woody crop poplar" is focused on identifying the genomic novelties that enable the symbiotic relationship between nodulating plants and nitrogen-fixing bacteria to support genetically engineering this capability into bioenergy crops. Our specific aims are:
(1) Phylogenomic discovery of the underlying mechanism of root nodule symbioses using a comparative phylogenetic framework to contrast related species that possess and lack the nodulation ability.
(2) Verification of molecular mechanisms of root nodule development in *Medicago truncatula* (nodulating) and poplar hairy root cultures (non-nodulating).
(3) Engineering nodulation capability into the bioenergy woody crop *Populus* spp. and testing the impact of these structures on nitrogen-fixation.

Nitrogen (N) availability is critical for high biomass productivity of bioenergy crops, particularly in marginal lands. Despite the abundance of N_2 in the atmosphere and critical importance for growth and development, plants cannot access it. Instead, plants must absorb available N in the soil as nitrates or ammonium, provided through intensive fertilization in a process that is costly, environmentally damaging and potentially hazardous to human health. More efficient and cost-effective approaches are needed to enable bioenergy crops to acquire the N required to maximize growth while minimizing inputs and environmental impact. Some plants have the capability to obtain N through a mutualistic relationship with bacteria and archaea, which can convert N_2 in the atmosphere to NH_3. This capability is limited in the plant kingdom, and absent from most bioenergy crops. All flowering plant lineages known to undergo root nodule N-fixation with bacterial symbionts are within a single clade, where multiple origins and losses of symbioses have occurred.

Phylogenetic approaches and comparative genomics can be used to discover clade innovations and uncover genes that that control N-fixing symbioses by comparing close relative species that have maintained or lost this ability. In order to achieve this goal, it is necessary to obtain a well-resolved and well-sampled N-fixing clade phylogeny. More specifically, we sampled 15,000 species from the N-fixing clade and are currently performing targeted-enrichment of DNA for 129 key functional genes involved in

nodulation, as well as over one hundred phylogenetic markers (further details can be found in the poster “NitFix: Global-Scale Phylogenomics of the Nitrogen-Fixing Clade”). In parallel we are pursuing a series of comparative genomic and transcriptomic approaches. These include the identification of conserved non-coding sequence (NCS) based on the alignment of 34 plant genomes, including nodulating species as well as an outgroup (details are described in the poster “NitFix: Identification of conserved non-coding sequences in nitrogen fixing plants”). We also characterized the root transcriptome and chromatin accessibility of the two species of key interest in this study, *Medicago truncatula* (nodulating) and *Populus* (non-nodulating), treated with lipo-chitooligosaccharides (Nod factors). These analyses identified a series of transcripts and regions of chromatin accessibility that respond to exposure to Nod factors within each species, as well as distinct differences in response between them. These may reflect differences in their ability to nodulate, and are being further evaluated in combination with genome comparisons, to detect target genes to be modified.

In parallel to the discovery of target genes, we are manipulating known genetic components involved in nodule development in a cell-specific manner in poplar (more details in the poster “NitFix: Engineering Root Nodule Symbiosis in *Populus* sp.”). The latter approach has involved developing root cortex and epidermis-specific promoters to modify the expression of genes involved in nodule development, such as Nodule INception (NIN) and cytokinin receptors. This targeted approach has resulted in a significant increase in the number of root lateral organs. We are currently characterizing these lateral organs in poplar and assessing their capability to support the establishment of a symbiotic relationship with N-fixing bacteria.

This project is funded by the Biosystems Design Program from the Biological and Environmental Research (BER) Office of Science at the U.S. Department of Energy (grant #DE-SC0018247).

NitFix: Identification of Conserved Non-Coding Sequences in Nitrogen Fixing Plants

Wendell J. Pereira,[1] Jean-Michel Ané,[2] Pamela Soltis,[3] Douglas Soltis,[3] Robert Guralnick,[3] Ryan Folk,[3] Sushmita Roy,[2] and **Matias Kirst**[1]* (mkirst@ufl.edu)

[1]University of Florida, Gainesville, FL; [2]University of Wisconsin, Madison, WI; [3]University of Florida, Florida Museum of Natural History, Gainesville, FL

http://nitfix.org; Twitter: @nit_fix

Project Goals: The collaborative project "Phylogenomic discovery and engineering of nitrogen fixation into the bioenergy woody crop poplar" is focused on identifying the genomic novelties that enable the symbiotic relationship between nodulating plants and nitrogen-fixing bacteria to support genetically engineering this capability into bioenergy crops. We are using a comparative phylogenetic framework to contrast related species that possess and lack this ability. Genomic novelties are being evaluated for their effect on root nodule development in *Medicago* (nodulating) and poplar (non-nodulating). Genes identified are being engineered into poplar, to test their impact on nitrogen-fixation, whole-plant development and biomass productivity and composition. This abstract focuses on one component of the comparative phylogenetic framework, in which conserved non-coding sequences that may contribute to nodule development are being identified among species of the nitrogen-fixing clade.

Nitrogen (N) is one of the most important nutrients required for plant growth because of its fundamental role as a component of DNA, RNA, and amino acids. However, plants are not able to obtain N directly from the atmosphere, but depend on its availability in the soil, in the form of nitrate, ammonium, or amino acids. Yearly, approximately 120 million metric tons of N fertilizer are produced and used in agriculture. Nitrogen fertilization is responsible for a significant cost of crop production, and an important source of environmental contamination.

Several species of four angiosperm orders (Fabales, Fagales, Rosales and Cucurbitales), developed the capacity of establish a symbiotic relationship with bacteria that are capable of conversion of atmospheric N to ammonium. These bacteria are attracted by signaling molecules secreted by the plant. After the infection is established, root nodules develop to host the bacteria, while the plant benefits by absorbing the produced ammonia. Together, these plant orders are known as the nitrogen-fixing root nodule (NFN) clade.

The restricted phylogenetic occurrence of plant symbiosis with nitrogen-fixing bacteria suggests that a unique genetic event established the foundation for this symbiotic relationship. Previous studies have tried to identify the genetic component of this evolutionary event by searching for conserved genes among species on the NFN clade. However, a complete understanding of the genetic origin of the symbiosis, that encompasses entire genome sequences (including putative regulatory sequences) is lacking. In recent years, the role of conserved, non-coding sequences (CNS) in gene regulation has been extensively demonstrated. In plants, a few studies showed the

presence of many CNS, suggesting that they are under selection and possibly involved in critical biological processes. The main objective of this study was to identify CNS among species that are able to promote nitrogen-fixing root nodule formation. Additionally, by comparing CNS with those reported in species that don't belong to the NFN clade, we sought to detect those that may be involved in the development of NFN symbioses in plants.

The search for conserved sequences across genomes of species capable of nitrogen fixation was carried out by applying a CNS pipeline (Liang et al., 2018) to the analysis of 34 plant genomes. To select the species to be used in this study, all genomes of the orders belonging to the NFN clade and available in the NCBI databases, RefSeq and GenBank, were classified into those capable of N fixation or not. If improved version of these genomes were identified in other databases (e.g. Phytozome), they were used instead. Additionally, nine species with a genome deposited in RefSeq were chosen as an outgroup, representing different orders outside of the NFN clade. All genomes were used in pairwise whole genome alignments using LAST, with the *Medicago truncatula* genome v.5.0 (Pecrix et al., 2018) used as reference. The reference-dependent multiple alignment was produced using ROAST. Next, the identification of conserved element was made according to a scoring system (Liang et al., 2018). The selection of CNS sequences potentially related to the NFN event was done by excluding those detected in all genomes, from those observed in the NFN clade. Finally, the genome context of CNS was inferred from *Medicago truncatula* annotation using bedtools and R scripts. After the exclusion of CNS present in all species, and considering a conservation score of 0.9, 76,989 CNS that are potentially exclusive of the NFN clade were identified. From these, the vast majority (>95%) are within 5Kb of annotated elements (genes, TEs or miRNAs), demonstrating a potential rule in promoting *cis*-regulation of gene expression.

Data derived from this experiment is being continuously improved by the inclusion of new genomes. It will be used now in combination with chromatin accessibility and transcriptome data to further filter for target genes to be evaluated in poplar root cultures.

References

Liang Y et al. (2018) Single-base resolution map of evolutionary constraints and annotation of conserved elements across major grass genomes. Genome Biol. Evol. 10:473–488.

Pecrix Y, et al. (2018) Whole-genome landscape of Medicago truncatula symbiotic genes. Nature Plants, 4:1017–1025.

This project is funded by the Biosystems Design Program from the Biological and Environmental Research (BER) Office of Science at the U.S. Department of Energy (grant #DE-SC0018247).

NitFix: Engineering Root Nodule Symbiosis in *Populus* sp.

Lucas Gontijo Silva Maia[1*$] (lucas.maia@wisc.edu), Thomas Barrance Irving,[1$], Daniel Conde,[2] Matias Kirst,[2] and **Jean-Michel Ané**[1]

[1]University of Wisconsin, Madison, WI; [2]University of Florida, Gainesville, FL;

[$]These authors have contributed equally to this work

Websites: https://nitfix.org/ and https://anelab.wisc.edu/

Project Goals: The collaborative project "Phylogenomic discovery and engineering of nitrogen fixation into the bioenergy woody crop poplar" is focused on identifying the genomic novelties that enable the symbiotic relationship between nodulating plants and Nitrogen-fixing bacteria to support genetically engineering this capability into bioenergy crops. The third aim of this project is engineering nodulation. Root nodules are specialized organs in legumes and actinorhizal plants that provide a proper environment for nitrogen-fixation bacteria such as rhizobia and *Frankia*. We aim to engineer root nodules in lineages of plants that do not form them to improve their associations with nitrogen-fixing bacteria. *Populus* sp. is a model plant unable to develop root nodules and is closely related to legumes and actinorhizal lineages. Genetics in model legumes identified genes such as *NIN* and *LHK1* that are essential and sometimes sufficient to induce nodule development. In this project, we aim to test in *Populus* sp. if these genes are sufficient to develop root nodules. In the future, more candidate genes coming from the Aim 1 of the DOE-funded project will be tested.

Engineering symbiotic nitrogen fixation in crops that do not associate with rhizobia or *Frankia* has been a long-term goal of the scientific community. *Populus* sp. is an essential bioenergy crop and presents natural characteristics that make it a unique platform for such an ambitious engineering project: (1) *Populus* sp. is phylogenetically very close to legumes and actinorhizal lineages. It contains all the known genes required for root nodule symbioses, suggesting that nodules evolved by co-opting pre-existing mechanisms such as lateral root development and symbiotic associations with arbuscular mycorrhizal fungi. (2) *Populus* sp. is easy to transform and to regenerate. The Ané lab has optimized transformation methods for the *Populus tremula x P. alba* hybrid (INRA 717-1B4) using *Agrobacterium rhizogenes* that allow us to generate transgenic roots in 2-3 weeks and full transgenic plants out of these transgenic roots in 2-3 months.

Genetic studied in model legumes, *Medicago truncatula* and *Lotus japonicus*, identified many genes required for nodule organogenesis and colonization by rhizobia. A large fraction of these genes is also involved in lateral root development and associations with arbuscular mycorrhizal fungi suggesting that legume nodulation evolved by recruiting genes from these processes. This observation is not entirely surprising because lateral roots and nodules are both "root lateral organs". Rhizobia and arbuscular mycorrhizal fungi are hosted inside of plant cells (endosymbiosis). Some of these genes are even sufficient to induce nodule development. For instance, overexpression of the NIN (1) transcription factor or the constitutive active LHK1 cytokinin receptor (*snf-2*) led to nodule organogenesis in *L. japonicus* in the absence of rhizobia (2).

In the first year of our DOE-funded project, we identified *Populus* sp. homologs of legume *NIN* and *LHK1*. *Populus* sp. has undergone a significant expansion of the *NIN* family, with seven full-length members and one pseudogene. In the near future, we will test more candidate genes coming from comparative phylogenomic approaches from Aim 1 of our DOE project. In our initial experiments, we expressed the constitutive active *L. japonicus* LHK1 cytokinin receptor (*snf-2*) either constitutively throughout the root or

specifically in the root cortex. We also overexpressed individual *Populus* sp. *NIN* genes in the hybrid INRA 717-1B4 roots.

We observed the surprising development of "root lateral organs" on *Populus* sp. roots expressing *snf-2* and *PtNIN2a*, *b* and *c*. It is too early at this stage to say if these structures are nodules or lateral roots, but some of them present the round morphology typical of root nodules. Further analysis of these structures using microscopy will determine the tissue of origin (cortex or pericycle) and their type of vasculature (central like lateral roots and actinorhizal nodules or peripheral like legume nodules).

Populus sp., like legumes and the vast majority of land plants, associates with arbuscular mycorrhizal fungi. A hallmark of the response of legumes to rhizobia and arbuscular mycorrhizal fungi are oscillations in nuclear and perinuclear calcium concentration, called “calcium spiking”. In the first year of this project, we demonstrated that *Populus* sp. responds to some rhizobia by eliciting “calcium spiking” just like legumes. This calcium spiking was also triggered in *Populus* sp. by the application of lipo-chitooligosaccharides (LCOs) produced by rhizobia and arbuscular mycorrhizal fungi. We have initiated RNA-seq and ATAC-seq experiments to test the response *Populus* sp. to various LCOs.

In the next years, we continue testing more genes and promoters for their ability to induce root nodules in *Populus* sp. We will determine if the structures observed are analogous to legume nodules. We will continue investigating the possibility of getting nodules induced and colonized by rhizobia by manipulating the LCO signaling pathway.

References

1. Soyano T, Kouchi H, Hirota A, Hayashi M (2013) Nodule inception directly targets NF-Y subunit genes to regulate essential processes of root nodule development in *Lotus japonicus*. PLoS Genet 9(3):e1003352.
2. Tirichine L, et al. (2007) A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. Science 315(5808):104–107.

Funding statement: This project is funded by the Biosystems Design Program from the Biological and Environmental Research (BER) Office of Science at the U.S. Department of Energy (grant #DE-SC0018247).

Towards durable resistance to *Septoria* stem canker and leaf spot: a molecular understanding of resistance.

Jared M. LeBoldus[1] (your email: jared.leboldus@science.oregonstate.edu), Susanna Keriö[1],

Wellington Muchero[3], Jin-Gui Chen[3], and Steven Strauss[4].

[1] Department of Botany and Plant Pathology, , Oregon State University, Corvallis, OR, USA.

[2]Forest Engineering, Resources and Management Department, Oregon State University, Corvallis, OR, USA.

[3] Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA.

[4] Department of Ecosystems and Society, Oregon State University, Corvallis, OR, USA.

Fuels developed from lignocellulosic biomass offer a potential renewable and clean alternative to conventional fossil-based energy sources. *Populus* is one of DOE's "flagship" plant species that is of special interest as a biofuel feedstock. *Septoria* canker is the major limiting factor in the use of *Populus* as a biomass feedstock in the central and eastern United States. An effective disease resistance-breeding program has not been developed due to an absence of information on the genetic basis of resistance. To identify resistance alleles, we combine the re-sequenced *P. trichocarpa* genome-wide association population with our robust disease resistance phenotyping platform. This has enabled us to: (1) identify and analyze alleles conferring resistance to *Septoria* stem canker and/or leaf-spot disease; (2) use a combination of stable and transient expression systems in *Populus* and *Nicotiana* to validate the function of the encoded proteins of the identified genes; and (3) field test resistant genotypes in order to validate their performance under changing environmental conditions.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0018196.

Title: EPICON: From Leaves to Roots to Microbes – How Sorghum and Its Microbiome Respond to Drought

Peggy G. Lemaux[1*] (lemauxpg@berkeley.edu), Jeffrey Dahlberg,[2] Devin Coleman-Derr,[1,3] Robert Hutmacher,[4] Christer Jansson,[5] Ronan O'Malley,[6] Elizabeth Purdom,[1] John Taylor,[1] Axel Visel[6] and John Vogel[1,6]

[1] University of California, Berkeley; [2]University of California, Kearney Ag Research & Extension Center, Parlier; [3]USDA-ARS Plant Gene Expression Center, Albany CA; [4]University of California, West Side Research & Extension Center, Five Points; [5]EMSL-Pacific Northwest National Laboratory, Richland WA; and [6]DOE Joint Genome Institute, Walnut Creek CA.

Project Goals:

Analysis of transcriptomic and epigenetic control mechanisms during spatiotemporal responses to water-limiting conditions are being performed on leaves and roots of field-grown, pre-flowering and post-flowering drought-tolerant *Sorghum bicolor* (L.) Moench. Also changes in associated bacterial and fungal communities in bulk soil, rhizosphere, leaves and roots of drought-stressed sorghum are also being studied. The goal of these efforts is to understand mechanisms functioning in acclimation to and recovery from pre- and post-flowering drought, using RNA-Seq, BS-Seq, proteomics, metabolomics, and histone profiling. To provide additional insights into sorghum's drought responses, impact of microbial populations is also being investigated, using metagenomics, metatranscriptomics, and metabolomics. Cumulative data will be used to devise models to better predict and control roles and interactions of transcriptional regulation, epigenetics and the microbiome in sorghum's response to drought. Ultimately, we will identify genes, molecular markers and microbes to devise strategies for improving drought tolerance in sorghum and other crops.

EPICON research focuses on drought, given that its frequency and severity will increase with climate change. It is broadly accepted that both transcriptomic and epigenetic changes play major roles in regulating drought responses. These are the issues driving EPICON research. Spatiotemporal responses of transcriptomic and epigenetic controls were examined in leaves and roots of field-grown, drought-tolerant *Sorghum bicolor* (L.) Moench. RNA-Seq, BS-Seq, proteomics, metabolomics, and histone profiling are used to gain a mechanistic understanding of plant acclimation to and recovery from drought. Shotgun metagenomics, metatranscriptomics, and metabolomics were also used to monitor sorghum's rhizosphere microbiome.

Sorghum bicolor was chosen because it is a widely cultivated, drought- and flood-tolerant cereal – important as a flexible bioenergy feedstock with a relatively reduced environmental footprint. Drought was imposed in fields in California's Central Valley, where rare summer rainfall permits controlled drought conditions. One pre-flowering and one stay-green, post-flowering drought-tolerant variety were planted in a replicated, split plot design, with normal watering and pre- and post-flowering drought. Weekly phenotypic measurements were taken during the growing season, grain and biomass yields at harvest. Most impacted yield-related phenotypes due to pre-flowering drought, were later flowering, shorter stature, lower forage/grain yields.

Year one transcriptional profiling of triplicate, weekly leaf and root samples, revealed widespread adaptations at all developmental stages, including within one week after watering pre-flowering droughted plants and after imposing post-flowering drought – with 44% of expressed genes significantly affected (Varoquaux et al., submitted). Based on 350 transcriptomes, fast, global, temporal transcriptomic responses were observed in leaves and roots, including modulation of well-known drought pathways. Roots showed greater transcriptional disruptions than leaves; pre-flowering drought had more complex temporal changes than post-flowering drought; large differences were found between genotypes.

In-depth studies were done on two drought-related transcriptional responses. (1) Genotypic differences were seen in core photosynthesis and reactive oxygen species (ROS)-scavenging pathways, suggesting possible mechanisms of drought tolerance and delayed senescence in the post-flowering, drought-tolerant, stay-green variety. Baseline gene expression differences between the two varieties may affect how equipped plants are for pre- or post-flowering drought. For example, the stay-green variety has constitutively higher mRNA expression of genes involved in ROS scavenging and osmoprotection, versus the non-post-flowering drought-tolerant variety. (2) Large-scale depletion in expression of genes critical to arbuscular mycorrhizal (AM) symbiosis occurred during drought, with corresponding drops in AM fungal root mass. That drop in mass during pre- and post-flowering drought corresponded with decreased mRNA expression in AM fungal symbiosis-induced genes, suggesting drought leads to reduced AM fungal survival and a loss in vital symbiotic interactions. These gene expression recovery differences were the largest genotype-specific drought response for a single functional category of genes, indicating AM symbiosis may explain some genotype differences during pre-flowering drought recovery.

Preliminary analysis of BS-Seq data, designed to explore DNA methylation patterns, revealed many regions in leaves where changes correlated with plant development, including varietal differences. Additionally, strong overexpression from transposable elements occurred under drought and, at times, continued after water resumption. Current efforts focus on elucidating connections between transposable elements and methylation changes. Also, LC-MS analysis of intact, untargeted leaf histones enabled discovery of novel drought- or development-related histone posttranslational modifications. Data suggest terminal clipping of histones H4 and H3 may regulate plant growth and drought tolerance differently in the varieties.

Metabolomic and proteomic changes in leaf and root samples, identical to those above, are being analyzed in samples in which proteins and metabolites are extracted from the same samples. GC-MS of root and leaf metabolites revealed greater disruptions in roots versus leaves and significant differences between the two varieties. Proline was significantly increased in both varieties during pre-flowering drought versus controls. Glycerol 3-phosphate was significantly increased in roots during pre-flowering drought, which correlated with rhizosphere monoderm enrichment. iTRAQ-labeled peptides revealed significant changes in roots with significantly different protein profiles between the varieties. At the most extreme pre-flowering drought time, ascorbate metabolism, flavonoid and carotenoid biosynthesis, and porphyrin and chlorophyll metabolism pathways were more significantly changed in the stay-green variety versus the pre-flowering drought-tolerant variety. This validates transcriptomic data where genes involved in ROS-scavenging and photosynthesis are more significantly altered in the stay-green variety.

Using soil, root, leaf and rhizosphere samples, collected weekly from the same plants, dramatic shifts in bacteria and fungi followed drought and re-watering. Gene function in those populations was inferred from shotgun metagenomic and metatranscriptomic analyses. Rapid changes in bacterial community composition occurred following pre-flowering drought, revealing relative enrichment in most monoderm (Gram-positive) bacterial lineages (Xu et al., 2018). After re-watering, reversion occurred within one week, leading to domination by diderm (Gram-negative) lineages. Monoderm enrichment during drought was accompanied by increases in transcriptional activity, specifically for genes related to carbohydrate and amino acid transport and metabolism. From metabolomic analyses, drought-treated roots were enriched in many of the same carbohydrate and amino acid metabolites, suggesting interplay between plant metabolism and bacterial community activity. Second year data recapitulates these findings and showed that the root microbiome returns to a diderm-dominated state within eight hours of rewatering.

From year one fungal data, both pre-flowering and post-flowering drought exerted significant effects on fungal diversity and community composition. However, the two drought conditions exerted asymmetrical changes in community composition in roots and rhizosphere. Abundance of plant pathogens, *i.e., Fusarium*, *Gibberella* and *Sarocladium*, decreased in pre-flowering drought but increased in post-flowering drought in rhizosphere, only partially in roots. The rhizosphere fungal community largely followed the root fungal community pattern, but the soil fungal community was not substantially affected. Symbiotic arbuscular mycorrhizal (AM) fungi were found in root, rhizosphere and soil (Gao et al 2018); however, diversity and community composition were not affected by drought. Instead, AM fungal biomass decreased during pre- and post- flowering drought (Varoquaux et al, submitted).

Scale and scope of EPICON data provide unprecedented platforms for in-depth exploration of molecular mechanisms of drought tolerance and its interplay with the plants' biotic environment. Data generated provide many avenues for future research on sorghum and drought – likely relevant to other crops. Ultimately, genes, molecular markers and microbes causally associated with drought tolerance will be identified that improve yield and fitness under drought.

References:

1. Xu et al., 2018. *Proc Natl Acad Sci USA.* www.pnas.org/cgi/doi/10.1073/pnas.1717308115
2. Gao et al., 2018. *Int Soc Micro Ecol J* https://doi.org/10.1038/s41396-018-0264-0
3. Varoquaux et al., *Genome Biology* (submitted)

Funding: DOE Office of Biological and Environmental Research (DE-SC0014081).

Novel Insights into Rare Earth-Mediated Metabolisms, Aerobic Methanotrophy under Hypoxia, and the Role of Nitrogen Species in Methane Oxidation

Yu Z.[1], Zheng Y.[1,2], Huang J.[1], Groom J. [1], **Chistoserdova L.*** (milachis@uw.edu)[1], and **Lidstrom M.E.**[1]

[1]University of Washington, Seattle, USA; [2]Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, China

Project Goals: This project addresses the structure and function of microbial communities active in methane consumption, using lake sediment as a model. Through manipulation of synthetic communities and systems biology approaches, we are striving to understand the molecular mechanisms that form a basis for interspecies interactions in microbial oxidation of methane. In this phase of the project our goals were: 1) Expand the biochemical knowledge on rare earth alcohol dehydrogenases in keystone methylotroph species and beyond, through enzyme purification and analysis; 2) Evaluate response of methane-oxidizing communities to hypoxia, pinpoint main mechanisms for adaptation, and test their significance via mutation; and 3) Directly address the interplay between denitrification and methane oxidation via mutant manipulation.

Through protein expression and analysis, we more than doubled the extant collection of the biochemically characterized rare earth-dependent enzymes, demonstrating a range of catalytic properties and substrate and cofactor specificities. Many of these enzymes reveal propensity for oxidation of methanol. This observation, in combination with genome-based reconstruction of methylotrophy pathways in select species suggests a much wider occurrence of this metabolic capability among bacterial species, and thus further suggests the importance of methylated compounds as parts of the global carbon cycle. We also compiled an inventory of genes potentially encoding rare earth-dependent enzymes closely related to the characterized enzymes, demonstrating their wide distribution among some of the most numerically abundant and environmentally important taxa, suggesting that reliance on rare earth-mediated biochemistries is much more widespread in the microbial world than previously assumed. Overall, our new data both firmly establish rare earth elements as the essential life metals and suggest their universal role in metabolisms that drive major biogeochemical processes, including but not limited to the conversion of methanol.

Through comparative (meta)transcriptomics, we identified potential mechanisms for hypoxia tolerance by *Methylobacter*, a globally occurring cosmopolitan species playing key role in methane consumption in oxic, hypoxic and even anoxic environments. We zoomed into one gene cluster that encodes a hybrid cluster protein and cognate oxidoreductase, along with regulatory proteins, unique to *Methylobacter* species. Mutations in these genes lead to decreased growth with nitrate but not ammonium, suggesting a role for these genes in nitrosative stress.

The role of denitrification was directly tested via mutation, again revealing decreased growth of the mutant in respiratory nitrate reductase, despite low expression of the denitrification genes under all conditions tested.

Overall, our new data provide additional insights into the functionality of methane-oxidizing communities and reveal physiological mechanisms that may be responsible for the dominant role of *Methylobacter* in methane cycle in natural environments with high fluxes of methane, including hypoxic and even anoxic environments. The source of dioxygen for methane activation under the latter condition remains obscure and requires further experimentation.

Publications

1. Zheng Y, Chistoserdova L (2019) Multi-omics understanding of methanotrophs. In: Methanotorphs: microbiology fundamentals and biotechnological applications. EY Lee Ed, Springer Nature. Submitted.
2. Huang J, Yu Z, Groom J, Cheng J-F, Traver A, Yoshikuni Y, Chistoserdova L (2019) Rare earth element-dependent alcohol dehydrogenases widely occur among globally distributed, numerically abundant and environmentally important microbes. Submitted.
3. Puri A, Liu D, Schaefer AL, Yu Z, Pesesky M, Greenberg EP, Lidstrom ME (2019) Interspecies chemical signaling in a methane-oxidizing bacterial community. Submitted.
4. Yu Z, Zheng Y, Huang J, Chistoserdova L (2019) Systems biology meets enzymology: recent insights into communal metabolism of methane and the role of lanthanides. Curr Issues Mol Biol. In Press.
5. Kalyuzhnaya MG, Collins D, Chistoserdova L (2019) Microbial cycling of methane. In: Encyclopedia of Microbiology Fourth Edition. In Press.
6. Chistoserdova L (2019) New pieces to the lanthanide puzzle. Mol Microbiol doi: 10.1111/mmi.14210.
7. Huang J, Yu Z, Chistoserdova L (2018) Lanthanide-dependent methanol dehydrogenases of XoxF4 and XoxF5 clades display differential distribution among methylotrophic bacteria and different biochemical properties. Front Microbiol 9:1366.
8. Zheng Y, Harris DF, Yu Z, Fu Y, Poudel S, Ledbetter RN, Fixen KR, Yang ZY, Boyd ES, Lidstrom ME, Seefeldt LC, Harwood CS (2018) A pathway for biological methane production using bacterial iron-only nitrogenase. Nat Microbiol 3:281-286.
9. Zheng Y, Huang J, Zhao F, Chistoserdova L (2018) Physiological effect of XoxG(4) on lanthanide-dependent methanotrophy. mBio pii: e02430-17.
10. Chistoserdova L (2018) Applications of methylotrophs: can single carbon be harnessed for biotechnology? Curr Opin Biotechnol 50:189-194.
11. Chistoserdova L, Kalyuzhnaya MG (2018) Current trends in methylotrophy. Trends Microbiol 26:703-714.
12. Chistoserdova L (2018) Methanotrophy: an evolving field. In: Kalyuzhnaya M, Xing X-H Eds, Methane biocatalysis: Paving the way to sustainability. Springer, Springer-Verlag, Berlin Heidelberg, Germany, pp. 1-15.
13. Yu Z, Beck DAC, Chistoserdova L (2017) Natural selection in synthetic communities highlights the roles of *Methylococcaceae* and *Methylophilaceae* and suggests differential roles for alternative methanol dehydrogenases in methane consumption. Front Microbiol 8:2392.
14. Yu Z, Chistoserdova L (2017) Communal metabolism of methane and the rare Earth element switch. J Bacteriol pii:JB.00328-17.
15. Chistoserdova L (2017) Omics approaches to studying methylotrophs and methylotroph communities. Curr Iss Mol Biol 24:119-142.

This research is supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC-0016224.

Biolistic Transformation of *Miscanthus S*pecies. Preliminary Studies.

Joanna Cerazy-Waliszewska,[1,2]* (joannac@illinois.edu), Katarzyna Głowacka,[3] Johannes Kromdijk,[4] Tomasz Pniewski,[2] Karolina Sobańska,[2] and **Prof. Stephen Long**

[1]University of Illinois, Urbana, US; [2] Institute of Plant Genetics, Polish Academy of Sciences, Poznań, Poland; [3] University of Nebraska, Lincoln, US, [4] University of Cambridge, UK

https://rogue.illinois.edu/

Project Goals:
Renewable Oil Generated with Ultra-productive Energycanes—or ROGUE—is engineering the two most productive American crops—energycane and Miscanthus—to produce a sustainable supply of biodiesel, biojet fuel, and bioproducts.
The aim of this work is to engineer energycane and Miscanthus to produce an abundance of natural oil that can be converted into biodiesel, biojet fuel, and bioproducts.

Abstract:
The Miscanthus x giganteus has proved a highly productive and cold tolerant C4 biomass feedstock. As a sterile interspecific hybrid, it lacks genetic variation. Bioengineering offers the opportunity to add desirable traits to this productive chassis, such as genes for triacylglyceride accumulation and improved photosynthetic efficiency. Since the crop is clonally propagated, once successfully transformed, the issues that face seed propagated plants in moving to farm-ready propagules, such as copy number, segregation, silencing and breeding into elite cultivars are by-passed.

Three plasmid vectors were used for transformation. The vector pCAHGA (optimized pCAMBIA 1201) with the *ohpt* marker gene witch codon from monocots under the Ubi gene promoter from *Zea mays* (GenBank JX947345.1), and the reporter gene *uid* (GUS) under the Act1 promoter from *Oryza sativa* (GenBank S44221.1). Two other transformation vectors, engineered to encode genes known to modulate the photoprotection pathway in plants were also used [1, 2].

The Miscanthus transformation was carried out using the microparticle bombardment (biolistic method) established by Bioengineering Team from Institute of Plant Genetics, Polish Academy of Sciences, Poznań, Poland.

Callus from *Miscanthus sinensis* and *M.* x *giganteus* was inducted from immature inflorescences on the C-17 medium (with 5 mg/l 2.4-D, 0.5 mg/l BAP and 90 g/l maltose). After 8-10 weeks callus was transferred on the osmotic medium (MIC with 36.4 g/l sorbitol and mannitol each) 24 hours before transformation. Biolistic® PDS-1000/He gene gun and 1.0 μm gold particles were used for transformation. The distance between the stopping screen and target cells was 9 cm. Plant regeneration were conducted on 190-2 medium with 0.5 mg/l KIN and 0.5 mg/l NAA and 200 mg/l of kanamycin or 5 mg/l hygromycin B [3].

Currently, molecular analyses are carried out to check the presence of the transgene in the obtained plants.

References

[1] Głowacka, K., Kromdijk, J., Kucera, K., Xie, J., Cavanagh, A. P., Leonelli, L., Leakey, A. D. B., Ort, D. R., Niyogi, K. K., and Long, S. P. 2018. Photosystem II Subunit S overexpression increases the efficiency of water use in a field-grown crop. *Nature Communications* 9, 1, 868.

[2] Kromdijk, J., Głowacka, K., Leonelli, L., Gabilly, S. T., Iwai, M., Niyogi, K. K., and Long, S. P. 2016. Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science (New York, N.Y.)* 354, 6314, 857–861.

[3] Ślusarkiewicz-Jarzina, A., Ponitka, A., Cerazy-Waliszewska, J., Wojciechowicz, M. K., Sobańska, K., Jeżowski, S., and Pniewski, T. 2017. Effective and simple in vitro regeneration system of Miscanthus sinensis, M. × giganteus and M. sacchariflorus for planting and biotechnology purposes. *Biomass and Bioenergy* 107, 219–226.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (Award Number DE-SC-0018254).

Improving Photosynthesis in C_4 Bioenergy Crops

Kher Xing, Chan[1] (cindyckx@illinois.edu), Fredy, Altpeter[2] and **Stephen P., Long**[1]

[1]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana Illinois. [2]University of Florida, Gainesville, Florida.

https://rogue.illinois.edu

Project Goals

The Renewable Oil Generated with Ultra-productive Energycane (ROGUE) project aims to engineer the two most productive American biofuel crops, energycane and *Miscanthus*, to produce a sustainable supply of biodiesel, biojet fuel and bioproducts. The three main objectives of this work are:

1) **To improve the conversion of sunlight into plant biomass/metabolites through photosynthesis without the need for increased quantities of either water, or fertilizer.**
2) **To transfer ROGUE technologies from the lab bench to crops through an efficient pipeline.**
3) **To test technologies through replicated field trials.**

Abstract

Many important C4 bioenergy crops, such as maize, sorghum and sugarcane, are tropical, warm-season forage grasses with limited productivity under chilling conditions, as compared to temperate C4 grasses, e.g. Miscanthus × giganteus, Panicum virgatum and Spartina pectinata (1). This is due to the limited amount of pyruvate orthophosphate dikinase (PPDK) and rubisco which restrain phosphoenolpyruvate (PEP) regeneration (1–3). In order to expand potential growing seasons and increase the range of latitudes where biofuel crops can grow, two main hypotheses are formed to increase photosynthetic ability of these crops under different conditions. We hypothesize that the upregulation PPDK in C4 crops by engineering PPDK native gene from closely-related species, e.g. sorghum, elevates photosynthesis and biomass. This will be tested by introducing a synthesized native Sorghum bicolor PPDK gene into Miscanthus and energycane to determine the level of photosynthetic improvement in the greenhouse and by field trials. We also plan to modulate the level of PPDK regulatory protein (PDRP) which controls posttranslational modifications of the rate-limiting PPDK under different light conditions (4–6) increases PPDK activity in plants. We will first perform the in silico modelling to understand the potential impact of modulating PDRP activity, then assemble the transgenic construct followed by transient expression of the gene in protoplasts of bioenergy crops, and finally perform physiological assessments in resulting transgenic crops.

References

1. Long SP, Spence AK (2013) Toward Cool C_4 Crops. *Annu Rev Plant Biol* 64(1):701–722.
2. Wang D, Naidu SL, Portis AR, Moose SP, Long SP (2008) Can the cold tolerance of C_4 photosynthesis in Miscanthus x giganteus relative to Zea mays be explained by differences in activities and thermal properties of Rubisco? *J Exp Bot* 59(7):1779–1787.
3. Naidu SL, Moose SP, Al-Shoaibi AK, Raines CA, Long SP (2003) Cold Tolerance of C4 photosynthesis in Miscanthus x giganteus: Adaptation in Amounts and Sequence of C4 Photosynthetic Enzymes. *Plant Physiol* 132(July):1688–1697.
4. Ohta S, Ishida Y, Usami S (2006) High-level expression of cold-tolerant pyruvate, orthophosphate dikinase from a genomic clone with site-directed mutations in transgenic maize. *Mol Breed* 18(1):29–38.
5. Chastain CJ, et al. (2011) Functional evolution of C4 pyruvate, orthophosphate dikinase. *J Exp Bot* 62(9):3083–3091.
6. Burnell JN, Chastain CJ (2006) Cloning and expression of maize-leaf pyruvate, Pi dikinase regulatory protein gene. *Biochem Biophys Res Commun* 345(2):675–680.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (Award Number DE-SC-0018254).

Overcoming Recalcitrance of Genetic Transformation in Energycane for Improving Cold Tolerance and Biomass Yield

Baskaran Kannan[1]* (kbaskaran@ufl.edu), Matthew Schneider[1], Thaibinhduong Nguyen[1], Paul South[2], Don Ort[2], Steve Long[2] and **Fredy Altpeter[1]**

[1]University of Florida, Gainesville, FL; [2]University of Illinois, Urbana, IL

https://rogue.illinois.edu/

Project Goals:

The Renewable Oil Generated with Ultra-productive Energycane (ROGUE) project aims to engineer the two most productive American biofuel crops, energycane and *Miscanthus*, to produce a sustainable supply of biodiesel, biojet fuel and bioproducts.

1) **Overcoming recalcitrance in tissue culture of energycane to establish an efficient genetic transformation protocol.**
2) **Over expression of pyruvate Pi dikinase in energycane to enhance biomass yield and cold tolerance.**

Abstract:

Energycane, an interspecific hybrid within the genus saccharum, has a superior photosynthetic efficiency, biomass accumulation and persistence in tropical and subtropical regions. In contrast to sugarcane, energycane has a high proportion of the ancestral species *Saccharum spontaneum* in its genome which contributes to higher tiller number, biomass yield and persistence in addition to a reduced stem diameter and sugar content. Pyruvate orthophosphate dikinase (PPDK) has been proposed as rate limiting enzyme in C4 photosynthesis. It regenerates the substrate phosphoenol pyruvate (PEP) for the initial carbon-fixation step. C4 plants are also severely limited by low temperature, possibly because PPDK is highly cold-labile and partially dissociates below 14 ºC. A previous study suggested that *Miscanthus x giganteus* achieves cold tolerance by increasing the amount of the enzyme rather than its intrinsic properties. Therefore, we decided to explore the over-expression of *Miscanthus x giganteus* PPDK in energycane under its native regulatory sequences. However, energycane is far more recalcitrant in tissue culture than sugarcane. Visual browning of the newly excised explants is a major hurdle that needs to be overcome to establish an efficient genetic transformation protocol for this target species.

Experiment I: We investigated effects of several supplements (e.g. anti-oxidants/anti-browning agents) in the tissue culture medium on visual tissue browning in energycane. Reduction in visual tissue browning, callus induction and regeneration response were evaluated to determine the optimal concentration and combination of these anti-browning agents when added to culture media. The combination of 2 to 3 anti-browning agents significantly reduced visual tissue browning while increasing the number of regenerating plantlets from energycane callus.

Experiment II: Calli were generated from five different genotypes of energycane in order to evaluate the genotype response to callus induction and regeneration. Two energycane genotypes with the highest callus induction and regeneration frequency were selected for biolistic transformation. Transgenic calli were regenerated on the media containing selection agent and regenerated plantlets were transferred to the soil. Independent transgenic events were confirmed by PCR.

Experiment III: In the present study, Miscanthus PPDK gene (*M-PPDK*) was introduced under its native regulatory sequences into energycane callus by biolistic gene transfer. The transgenic calli were regenerated on the media containing selection agent. The regenerated plants will be evaluated for the effect of *PPDK* overexpression on photosynthetic efficiency, cold tolerance and biomass accumulation.

References

Matsuoka S., Kennedy A.J., dos Santos E.G.D., Tomazela A.L., and Rubio L.C.S. 2014. Energy cane: Its concept, development, characteristics, and prospects, Adv. Bot., Article ID 597275.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (Award Number DE-SC-0018254).

Techno-economic Feasibility Analysis of Biodiesel and Ethanol Co-production from Lipid-producing Energycane

Deepak Kumar[1]* (kumard@illinois.edu), Stephen Long[1] and **Vijay Singh[1]**

[1]University of Illinois, Urbana

https://rogue.illinois.edu/

Project Goals

Renewable Oil Generated with Ultra-productive Energycanes—or ROGUE—is engineering the two most productive American crops—energycane and Miscanthus—to produce a sustainable supply of biodiesel, biojet fuel, and bioproducts. The goal of this work is to determine the techno-economic feasibility and competitiveness of using engineered energycane as feedstock for biodiesel and ethanol production at commercial scale.

Abstract

Biodiesel is a promising renewable alternative fuel to petro-diesel. However, high feedstock cost and low oil yields per unit land from temperate oilseed crops limits the growth of commercial scale biodiesel production. Through the ROGUE project, energycane, one of the most productive crops in the US, is being engineered to accumulate lipids, which would open the way for production of far more industrial vegetable oil per unit land than previously possible.

The objective of this work is to perform a techno-economic analysis for commercial scale biodiesel production from the engineered energycane. Compared with conventional sugarcane (Saccharum spp.), energycane is dedicated bioenergy feedstock rich in fiber, low in sucrose, and more persistent on marginal soils [1-2]. Energycane contains 22.5% soluble sugars (in juice) and 70.6% fiber on dry mass basis. The fiber consists of 38.8% cellulose, 23.4% hemicellulose, and 21.5% lignin [3]. As first approximation, if all energy from the sucrose that normally accumulates in the stem is diverted to triacylglycerides (TAG), energycane could accumulate up to 8% lipid by weight in its stem (dry mass basis). Comprehensive process models for bio-refinery for coproduction of biodiesel and ethanol from energycane, with 1600,000 MT/year (200 operating days) energycane processing capacity, were developed in SuperPro Designer, assuming 2, 5, and

8% lipid concentration in the harvested stem. The bio-refinery was modeled to produce biodiesel (by transesterification of extracted lipids) and ethanol (from fermentation of sugars in juice and carbohydrates in fiber) as main products, while glycerol and electricity are produced as coproducts. The process simulations results would provide the production yields of biodiesel and ethanol, unit biofuel production cost, and economic profitability, indicating the competitiveness of energycane as a feedstock for biodiesel production compared to conventional oil crops.

Reference

1. Matsuoka, S.; Kennedy, A.J.; dos Santos, E.G.D.; Tomazela, A.L. and Rubio, L.C.S., Energy cane: Its concept, development, characteristics, and prospects. *Adv Botany* **2014**, Article ID 597275
2. Kim, M. and Day, D.F., Composition of sugar cane, energy cane, and sweet sorghum suitable for ethanol production at Louisiana sugar mills. *J Ind Microbiol Biotechnol* **2011**, 38:803–807
3. Aragon, D.; Lu, S. and Kochergin, V., Conversion of energy cane and sweet sorghum into biofuels and chemicals: a modeling approach. In New crops: Bioenergy, biomaterials, and sustainability, ed. J. Janick, A. Whipkey, and C. Von Mark. Tucson, AZ: The University of Arizona, **2015**

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (Award Number DE-SC-0018254).

Characterizing photosynthetic capacity of sugarcane under fluctuating lights

Moonsub Lee[1]* (mlee128@illinois.edu), Fredy Altpeter[2] and **Donald Ort[1]**

[1]University of Illinois, Urbana and [2] University of Florida, Gainesville

https://rogue.illinois.edu/

Project Goals:
Renewable Oil Generated with Ultra-productive Energycanes—or ROGUE—is engineering the two most productive American crops—energycane and Miscanthus—to produce a sustainable supply of biodiesel, biojet fuel, and bioproducts. Project goals are to:
1) Engineer energycane and Miscanthus to produce an abundance of natural oil that can be converted into biodiesel, biojet fuel, and bioproducts.
2) Improve how plants convert sunlight into plant matter through photosynthesis without more water or fertilizer.

Abstract
C_4 bioenergy crops such as sugarcane and *Miscanthus* have attracted attention as feedstocks for alternative energy sources. The productivity of C_4 bioenergy crops is reliant on the high photosynthetic efficiency of C_4 photosynthesis. However, the rapidly changing light environment within canopies of these C_4 grasses prevents C_4 photosynthesis from operating at its highest efficiency. How fluctuating light effects C_4 bioenergy crop productivity is not fully understood. We hypothesize that the metabolite pools of C_4 photosynthesis buffer photosynthetic fluctuations during fluctuating light conditions in C_4 leaves. The previous modeling suggested the metabolite buffering capacity of C_4 photosynthesis could be capable of sustaining high rates of photosynthesis for up to 15s following a high light to low light transition or facilitate a more rapid return to high rates of photosynthesis following a low to high light transition. Thus we hypothesize that increased chloroplast volume may increase this metabolite buffering capacity and thus enhance the performance of photosynthesis in fluctuating light. We tested leaf gas exchange of transgenic sugarcanes that have increased chloroplast volume due to the insertion of an FtsZ RNAi construct to determine whether increased chloroplast volume could enhance the performance of photosynthesis in fluctuating lights. Our preliminary results indicate that the transgenic sugarcane more rapidly recovers high photosynthetic rates during low to high light transition when compared to wild-type sugarcane. Simulating dynamic light conditions can provide insight into the interaction between C_4 photosynthesis and dynamic light.

References

1. Slattery, R.A., Walker, B.J., Weber, A.P., and Ort, D.R. (2018). The impacts of fluctuating light on crop performance. Plant physiology, 176, 990-1003.
2. Grennan, A.K., Ort, D.R., Moose, S.P., Bilgin, D.D., Clemente, T., Altpeter, F., and Long, S.P. (2016). U.S. Patent Application No. 14/995,109.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (Award Number DE-SC-0018254)

Modular Assembly of Gene Constructs for Engineering Lipid Accumulation into Energycane

Guangbin Luo[1*] (g.luo@ufl.edu), John Shanklin[2] and **Fredy Altpeter**[1]

[1]Agronomy Department, IFAS, University of Florida, Gainesville, FL
[2]Biosciences Dept., Brookhaven National Lab, Upton, NY

https://rogue.illinois.edu/

Project goals

The Renewable Oil Generated with Ultra-productive Energycane (ROGUE) project aims to engineer the two most productive American biofuel crops, energycane and *Miscanthus*, to produce a sustainable supply of biodiesel, biojet fuel and bioproducts. The main aim of this work is:

1) **Adding value to energycane for production of advanced biofuels.**
2) **Converting energycane into a lipid producing crop while retaining its superior biomass accumulation.**

Abstract

Metabolic engineering to divert carbon flux from sucrose to oil in a high biomass crop like energycane has been proposed as a strategy to boost lipid yields per acre for biodiesel production (Zale et al. 2016). The energy content of plant oils in the form of triacylglycerols (TAGs) is two-fold greater compared to carbohydrates. However, vegetative plant tissues do not accumulate oil to a significant amount since fatty acid synthesis in these tissues serves primarily membrane construction, in addition TAGs undergo rapid turnover. Therefore, our objectives include:

1) Increasing fatty acid synthesis by expressing a transcription activator of fatty acid biosynthetic genes,

2) Increasing TAG synthesis from diacyl-glycerol and acyl-CoA by over-expression of rate limiting enzymes,

3) Optimizing TAG storage by limiting the access of lipases to TAG storage compartments.

Experiment

To explore the effect of different versions of target genes under different regulatory signals on TAG accumulation in vegetative tissue we developed a library of regulatory elements and open reading frames. These components were used for modular assembly into multi-gene constructs by Golden Gate cloning. Gene expression cassettes were co-delivered with the selectable *npt*II expression cassette by biolistic gene transfer into energycane callus. Plants are currently regenerating on geneticin containing culture medium and will be analyzed for presence and expression of target constructs by PCR and quantitative RT-PCR, respectively. Plants will also be analyzed for TAG content by Gas-Chromatography and Mass Spectrometry (GC-MS).

References

Zale, J., J.H. Jung, J.Y. Kim, B. Pathak, R. Karan, H. Liu, X. Chen, H. Wu, J. Candreva, Z. Zhai, J. Shanklin, F. Altpeter. 2016. Metabolic engineering of sugarcane to accumulate energy-dense triacylglycerols in vegetative biomass. Plant Biotech. J. 14: 661-669.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (Award Number DE-SC-0018254).

Identifying Physiological and Metabolic Response Mechanisms Associated with Divergence in Chilling and Freezing Tolerance Between Upland and Lowland Switchgrass Cultivars.

Anne-Sophie Bohrer[1,2]* (bohreras@msu.edu), Linus Gog[1,3], Darlene Brennan[1,4], Anthony Schilmiller[2,5], A. Daniel Jones[2,5], Thomas D. Sharkey[1,3], Hideki Takahashi[1,2], and **David B. Lowry**[1,4]

[1]Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, Michigan; [2]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan; [3]MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, Michigan; [4]Department of Plant Biology, Michigan State University, East Lansing, Michigan; [5]Mass Spectrometry and Metabolomics Core, Michigan State University, East Lansing, Michigan.

URL: https://www.glbrc.org/research/sustainable-cropping-systems

Project Goals

The goal of this project is to identify traits underlying divergence in cold tolerance (chilling and freezing) and flowering time between cold-tolerant (northern upland) and cold-sensitive (southern lowland) cultivars of switchgrass (*Panicum virgatum*). Accomplishing this goal will facilitate selection for high-yielding, cold-tolerant cultivars that can thrive despite adverse winter (freezing temperatures) or early spring (chilling events) conditions in the northern United States.

Abstract

Research on bioenergy crop production on marginal lands is important for environmentally sustainable production of biofuels and bioproducts. Switchgrass (*Panicum virgatum*), a potential lignocellulosic bioenergy crop species, is a native North American prairie grass with a range of genetic diversity, environmental adaptability, and high biomass yield.

Upland cultivars of switchgrass can survive cold winters in the northern U.S., but have lower biomass yield, in part because of their early flowering traits. In contrast, lowland cultivars accumulate more biomass, are more nutrient-use efficient, are more tolerant to drought and heat, and are more resistant to pathogens, but are less tolerant to cold environmental conditions.

In this study, we focused on the physiological and metabolic responses to a chilling event of two parental cultivars, VS16 (Nebraska-Summer-upland) and WBC3 (Texas-lowland), and their F1 hybrid (WBC3xVS16). For this purpose, plants were grown on a 14 h light – 10 h dark cycle with a light intensity of 1200 µmol/m^2/sec at 25°C during the day and 20°C at night (day 1, pre-chilling event). Plants were then submitted to a 24-hour chilling event, where the temperature was lowered to 10°C during the day and 5°C at night (day 2). On day 3 (post-chilling event), the temperature settings were identical to the initial condition (day 1).

For each cultivar, photosynthetic parameters including carbon assimilation (A), photosynthetic efficiency of photosystem II (F_v/F_m and $F_{v'}/F_{m'}$), stomatal conductance (g_s), internal CO_2 concentration (C_i) and intrinsic water use efficiency (iWUE) were measured. Pre-chilling event, the upland cultivar WBC3 and the F1 hybrid have the highest photosynthesis rate. Although all three cultivars are significantly affected by the chilling event (day 2), the WBC3 cultivar is more sensitive, with lasting damage to the photosynthetic electron transport chain (day 3). The cold-tolerant cultivar VS16 is distinct from the WBC3 cultivar and the F1 hybrid as it closes its stomata in response to the chilling event but continues to assimilate carbon at high rates. As a result, it

maintains good iWUE. This feature may explain how VS16 achieves cold-tolerance. The F1 hybrid, although affected by the chilling event, recovered in a similar manner to VS16. Overall, the F1 hybrid shows combined characteristics of both parental lines. To assess the metabolic responses of each cultivar to the chilling event, leaf samples were collected and their amino acid profiles were analyzed using LC-MS/MS.

In addition to chilling, we are currently conducting studies on these cultivars to understand the mechanisms of differential tolerance of upland and lowland rhizomes to freezing. Overall, the results of these studies will provide a new understanding of the divergence in cold tolerance between northern upland and southern lowland cultivars, which will facilitate the development of high yielding switchgrass cultivars that can survive northern winter conditions.

Funding statement

This work is supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409.

Identification of Adaptive Fungal Pathogen Resistance Loci in Switchgrass

David B. Lowry[1,*] (dlowry@msu.edu), Acer VanWallendael[1], Shawn C. Kenaley[2], Gary C. Bergstrom[2]

[1]Michigan State University, East Lansing, MI; [2]Cornell University, Ithaca, NY

Project Goals:

Switchgrass (*Panicum virgatum*) is an important target species for domestic production of cellulosic biofuels, but it is susceptible to multiple fungal pathogens. The principle aim of our research is to identify the loci responsible for disease resistance in switchgrass and determine how effective those resistance loci are across geographic space. To accomplish this overarching goal, we will: 1) Characterize the pathogens associated with disease in switchgrass and quantify their geographic distributions. 2) Discover genetic loci for effective switchgrass disease resistance across different geographic locations. 3) Validate QTLs for pathogen resistance through controlled experiments. Overall, the discovery of loci and genes involved in resistance to specific pathogens will make crucial improvements of switchgrass cultivars possible through future breeding and gene editing efforts.

Abstract:

Switchgrass is an important target species for domestic production of cellulosic biofuels. The principal aim of most switchgrass breeding programs is to develop high-yielding cultivars. However, as feedstock plantings expand, so will pathogen pressure. Unless controlled, fungal pathogens with explosive disease potential will likely drive yield declines and economic losses. Pathogen resistance can be developed through breeding programs that exploit natural genetic variation in disease resistance. Much of the functional genetic variation in switchgrass, including pathogen resistance, is distributed clinally with latitude as well as between ecotypes. In general, southern lowland cultivars are more resistant to fungal pathogens than northern upland cultivars.

To identify quantitative trait loci (QTL) responsible for pathogen resistance, we conducted QTL mapping of rust infection on a mapping population planted across the central United States. The mapping population was created by crossing two northern upland accessions with two southern lowland accessions to create a four-way outbred population. This outbred mapping population was planted in 2015 at an unprecedented geographical scale, spanning ten common garden field sites distributed over 17 degrees of latitude in the central United States. Over the following three years (2016-2018), we collect ~149,000 rust infection scores on the mapping population across eight field sites. This data allowed us to identify two major QTLs and many minor QTLs contributing to rust resistance across field sites. The two major QTLs had large effects at northern field sites, but had little or no detectable effect at the southern field sites. This suggests that there are either different rust types in the north or that resistance is modified by local environmental conditions. To examine whether the abundance of different rust species was diverged between the north and the south, we used a combination of microscopy and molecular methods on field collected samples. This analysis found that *Puccinia panici*

was by far the dominant rust species present at all of the sites we surveyed, which suggests that species-level differences in resistance are not responsible for the QTL differences between the north and the south. We are currently developing population genetic approaches to determine whether population structure within *P. panici* could explain the latitudinal QTL patterns. We are also developing GIS maps of each field site to evaluate potential environmental covariates that could explain QTL variation among sites. In the summer of 2018, we planted a genome-wide association study (GWAS) mapping population at all of the field sites, as well as two field sites in central Mexico. We will score rust infection on those plantings and conduct a GWAS analysis, which should allow us to identify genes responsible for variation in rust resistance. Overall, our efforts are providing foundational research that will facilitate the future development of more rust resistant cultivars of switchgrass.

Funding Statement:

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0017883.

Genetic and Environmental Contributions to Switchgrass Biofuel Traits

Acer VanWallendael[1,2,*] (vanwaax07@gmail.com), Rebecca G. Ong,[3] Laura E. Bartley,[4] Li Zhang,[5] Jason Bonnette,[5] Robert B. Mitchell,[6] Thomas E Juenger,[5] and **David B. Lowry[1,2]**

[1]Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI; [2]Department of Plant Biology, Michigan State University, East Lansing, MI; Department of Chemical Engineering, Michigan Technological University, Houghton, MI; [4]Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK; [5]Department of Integrative Biology, The University of Texas at Austin, Austin, TX; [6] USDA-ARS Wheat, Sorghum, and Forage Research Unit, University of Nebraska, Lincoln, NE

Project Goals:

1. Quantify traits associated with biofuel yields from switchgrass across environmental conditions.
    a. Using controlled digestion, quantify the yield of glucose and pentose sugars from three clonally replicated field plantings in Texas, Missouri, and Michigan
    b. Using Near-infrared spectroscopy (NIRS), estimate several key compounds in switchgrass biomass for switchgrass planted in eight field sites in Texas, Oklahoma, Missouri, Nebraska, Kansas, and Michigan.
2. Elucidate the genetic architecture of biofuel traits.
    a. Using quantitative trait locus (QTL) mapping, identify loci associated with important biofuel traits across field sites.
3. Evaluate the degree to which traits vary across different environmental conditions.
    a. Quantify how QTL effects for biofuel traits vary across space to evaluate genotype x environment (GxE) interactions.

Abstract

A key to successful bioenergy production will be economically efficient decomposition and conversion of biomass to liquid fuels. Unfortunately, recent research suggests that environmental variation in growing conditions can alter biomass in way that can interfere with decomposition and conversion. Therefore, it will be necessary to understand how both genotype and environmental variation in feedstock crops contributes to variation in traits associated with cell wall digestibility as well as influences downstream decomposition and conversion processes.

In this study, we evaluated digestibility and cell wall traits for a switchgrass genetic mapping population planted at multiple locations throughout the central United States, from Texas to Michigan. We used two complementary techniques to evaluate key traits of switchgrass biomass collected from each site. The first technique, digestibility analysis, involved controlled digestion of switchgrass biomass followed by the quantification glucose and pentose sugar yields. The second technique, NIRS (Near-infrared spectroscopy), uses calibrated spectral data from NIR light to infer chemical composition of a sample for a number of important compounds, including nitrogen, ethanol, calcium, and several types of sugars. For each of the quantified traits, we

conducted QTL mapping to understand the genetic architecture underlying variation in each trait. Because traits had been scored at multiple field sites, we were also able to identify QTL x Environment interactions contributing to trait variation. We performed controlled digestion for plants at three sites and NIRS analysis at eight field sites.

For the digestibility analyses, we identified one QTL for glucose yield, and three pentose yield QTLs, all three of which had significant interactions with the environment. For the NIRS analysis, we faced challenges in the calibration of several compounds. Using the measure of divergence between spectra (global H), traits deviated significantly from switchgrass biomass calibrations ($H > 3.0$). Therefore, we used the general calibrations for forage biomass to calculate compound concentrations. Overall, we mapped twenty NIRS QTLs for biomass collected in the middle of the season, and 27 for biomass collected at the end of season. Importantly, nearly all of the QTLs had significant GxE. Further, while several traits showed significant QTLs at both mid- and end-season timepoints, the majority only were detected at one of the two time points. This suggests that phenology plays an important role in the chemical composition of switchgrass biomass.

Overall, our results indicate that there is significant genotype, environment, and genotype x environment effects on key traits that will influence decomposition and conversion of switchgrass feedstocks to liquid fuels. Thus, researchers using switchgrass biomass should expect geographic and genetic variation in chemical composition that may impact downstream biofuel yields. We are currently in the process of determining the extent to which these sources of variation contribute to those downstream decomposition and conversion processes.

Funding statement:

This work is supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409. Additional funding was provided by a DOE Sustainability grant under Award Number DE-SC0014156.

Optimizing RNA *in situ* Hybridization for Stem Parenchyma Cell-specific Promoter Characterization in Energycane

Jiang Wang[1,2], Ya Chi Yu[1] and **Li-Qing Chen**[1,2]

[1] Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[2] Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA

https://rogue.illinois.edu/

Project Goals:
Renewable Oil Generated with Ultra-productive Energycanes—or ROGUE—is engineering the two most productive American crops—energycane and Miscanthus—to produce a sustainable supply of biodiesel, biojet fuel, and bioproducts. This work investigates mature stem parenchyma cell-specific genes with the aim to increase oil production in the stem.

Abstract:
Triacylglycerols (TAG) are major components of plant oil. Engineering energycane to produce an abundance of TAG for further conversion into biodiesel, biojet fuel, and bioproducts, is an important goal of ROGUE. Constitutively engineering TAG biosynthesis throughout the plant may produce pleiotropic effects, therefore manipulations should be restricted in mature stem parenchyma cells, where an efficient conversion of stored photoassimilates into TAG is practicable. However, the information of mature stem parenchyma cell-specific genes still lacks up to now. RNA *in situ* hybridization is a powerful tool to determine gene tissue specific expression at the cellular level. The lack of the previous report in energycane and the hard texture of mature stem hinders the histological analysis using RNA *in situ* hybridization. Here, we presented an optimized method for RNA *in situ* hybridization in energycane. Eight-micron thick paraplast-embedded cross sections can be achieved for both immature stem (internode 5) and mature stem (internode 16). This optimized method has been validated by hybridizing with an anti-sense probe against a *ScLSG* gene, which has been reported to express in the stem. Now, the method in energycane works consistently. We will be using this method to screen for the best candidate gene that will have higher abundance in pith parenchyma cells of the mature stem. Once the mature stem parenchyma cell-specific promoter is being identified, it will be used for engineering TAG biosynthesis in energycane.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (Award Number DE-SC-0018254).

Systems Biology to Improve Camelina Seed and Oil Quality Traits

John Browse[1], Jesse Bengtsson[1], Wilson Craine[1], Scot Hulbert[1], Kevin King[2], Huang Li[2], GunNam Na[2], Mehmet Ozseyhan[2], and **Chaofu Lu**[2*] (clu@montana.edu)

[1]Washington State University, Pullman; [2]Montana State University, Bozeman

Project Goals: Camelina has great potential to become a sustainable high energy-yielding source of biofuel in the US. This project aims to address two critical needs for realizing this potential: to increase seed size and oil content for improved seedling establishment and oil yield, and to optimize oil quality for satisfactory fuel properties. Specifically, quantitative trait loci (QTL) and molecular markers associated with these important traits will be identified using high-density genome maps and repeated field trials in Montana and Washington states. Modern genomics and biotechnological approaches will be employed to uncover novel molecular mechanisms (including genes and gene networks regulated by microRNAs and transcription factors) regulating fatty acid modification, oil accumulation and seed size in Camelina.

Progresses have been made in the following areas during the last reporting period:

1. To identify quantitative trait loci (QTL) and molecular markers associated with seed size, oil content and other important agronomic traits. 1) A population of recombinant inbred lines (RILs) derived from the cross of two contrasting varieties was used for linkage mapping. SNP markers were generated by genome sequencing, and a linkage map comprising nearly 2400 SNPs was constructed that spanned a genetic distance of about 2035cM covering ~78% of the camelina genome. The RIL population (at the F5 and F6 generations of single seed descent) was grown in Bozeman, MT for phenotypic evaluations. A total 38 QTL were detected for seed and pod sizes (e.g., area, length and width). Several QTL were also detected for other important traits such as oil content, plant height, and flowering time. Significant QTL were chosen for future fine mapping and candidate gene identification. 2) A panel consisting 230 accessions of *Camelina sativa* was used for genome wide association studies (GWAS). These lines were planted in the fields in Bozeman, MT and Pullman, WA in 2017 and 2018. Several important traits evaluated, including seed size, oil content and flowering time, showed wide variations. Genotyping by genome resequencing is being conducted in collaboration with the Joint Genome Institute (JGI).
2. To discover novel molecular mechanisms that regulate fatty acid modification and seed size in camelina. 1) Overexpressing microRNA167A (miR167OE) in camelina under a seed-specific promoter changed fatty acid composition and increased seed size [1]. The miR167OE seeds had a lower α-linolenic acid with a concomitantly higher linoleic acid content than the wild type. This decreased level of fatty acid desaturation corresponded to a decreased transcriptional expression of the camelina fatty acid dsesaturase3 (*CsFAD3*) in developing seeds. Chromatin immunoprecipitation experiments combined with gene expression studies indicated that the miR167 targeted CsARF8, which bound to promoters of camelina *bZIP67* and *ABI3* genes. These transcription factors directly or through the ABI3-bZIP12 pathway regulate *CsFAD3* expression and affect the α-linolenic acid accumulation. Also, comparative transcriptome analysis was conducted to study mechanisms that caused increased seed size in mir167OE. Expression levels of many genes were altered in miR167OE compared to wildtype, including orthologs that have previously been identified to

affect seed size in other plants. Most notably, genes for seed coat development such as suberin and lignin biosynthesis were down-regulated. These results suggested possible mechanisms that we will test to increase seed size in camelina. 2) Seed-specific suppression of ADP-glucose pyrophosphorylase (AGPase) increased seed size and weight [2]. The RNAi suppression reduced AGPase activities which concurred with moderately decreased starch accumulation during seed development. Transcripts of genes examined that are involved in storage products were not affected, but contents of sugars and water were increased in developing seeds. The transgenic seeds were larger than wild-type plants due to increased cell sizes in seed coat and embryos, and mature seeds contained similar oil but more protein contents. The larger seeds showed advantages on seedling emergence from deep soils. This study indicated that changing starch and sugar metabolism during seed development may increase the size and mass of seeds without affecting their final oil content in Camelina. Increased seed size may improve seedling establishment in the field and increase seed yield.

3. Modification of fatty acid composition in camelina seeds. 1) Artificial microRNA was used to down-regulate the expression of *FATB* in camelina seed [3]. Over 40% reduction of saturated fatty acids (16:0+18:0) was observed in transgenic seeds compared to the non-transgenic wild type. 2) A two-gene strategy for reducing oil saturates in camelina oil. To reduce saturated fatty acids (16:0) in the oil, we expressed a DES9* glycerolipid desaturase, which we developed by directed evolution, in seeds of Camelina. As we reported in our 2018 abstract, this resulted in substantial lowering of 16:0. However, further experiments indicated that oil content and seed weight of the transgenic plants was severely compromised. We have also used seed-specific expression of the FAT5 acyl-CoA desaturase from *C. elegans*. However, by itself FAT5 resulted in only small reductions in 16:0 content of the oil. Because the FAT5 and DES9* desaturases act on different lipid substrates and at different stages of oil synthesis, we hypothesized that more modest expression of the two enzymes together might provide for a better result. To accomplish this, we cloned the FAT5 and DES9* coding sequences into a single binary vector under control of different seed-specific promoters, and used the vector for transformation of camelina plants. Our initial results from FAT5-DES9* transgenic plants indicate that this strategy has worked. The double transgenic plants have only 1.8% 16:0 in the oil (compared to 7.5% in the parental line), while retaining oil content and seed weight comparable to the parental line. We are now undertaking detailed biochemical and physiological studies of seed metabolism and physiology to compare our FAT5-DES9* lines with untransformed controls. Taken as a whole our work will not only lead to camelina lines with greatly improved oil characteristics, but will also result in an improved understanding of the metabolism and physiology of oilseed crops.

Publications

1. Na, G., Mu, X., Grabowski, P., Schmutz, J. & Lu, C. (2019) *Plant J*, doi:10.1111/tpj.14223.
2. Na, G., Aryal, N., Fatihi, A., Kang, J. & Lu, C. (2018) *Biotechnol Biofuels* **11**, 330.
3. Ozseyhan, M. E., Li, P., Na, G., Li, Z., Wang, C. & Lu, C. (2018) *Biochem Biophys Res Commun* **503**, 621-624.

This research project is supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0016400. The authors thank the Joint Genome Institute to assist in camelina genome and RNA sequencing experiments.

Designing Microbial Consortia with Defined Social Interactions

Wentao Kong,[1-2] David R. Meldgin,[3] James J. Collins,[4-6] and **Ting Lu**[1-3]

[1]Department of Bioengineering, 2Institute for Genomic Biology, and [3]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801; [4]Institute for Medical Engineering and Science, [5]Department of Biological Engineering, and [6]Synthetic Biology Center, Massachusetts Institute of Technology, Cambridge, MA 02142, USA

Project Goals: We developed social interaction engineering as a systematic strategy for the design and construction of synthetic microbial ecosystems.

Designer microbial consortia are communities of rationally designed, interacting microorganisms that are capable of producing desired behaviors. Compared to engineered isogenic populations, synthetic communities offer an increased degree of robustness for designed cellular functions and an expanded spectrum of functional programmability for complex tasks. They have also emerged as a promising engineering tool to manipulate microbiomes, which helps to realize the enormous potential of microbiomes for therapeutic, environmental, and agricultural purposes. However, despite increasing exciting proof-of-concept demonstrations, the utilization of such synthetic ecosystems is hampered by our limited ability in rapidly developing microbial ecosystems with desired temporal and spatial dynamics. Inspired by the facts that social interactions such as competition and cooperation are both ubiquitous and essential in microbial communities. findings, here we present a systematic framework to the design, construction and characterization of synthetic microbial communities, namely, social interaction programming that combines modular pathway reconfiguration with model creation. Specifically, we employed a modular pathway reconfiguration approach to create six distinct consortia whose dynamics is specified by their underlying interaction modes. Using a modular approach similar to our experimental construction, we also derived quantitative models that captured experimentally observed population patterns. We further showed that the models from two-strain consortia can be used to design and build three- and four-strain ecosystems with predictable behaviors. Together, we established social interaction engineering as an effective route for ecosystem programming.

References

1. W. Kong, D. Meldgin, J. Collins, and T. Lu, Designing microbial consortia with defined social interactions, Nature Chemical Biology, 14: 821–829 (2018).

Biological Design of *Lemnaceae* Aquatic Plants for Biodiesel Production

Evan Ernst,[1,2] James Birchler,[3] Eric Lam,[4] Jorg Schwender,[5] John Shanklin,[5] and **Robert A. Martienssen* (martiens@cshl.edu)**[1,2]

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; [2]Howard Hughes Medical Institute, Cold Spring Harbor, NY; [3]University of Missouri, Columbia, MO; [4]Rutgers University, New Brunswick, NJ; and [5]Brookhaven National Laboratory, Brookhaven, NY

Project Goals:

1. **Leveraging our transformation methods, we will develop a comprehensive toolset for genetic manipulation of *Lemnaceae*. We will establish CRISPR/Cas9 genome editing to complement our artificial miRNA silencing methods. We will construct artificial chromosomes in *Lemna minor* to potentiate whole pathway engineering.**
2. **Resting and over-wintering fronds have higher starch content than corn kernels, but the energy density of oil is more than twice that of starch. We will use regulatory network and metabolic flux modeling to re-engineer the carbon allocation pathways to optimize triacylglyceride (TAG).**
3. **We will use comparative genomics of multiple *Lemnaceae* genome sequences, an extensive living collection of global accessions, and systems network analysis to characterize gene expression networks underpinning developmental and environmental responses to maximize bioenergy products while preserving rapid biomass accumulation. Nutrient deprivation and CO_2 irrigation will be used to enhance yield.**

Lemnaceae species (commonly called duckweeds) are the world's smallest aquatic flowering plants. They have a much reduced morphology comprising leaf-like growing fronds, starch-filled resting fronds, and simple roots. *Lemnaceae* in optimal conditions have an exponential growth rate that can double the number of fronds in 30 hours and produce 64 grams of biomass per gram starting weight in a week, which is far beyond that of terrestrial crops such as corn (2.3 g/g /week), and unencumbered by secondary products such as lignin. *Lemnaceae* offer an attractive alternative to algae as biofuel feedstocks because of their robust growth in open ponds and the relative ease of harvesting dry material. Convenient metabolic labeling in culture makes Lemna a good system for pathway modeling and engineering, as nutrients are taken up from liquid growth media, and non-responsive stomata can utilize very high levels of atmospheric CO_2. Our goal is to divert a substantial portion of accumulated carbon from starch to oil metabolism in *Lemnaceae*, using resting fronds as the storage tissue. Clonal propagation, limited seed set, and variable chromosome number are shared with sugarcane and *Miscanthus*, and many of the design principles and technologies we develop will have applications in other energy crops.

Under prior support from DOE, the Shanklin Lab completed a survey of fatty acid and TAG composition across 30 *Lemnaceae* species, while the Schwender lab has constructed a constraint-based model of carbon flux. A reliable and rapid protocol for stable transformation of *Lemna minor* was published by the Martienssen and Shanklin labs, along with gene-knockdown by artificial miRNA. The Lam and Martienssen labs have contributed to the sequence and gene content of three Lemna genomes complete with chromosome structures, methylomes, small

RNA transcriptomes, and structural variant analysis across accessions. Current genome assemblies have yielded validated orthologs in all the major lipid biosynthesis pathways.

We have performed single molecule long-read genome sequencing of diploid *L. gibba* and allotetraploid *L. minor* clones, using Oxford Nanopore technology followed with Hi-C to link scaffolds into chromosome-scale assemblies. The 21 chromosomes of diploid *L. gibba* are highly colinear with each of the subgenomes of allotetraploid *L. minor*. A comparison with genome sequences of *S. polyrhiza* and *W. australiana* reveals similar gene content that is highly reduced compared to terrestrial monocots such as rice and Brachypodium. Whole methylome sequencing has shown a dramatic reduction in asymmetric cytosine methylation in *Lemna* spp., which is similar to *Spirodela polyrhiza* in this respect. Spirodela has a much-reduced retrotransposon content, which accounts for further reductions in symmetric CG methylation, while *L. minor* retains a similar retrotransposon content to other monocot genomes. Small RNA sequencing has revealed dramatic differences between the three genera consistent with known pathways of RNA directed DNA methylation. We have analyzed orthologous gene content across the *Lemnaceae*, 17 other monocots and 11 non-monocots, revealing variations that likely account for some of these differences, as well as for reduced morphology, clonal reproduction, and aquatic growth habit. With these tools at hand we will be able to more easily identify gene-regulatory bottlenecks limiting oil production.

Critically, we have already developed engineered *L. minor* exhibiting a significant increase in oil content, building on the successful engineering of sugarcane to achieve 2-5% leaf TAG in the Shanklin lab under ARPA-E support. Engineered lines include stable overexpressors of WRINKLED, DGAT and PDAT1, all exhibiting marked increases in TAG content. We are addressing expected growth defects in the lines by developing CRISPR/Cas9 knockouts targeting SDP1, and multigene overexpression lines including OLE1 which have proven to mitigate FA cytotoxicity in other systems. The groundwork for construction of artificial chromosomes and transgene stacking systems is being established by the introduction of a landing pad construct developed in the Birchler Lab.

Publications

1. Yan, Yiheng, Jason Candreva, Hai Shi, Evan Ernst, Robert Martienssen, Jorg Schwender, and John Shanklin. 2013. “Survey of the Total Fatty Acid and Triacylglycerol Composition and Content of 30 Duckweed Species and Cloning of a Δ6-Desaturase Responsible for the Production of γ-Linolenic and Stearidonic Acids in Lemna Gibba.” *BMC Plant Biology* 13 (1). BioMed Central:201.
2. Cantó-Pastor, A., A. Mollá-Morales, E. Ernst, W. Dahl, J. Zhai, Y. Yan, B. C. Meyers, J. Shanklin, and R. Martienssen. 2015. “Efficient Transformation and Artificial miRNA Gene Silencing in Lemna Minor.” *Plant Biology* 17 Suppl 1 (January). Wiley Online Library:59–65.
3. Michael, Todd P., Douglas Bryant, Ryan Gutierrez, Nikolai Borisjuk, Philomena Chu, Hanzhong Zhang, Jing Xia, et al. 2017. “Comprehensive Definition of Genome Features in Spirodela Polyrhiza by High-Depth Physical Mapping and Short-Read DNA Sequencing Strategies.” *The Plant Journal: For Cell and Molecular Biology* 89 (3). Wiley Online Library:617–35.

Funding for this project is provided by the DOE Office of Biological & Environmental Research (DE-SC0018244).

Using Gene Editing and an Accumulated Bioproduct as a Reporter for Genotypic and Phenotypic Heterogeneity in Growth-vs-Production for *Methylobacterium extorquens* Conversion of Aromatics to Butanol

Tomislav Ticak[1*] (tticak@uidaho.edu), Sergey Stolyar,[1] Ankur B. Dalia,[2] Andreas E. Vasdekis,[1] and **Christopher J. Marx**[1]

[1]University of Idaho, Moscow, ID; [2]Indiana University, Bloomington, IN

https://marxlab.org/doe-biosystems-project/

Project Goals: With a unique capacity to assay growth and production – for either a tremendous number of genotypes in a mixture, or for individual cells – we will provide an unprecedented view of the critical tradeoff between growth and production. This will be used to guide development of *M. extorquens* as a novel platform for conversion of methoxylated aromatics to butanol. We will accomplish this work through the following aims:

1. Engineer/evolve improved use of methoxylated aromatics in M. extorquens

2. Explore growth-vs-production tradeoffs for genetic and phenotypic variation in PHB production

3. Combine improvements in substrate use and production capacity

4. Exchange PHB synthesis for butanol synthesis to test best genotypes

Abstract. Lignin-derived compounds from plant biomass are amongst the most recalcitrant for microbial conversion. Hydrolysates contain a wide variety of aromatic molecules, and a particular issue with these molecules is that many of them are methoxylated: these methoxy groups are released as formaldehyde during degradation, which can overload the detoxification ability of standard heterotrophs. Methylotrophic bacteria, on the other hand, not only rapidly generate internal formaldehyde from oxidation of single-carbon compounds, like methanol, but also can oxidize it fast enough to prevent toxicity. In the course of an earlier DOE project, we discovered that some *Methylobacterium* strains grow exceptionally well on aromatics, and do not release formaldehyde into the medium from the methoxy groups present, unlike classic systems for aromatic degradation (*e.g.*, *Pseudomonas putida*). We have since demonstrated that the pathways for methoxylated aromatic use can be introduced into the emerging model organism, *Methylobacterium extorquens*, and enable it to grow on aromatics. Over a short period of experimental evolution, strains of *M. extorquens* PA1 now grow quite well on protochatechuic acid via a pathway containing a novel beta-cleavage enzyme. Furthermore, introduction of genes encoding vanillic acid use into this background now permit growth on this methoxlated aromatic.

Aromatics are advantageous as a feedstock due to the fact that they are degraded via acetoacetyl-CoA, which links well to this organism's unique glyoxylate-regeneration pathway and therefore can serve as a pipeline for generating reduced products such as butanol.

The goal of this project is to develop *M. extorquens* as a catalyst to convert methoxylated aromatics from lignin hydrolysate into a model bioproduct, 1-butanol. In order to generate variation, we will apply a combinatorial genetic approach to alter the expression level of genes involved in these pathways via promoter swaps, which will generate $>10^4$ variants. We can assess the effect of each combination of edits upon both growth and production while in a complex mixture using deep-sequencing. To do this, we take advantage of an internally-accumulated compound, poly-β-hydroxybutyrate (PHB), as a reporter for the ability to generate butanol, because both PHB and butanol are produced along the same pathway. This approach permits linkage of production to genotype: PHB levels can be sorted by fluorescence-activated cell sorting (FACS) following addition of a fluorescent dye. Altogether, we will be able to determine $>10^4$ growth and production phenotypes simultaneously, illuminating at a broad scale the pleiotropic consequences of each allelic combination.

Furthermore, we will take advantage of single-cell approaches and the ability to assay PHB in single cells, to examine phenotypic heterogeneity in both growth and production. We, and others, have observed tremendous differences in PHB production between cells in an isogenic population. This heterogeneity likely also occurs for secreted products like butanol, but currently this cannot be observed at the resolution of individuals. We hypothesize that the same physiological non-linearities that lead to cell-to-cell variability will also exist for mutations that affect the same processes. To our knowledge, however, there has never been a large-scale comparison between genotypic diversity (10^4 genotypes) and phenotypic heterogeneity (10^3 cells), for either growth or production.

We will combine two sets of genetic improvements in the same organism: improved utilization of methoxylated aromatics, and enhanced production of PHB without compromised growth. Finally, for the most successful of these genotypes, we will eliminate PHB synthase and insert genes encoding the two-enzyme pathway for butanol production. Our ultimate goals are thus to not only develop *M. extorquens* for conversion of methoxylated aromatics to butanol, but to develop a novel approach that combines the advantages of gene editing and deep-sequencing, an internally-accumulated product as a proxy, and an analysis of phenotypic heterogeneity for both growth and production. These conceptual advances could broadly revolutionize work in DOE-relevant biosystems design.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0019436.

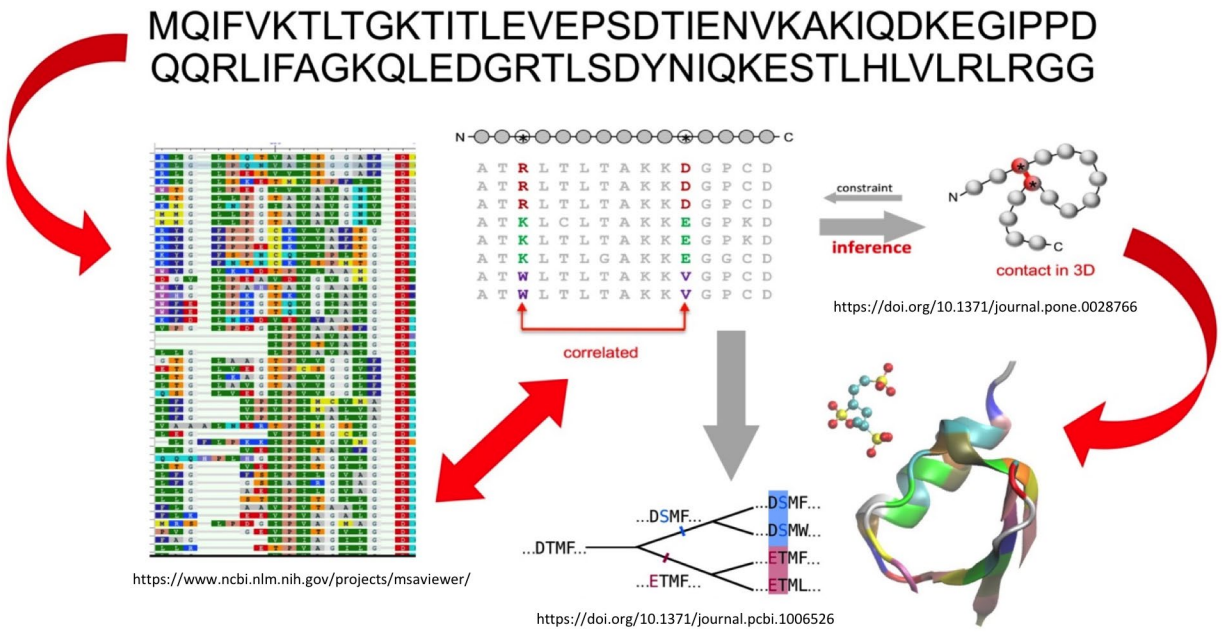
Genome to Structure: From Multiple Sequence Alignment to Virtual Ligand Screening Using Co-Evolutionary Protein-Residue Contact-Prediction in KBase

Ada Sedova[1]* (sedovaaa@ornl.gov), James G. Jeffryes[2], Loukas Petridis[1], Brian H Davison,[1] Christopher S. Henry[2], and **Julie Mitchell**[1]

[1]Oak Ridge National Laboratory, Oak Ridge, Tennessee
[2]Argonne National Laboratory, Argonne, Illinois

http://cmb.ornl.gov/index.php/research/bioenergy/dynamic-visualization-of-lignocellulose

Project Goals: The annotation of gene function benefits greatly from information on protein structure and binding. In this regard we have proposed the development of new functionality and workflows for KBase, in support of the Dynamic Visualization of Biological Structures SFA at Oak Ridge National Laboratory. The new structural branch of KBase includes tools to perform hidden-Markov model-based sequence searches with HHSuite[1], and CCMPredPy/CCMGen[2,3] tools to perform coevolutionary analysis of a protein sequence against metagenomic data, in order to predict residue contacts. These contact maps can be used to create phylogenetic trees, refine an existing multiple sequence alignment, and as restraints for use in protein folding algorithms. Finally, a protein-ligand docking workflow will allow for the use of the AutodockVINA[4] software to screen databases of metabolites, connecting metabolic-level function via physical molecular interactions to genomic information.



The annotation of gene function benefits greatly from information on protein structure and binding. In this regard we are developing a new structural branch for KBase, in support of the Biofuels SFA at Oak Ridge National Laboratory. The first workflow in this structural-biology

functionality starts from a protein sequence and uses HHSuite to perform a Multiple Sequence Alignment (MSA), followed by CCMPredPy to perform coevolutionary analysis to predict residue contact maps. These contacts can then be used as restraints in protein folding algorithms to generate a three-dimensional structural model for the protein from sequence. Information from the contact map will also provide value independent of any structure prediction, as the results indicate amino-acid residues that drive folding stability which is useful for experimental strategies in protein redesign and can also be used to refine the MSA and to create a phylogenetic tree. Several apps will be made available in KBase to utilize the HHSuite tools and the CCMPredPy/CCMPredGen tools. A second workflow will allow for the upload of a model protein structure or download of an experimental structure from the RCSB Protein Data Bank, followed by protein-ligand docking. The Ligand Screening App will use the AutodockVINA protein-ligand docking software to allow the protein structure to be screened against a set of metabolites. A potential use of this app will be in functional annotation of uncharacterized proteins: while sequence-based bioinformatics approaches have helped in annotating many genes to date, a structure-based treatment will facilitate additional insights into protein function that are obscured by lack of detectable sequence homology. In particular, for newly discovered enzymes, structural similarity along with ligand screening will help researchers anticipate likely functions. Along with the three KBase apps, we will develop sample narratives to illustrate their use.

Lignocellulosic biomass is a complex substrate that requires the synergistic action of a variety of enzymes for its efficient deconstruction. Biomass pretreatment generates byproducts, including solubilized lignin-derived aromatics, that inhibit enzymatic hydrolysis of cellulose[5]. Which bioproducts are formed depends on the biomass feedstock as well as the details of the pretreatment process. Applying the Ligand Screening App to predict which specific byproducts affect which particular enzymes can lead to an optimal selection of cellulolytic enzyme cocktails that minimize inhibition. A further barrier to biofuels and bioproduct production is that fermentation products, pretreatment solvents and byproducts can be toxic to microorganisms. We envision using CCMPred to predict the structures of novel proteins of any newly-sequenced microorganism and then docking solvents/byproducts on all proteins. Determining which proteins, and which protein residues, the small molecules bind to may lead to rational genetic engineering of those proteins and to microbes exhibiting improved tolerance to toxic pretreatment byproducts and solvents.

References

1) Söding J. Protein homology detection by HMM–HMM comparison. Bioinformatics. 2004 Nov 5;21(7):951-60.

2) Vorberg S, Seemayer S, Soeding J. Synthetic protein alignments by CCMgen quantify noise in residue-residue contact prediction. bioRxiv. 2018 Jan 1:344333.

3) Seemayer S, Gruber M, Söding J. CCMpred—fast and precise prediction of protein residue–residue contacts from correlated mutations. Bioinformatics. 2014 Jul 26;30(21):3128-30.

4) Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry. 2010 Jan 30;31(2):455-61.

5) Ximenes E, Kim Y, Mosier N, Dien B, Ladisch M. Deactivation of cellulases by phenols. Enzyme and microbial technology. 2011 Jan 5;48(1):54-60.

Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under contract no. DE-AC05-00OR22725. This program is supported by the U. S. Department of Energy, Office of Science, through the Genomic Science Program, Office of Biological and Environmental Research, under FWP ERKP917.

Evaluating Biogeochemical Processes Within Microenvironments Along the Root-rhizosphere-soil Continuum

James Moran[1]* (james.moran@pnnl.gov), Vivian Lin[1], Elizabeth Denis[1], Peter Ilhardt[1], Jamie Nuñez[1], Nicholas Huggett[1], Ryan Renslow[1], Timothy Linley[1], Mary Lipton[1]

[1]Pacific Northwest National Laboratory, Richland, Washington

https://science.pnl.gov/staff/staff_info.asp?staff_num=8559

Project Goals: This project seeks to elucidate key microbiological and geochemical controls on nutrient exchange within the rhizosphere and the role that spatial organization within the root-rhizosphere-soil continuum plays in directing nutrient acquisition by the host plant. Spatially-resolved understanding of nutrient exchange through this dynamic zone will identify key variables that may form part of an effective rhizosphere management program targeting enhanced plant productivity. Our aims are directed towards identifying the microbial and geochemical factors that stimulate enhanced plant investment (in the form of root exudation) into specific regions of the rhizosphere and assessing the implications of this carbon input on the microbial and geochemical response.

We hypothesize that localized regions within the rhizosphere act as foci for exchanging root-derived organic carbon with inorganic nutrients made available by a combination of soil microbial activity and inherent soil resource availability. Further, we hypothesize that the locations of these resulting nutrient exchange hotspots are not stochastically distributed throughout the rhizosphere but are controlled by microenvironmental conditions resulting from a combination of plant-derived carbon, microbiological activity, and soil geochemistry. To test these hypotheses, we are applying a suite of tools to evaluate the rhizosphere within a series of microcosms constructed with natural soil (Kellogg Biological Station, Hickory Corners, Michigan, USA) and switchgrass seedlings (variety Cave-in-Rock). We are specifically evaluating spatial heterogeneity in 1) root exudation, 2) microbial activity, and 3) soil geochemistry.

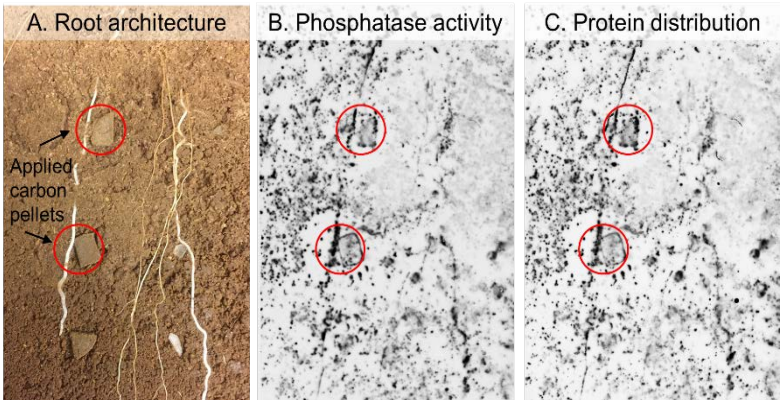


Figure 1: (A) Homogenized and pelleted root biomass was placed within the rooting zone of our soil microcosms. Samples of the mobile proteome were extracted, while maintaining their spatial localization, using a membrane transfer technique and then sequentially stained for (B) phosphatase activity and (C) total protein content. Increased protein content in areas of proximity between root and the applied resource islands suggest stimulation of overall biologic activity in these areas versus control regions: applied resource island with no root exposure, applied quartz (versus organic) inserts, and bulk soil.

Root exudation can provide a valuable carbon resource to subsurface environments which are frequently limited in this key nutrient. We are using a $^{13}CO_2$ tracer combined with laser ablation-isotope ratio mass spectrometry (LA-IRMS) to track variable rates of photosynthate flow into different roots and subsequently into the rhizosphere. We can clearly identify increased allocation of fresh photosynthate to specific roots over

others and are actively seeking to identify factors controlling the observed carbon distribution. We are also evaluating the spatial extent of root exudates into soil to identify locations of the rhizosphere that experience greater carbon investment by the host plant.

To help better understand the localized microbial response to root exudation, we are developing two methods to evaluate the microbial components of the system including 1) spatially resolved proteomics assays and 2) selective activity-based staining of specific enzymatic functions within the system. Our proteomic technique involves transferring mobile phase proteins (mainly exoproteins) onto a membrane while maintaining the native spatial distribution of the proteins. This technique is non-destructive to the system and enables timeseries analysis of the microbial community. Our enzymatic assays are designed to complement this approach to specifically map phosphatase activity onto the spatial distribution of proteins. In an initial experiment, we pelleted homogenized root biomass and dispersed this material into the root zone of our microcosms (Figure 1). The resulting images suggest there is both higher protein production and increased phosphatase activity where the root and the applied resource islands are in spatial proximity. We are working to identify how microbial diversity and protein expression may be stimulated by the combination of root exudation and bioavailable carbon.

Finally, in order to better characterize the geochemical microenvironment within and surrounding the rhizosphere, we developed a laser-induced breakdown spectroscopy (LIBS) technique to enable mapping of macro- and micro-nutrients in the soil and demonstrated its ability to identify specific elemental foci that may support hotspots of microbial activity (Ilhardt et al., 2019). We developed a quantitative image analysis package to identify gradients of nutrient concentration, such as carbon, calcium, potassium, phosphorus, and iron, at increasing distance from a root (Figure 2). Our ongoing work is focused on superimposing elemental gradients with microbial diversity and activity maps within the rhizosphere.

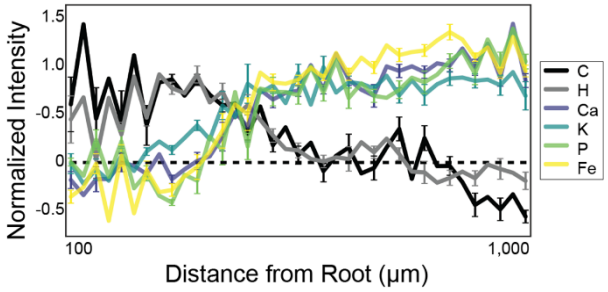


Figure 2: Measuring gradients of key nutrients with increasing distance from a root using LIBS (Ilhardt et al., 2019).

Overall, our developments allow us to track photosynthate into the rhizosphere and surrounding soil with high spatial resolution, and subsequently characterize the elemental and microbial composition of specific locations. Together, this data will reveal how soil geochemical microenvironments and microbial activity relate to the distribution of fresh photosynthate provided by the host plant.

References

1. Ilhardt, P., J. Nuñez, E. Denis, J. Rosnow, E.J. Krogstad, R. Renslow, J. Moran (2019) Rapid, high-resolution elemental mapping of the root-rhizosphere-soil continuum using laser-induced breakdown spectroscopy (LIBS). *Soil Biology and Biochemistry*, **131**, 119-132. doi: 10.1016/j.soilbio.2018.12.029

This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER). This contribution originates from an Early Career Research Award granted at the Pacific Northwest National Laboratory (PNNL).

Influence of Biogeochemistry and Engineering Controls on Microbial Growth, Membrane Features, and Interactions with Shale Matrices in Engineered Energy Systems

Jenna Luek,[1] Jia Xian Siew,[1] Brandon McAdams[2], J. Alexandra Hakala[2], and **Paula J Mouser**[1*] (Paula.Mouser@unh.edu)

[1]University of New Hampshire, Durham; [2]National Energy Technology Laboratory, Pittsburgh, PA

Project Goals: Humans are currently engineering the deep subsurface through energy-capture technologies, energy-storage capabilities, and environmental mitigation strategies. Despite the potential for microorganisms to alter the effectiveness of these techniques, we lack a clear understanding of how microbial cells interact with fluids and solids under deep subsurface conditions. The overarching goals of this research are to *1) advance our understanding of the microscopic complexity of nonsterile engineered hydrocarbon extraction systems as they evolve into mature use and 2) build new partnerships between UNH researchers and DOE National Laboratories*. Microbial membranes act as the first line of defense in adaptation to environmental conditions and are responsible for critical cellular functions, including matrix interactions. The specific project objectives are to **1) characterize variables influencing growth parameters and membrane features of fractured shale taxa, 2) characterize the interactions between shale matrices and microorganisms and 3) elucidate engineered and environmental processes driving biogeochemical signatures at the field scale.**

Abstract: During the first few months of this project, post-doctoral researcher (Luek) and graduate student (Siew) have focused their efforts on characterizing variables influencing growth and membrane features of the Utica shale derived taxon *Halanaerobium* (Objective 1). Due to the sensitivity of these bacteria to changes in pH and metabolite accumulation, and in order to determine growth parameters under shale-relevant conditions, we are using the Sartorius Biostat Q system continuous culture (chemostat) system to control for pH, temperature, media flow rate/growth rate, and gas purge rate in six replicate anaerobic bioreactors. Once we optimize parameters to ensure chemostat stability, we will collect cell and media for intact polar lipids methods testing and preliminary analysis of proteomics, lipidomics, and metabolites at the Environmental Molecular Sciences Laboratory (EMSL).

In addition to cultivation efforts, we have made inroads toward optimizing methods for detection and quantification of microbial intact polar lipids (IPLs) using an LC-qTOF-MS accessed at the National Energy Technology Laboratory (NETL) in Pittsburgh, PA during a visit by Luek in early January 2019. She worked with NETL post-doctoral researcher (McAdams) to train and perform experiments on NETL's LC-qTOF-MS, testing an IPL-specific HILIC method to determine the sensitivity and retention times for four lipid standards (phosphoethanolamine, ubiquitin, phosphatidylglycerol, and cardiolipin). Retention times for two of these standards could be directly compared to published research using this HILIC method and were found to be comparable, indicating the tested HILIC method would be appropriate for a wide range of IPLs. Luek will visit NETL again in the coming months to determine detection limits for standards, assess the fragmentation patterns for IPLs, develop a parallel reverse-phase LC-qTOF-MS method for more nonpolar lipids, and test the methods on lipid extracts from cultured

Halanaerobium and/or the field. Fatty acid standards have been purchased to begin methods development for polyunsaturated fatty acids using the GC-MS at UNH in March

In parallel with laboratory efforts, we are working with West Virginia University and Northeast Natural Energy to prepare for field sampling efforts at MSEEL II, a DOE NETL funded hydraulic fracturing research site. Drilling and sidewall coring is scheduled at the site for March 2019 and well completion (fracturing) is slated for May. We will collect core, fluid, and other drill media throughout these upcoming field efforts for characterization of membranes and/or use in future experiments.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (BER) and the Established Program to Stimulate Competitive Research (EPSCoR) program under Award Number DESC0019444.

Immune-suppressing pattern recognition receptors mediate host-driven recruitment of microbes

Carly Shanks[1], Timothy B. Yates[1], Kai Feng[1], Colin Brislawn[2], Lye Meng Markillie[2], Hugh D. Mitchell[2], Jing Hou[1], Sara Jawdy[1], Lee Gunter[1], Steven Lebreux[1], Jin Zhang[1], Stephen P. DiFazio[3], Lawrence Smart[4], **Wellington Muchero[1] (mucherow@ornl.gov)**

[1] Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [2] Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA; [3] Department of Biology, West Virginia University, Morgantown, WV; [4] Horticulture department, Cornell University, Geneva, NY

Project Goals: Identify and characterize species-divergence of immune-suppressing pattern recognition receptors in *Salix spp*.

Unlike models which posit that microbe-associated molecular patterns (MAMPs) are responsible for evasion of innate host immunity during colonization by microbes, we provide evidence of active host-driven recruitment of microbes by specific membrane-bound pattern recognition receptors (PRRs), which recognize carbohydrate moieties in microbial cell walls leading to global suppression of the host's own immune response to facilitate colonization. A hallmark feature of these PRRs is that they have conserved protein domain architecture which includes lectin-binding, S-locus, Plasminogen-Apple-Nematode (PAN), transmembrane (TM) and kinase domains. We hypothesize that upon recognition of microbial MAMPs, the intracellular kinase domain initiates signaling cascades which results in suppression of host defense mechanism to facilitate colonization. However, there is currently no knowledge of which specific domain is responsible for suppression of host immunity to facilitate invasion by microbes. Using comparative genomics across seven *Salix* species (, *S. purpurea*, *S. viminalis*, *S. udensis*, *S. integra*, *S. koriyanagi*, *S. alberti*, and *S. suchowensis*) coupled with molecular genetic approaches including heterologous transgenesis, transient assays and transcriptome profiling, we provide evidence that a key domain encompassing twenty amino acids is essential for suppression of host defense mechanisms. Mutating as few as six of these amino acids resulted in successful induction of defense response based on observed increases in expression of canonical defense marker-genes, including WRKY40, WRKY72 and NPR1. Implications of genomic variation in this domain across the *Salix* species will be illustrated by differential abundance of microbial symbionts in the root tissue as determined by 16S and fungal ITS sequencing. This discovery provides opportunities for engineering novel symbiotic interactions for sustainable production of plant-based industrial feedstocks.

Acknowledgement: This work was funded by the Department of Energy Office of Science under the Early Career Research Program.

Comparative Genomic and Transcriptomic Analyses for Pathway Discovery in *Chromochloris zofingiensis*

Fatima Foflonker* (foflonker@bnl.gov), **Crysten E. Blaby-Haas**

Brookhaven National Laboratory, Upton, NY

Project Goals: Our overarching research goal is to design and engineer high-level production of biofuel precursors in photoautotrophic cells of the unicellular green *alga Chromochloris zofingiensis*. Our strategy involves large-scale multi-'omics systems analysis to understand the genomic basis for energy metabolism partitioning as a consequence of carbon source. Enabled by cutting-edge synthetic biology and genome-editing tools, we will integrate the systems data in a predictive model that will guide the redesign and engineering of metabolism in *C. zofingiensis*. Toward these objectives, we are implementing a phylogenomics-guided approach that leverages evolutionary relationships between genomes and between proteins encoded on those genomes for contextualized and evidence-based protein function discovery.

The development of 'omics technologies has provided an unprecedented opportunity for the study of organisms as complex systems. With the genome-wide data these technologies provide, nearly any species can be fast-tracked to a level of understanding that was previously attainable for only a few "model organisms". The present challenge is shifting from acquiring genomic data, such as whole-genome sequences and transcriptomes, to using that knowledge to enable predictive biology and the rational redesign of biosystems. At the core of this challenge is the bottleneck in protein function understanding. Every genome-wide analysis, such as an RNA-Seq experiment, relies on functional annotations for translating data into information. As a result, the ability of an 'omics dataset to inform on a biological system is dependent on knowledge of that system's parts. However, for emerging organisms, nearly every functional annotation is a prediction that is typically based solely on sequence similarity to a database hit. Approaches are needed that place genomic resources in the hands of experimentalists allowing them to verify and test those function predictions in silico before taking a hypothesis to the bench.

With the successful implementation of an integrated systems biology approach to model, design and engineer high levels of biofuel precursors and value-added products, microalgae have the potential to become major sources of sustainable bioenergy and bioproducts. Toward this goal, over 100 algal whole-genome sequences are either presently available or are soon to be published. As fundamental resources, these data combined with thousands of published transcriptomes are precipitating a paradigm shift in the way we understand one of the most diverse, complex and understudied groups of photosynthetic eukaryotes. Remarkably, over half of the proteins encoded by algal genomes are of unknown function, highlighting both the volume of unique functional capabilities yet to be discovered and a fundamental knowledge gap that impedes successful biosystem design.

We are using a phylogenomic-based approach supported by a comprehensive, large-scale systems analyses for the discovery of novel, economically valuable, functional capabilities in the unicellular green alga *Chromochloris zofingiensis*. We are utilizing comparative genomics approaches to infer protein function from evidence-based associations. Ten chlorophyte algal genomes were used to find conserved gene neighborhoods, defined as: proximal orthologous genes within a 5 gene window, in a minimum of 4 species, and from at least two taxonomic classes to minimize effects of background synteny. This resulted in 183 conserved gene neighborhoods with predicted functionality in carotenoid biosynthesis, photorespiration, thiamine metabolism, nitrogen recycling, oxidative stress responses, and detoxification. Furthermore, relaxed constraints were used to capture proximal orthologous genes that co-occur in *C. zofingiensis* and at least one other algal genome. Gene fusions were identified by searching for domains that are encoded by two separate genes in at least two genomes and encoded by a single gene in at least two other genomes. These analyses were combined with a phylogenetic profile of orthologous proteins and co-expression analysis of condition-specific RNA-Seq data for the discovery and support of co-functional proteins and potential bioengineering targets for the production of value-added bioproducts and redesign of carbon and nutrient handling in *C. zofingiensis*.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), award number DE-SC0018301.

Single-Molecule, Whole-Transcript Sequencing of the Transcriptome of the Green Microalga *Chromochloris zofingiensis* to Accurately Annotate Gene Models and Identify Splice Variants

Sean D. Gallaher[*1] (gallaher@chem.ucla.edu), and **Sabeeha S. Merchant**[2]

[*] Presented By, [1] University of California, Los Angeles, CA, [2] University of California, Berkeley, CA

Project Goals: Our overarching research goal is to design and engineer high-level production of biofuel precursors in photoautotrophic cells of the unicellular green alga *Chromochloris zofingiensis*. Our strategy involves large-scale multi-'omics systems analysis to understand the genomic basis for energy metabolism partitioning as a consequence of carbon source. Enabled by cutting-edge synthetic biology and genome-editing tools, we will integrate the systems data in a predictive model that will guide the redesign and engineering of metabolism in *C. zofingiensis*.

Despite having a high-quality genome for *C. zofingiensis*, the current gene annotations are riddled with errors, which is an impediment to our 'omics studies. Here, we have sequenced whole transcripts as single-molecule, long reads from a pool of *C. zofingiensis* mRNA. We are using this data to re-annotate the genome to correct errors, join fragmented genes, and identify splice variants. This should greatly benefit our efforts to conduct multi-'omics analyses in the species.

Chromochloris zofingiensis is a unicellular green alga that is of interest for its ability to produce high quantities of TAGs, as well as other high-value bio-products. In 2017, we published a high quality, chromosome-complete, assembly for *C. zofingiensis*' 58 Mbp genome [1]. As part of this work, we used the AUGUSTUS ab initio gene caller trained on a set of de novo assembled transcripts to annotate 15,274 nuclear genes. However, since that work we have identified a number of short-comings with those annotations. Many genes appear to be erroneously split into two, three, or more gene models. This has the consequence of making it difficult to accurately assign functional annotations to the fragmented genes, and complicates transcriptomic analysis. Many gene models appear to be wrong, specifically with regards to the intron/exon junctions. This leads to inaccurate predictions of the encoded polypeptide, which confounds proteomic analysis. In addition, the current annotations recognize only one splice variant per locus, despite evidence that *C. zofingiensis* utilizes alternative splicing. Lastly, the AUGUSTUS gene annotations lack UTRs. Collectively, these limitations and errors in the currently available AUGUSTUS gene models are obscuring the complete picture of *C. zofingiensis* gene expression that our research demands.

To advance our research in *C. zofingiensis*, we have used a relatively new method, called Iso-Seq, in which PacBio long reads are used to sequence whole transcripts as single molecules. This approach has the advantage over the Illumina short read sequencing that was done previously in that each sequencing read identifies a complete splice variant that is part of the transcriptome. When mapped back to the genome assembly, these reads help to mark the precise locations of intron/exon boundaries, and identify the UTRs. For *C. zofingiensis,* cDNA was made from a pool of RNA collected from cultures grown under a wide range of conditions (phototrophic growth, heterotrophic growth, nutrient-deprived, oxidative stress, etc.), in order to capture the widest possible range of transcripts. The resulting cDNA was ligated to adapters and sequenced on the PacBio Sequel platform. After being subjected to appropriate quality filters, the transcript sequences were re-mapped to the *C. zofingiensis* genome assembly, and compared with the AUGUSTUS gene models.

This analysis identified 26,529 unique transcript isoforms, which mapped to 11,305 genetic loci. The genes identified by Iso-Seq included 9,174 of the 15,274 nuclear genes identified by AUGUSTUS (~60%), as well as 2,131 novel genes. The remaining AUGUSTUS gene models were absent from the Iso-Seq dataset either because those transcripts were under-represented in the Iso-Seq library, or because those genes were originally misannotated by AUGUSTUS. For those genes identified by Iso-Seq, ~50% had two or more splice variants. Importantly, dozens of genes that we had identified as being fragmented in the AUGUSTUS gene models were correctly merged in the Iso-Seq data. While the Iso-Seq data alone is insufficient for complete annotation of all genes, careful merging of this new data with the previously available AUGUSTUS gene predictions and Illumina RNA-Seq data should allow us to produce a highly accurate and detailed picture of the *C. zofingiensis* transcriptome and proteome.

References

1. Roth MS, Cokus SJ, Gallaher SD, Walter A, Lopez D, Erickson E, et al. Chromosome-level genome assembly and transcriptome of the green alga Chromochloris zofingiensis illuminates astaxanthin production. Proc Natl Acad Sci U S A. 2017;114. doi:10.1073/pnas.1619928114/-/DCSupplemental

Funding statement.

Funding was generously provided by the DOE / BER grant # DE-SC0018301 and by the Facilities Integrating Collaboration for User Science (FICUS) project of the Joint Genome Institute (JGI) and the Environmental Molecular Sciences Laboratory (EMSL).

Lipidomics analysis of the emerging model green alga *Chromochloris zofingiensis*

Yuntao Hu[1,2]*(yuntaohu@lbl.gov), Melissa Roth[2], Katherine Louie[1,3], Daniel Westcott[2], Tim Jeffers[2], Benjamin P Bowen[1,3], **Trent Northen**[1,3], and Krishna K. Niyogi[2]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA; [2]University of California, Berkeley, CA; [3]Joint Genome Institute (JGI), Walnut Creek, CA

Project Goals: Our overarching research goal is to design and engineer high-level production of biofuel precursors in photoautotrophic cells of the unicellular green alga *Chromochloris zofingiensis*. Our strategy involves using large-scale multi-'omics systems analysis to understand and model the genomic basis for how the energy metabolism of the cell is redirected partitioning based on the carbon source as a consequence of carbon source. Enabled by cutting-edge synthetic biology and genome-editing tools, we will integrate the systems data in a predictive model that will guide us in the redesigning and engineering of the metabolism of in *C. zofingiensis*. This presentation will focus on the lipidomics analysis of the *C. zofingiensis* trophic transitions to help us understand the key metabolic pathways involved in lipids synthesis and degradation.

Abstract text: The emerging model green alga *Chromochloris zofingiensis* is a promising producer of lipids and astaxanthin, in which lipids can comprise up to 50% of its dry biomass. One of its unique characteristics is that its photosynthesis can be switched off under the supplement of additional carbon sources (e.g. glucose)[1]. The changes of the metabolism result in the decomposition of thylakoid membranes and accumulation of lipids, e.g. triacylglycerols (TAGs). The abundance of TAGs increased over 20-fold within a few days with glucose making it ideal for the extraction of lipids as the biofuel precursor. In addition, thylakoids are reassembled and photosynthesis is resumed when the carbon source is depleted. This unique metabolism switch phenomenon offers us a great opportunity to reveal the key metabolic pathways of the algal lipid synthesis.

To gain insights into this metabolic switch we are developing lipidomic approaches for analyzing the algal lipids using high performance liquid chromatography coupled to Q Orbitrap Exactive mass spectrometry. Critically, we are also developing the cheminformatic capabilities to analyze the complex lipid profiles. MetaboliteAtlas is our main software tool and it requires a manually defined "Lipid Atlas" with defined mass-to-charge ratios and retention times. To date we have characterized algal lipids, e.g. phosphatidylcholine, phosphatidylethanolamine, lysophosphatidylcholine, lysophosphatidylethanolamine, and glycerolipids, to create the requisite Lipid Atlas, which currently includes ~1300 lipids divided into more than 14 lipid subclasses. Once complete, we will use this approach to analyze a time series experiment of *C. zofingiensis* to observe its trophic transitions by the addition of glucose. The resulting information will be used to refine and develop predictive models that will provide new biological insights into how this important organism is able to re-route flux during trophic transitions.

Funding statement. This work was supported by U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under contract No: DE-SC0018301.

References

1. Roth MS, Gallaher SD, Westcott DJ, Iwai M, Louie KB, Mueller M, Walter A, Foflonker F, Bowen BP, Ataii NN, Song J, Chen J-H, Blaby-Haas CE, Larabell C, Auer M, Northen TR, Merchant SS, Niyogi KK (In press) Regulation of oxygenic photosynthesis during trophic transitions in the green alga *Chromochloris zofingiensis*. Plant Cell

Iron and Nitrogen Regulate Trophic Transitions and Metabolism in the Oleaginous Green Alga *Chromochloris zofingiensis*

Tim Jeffers[1,2]* (tim_jeffers@berkeley.edu), Ryan McCombs,[1,2] Gauri Kapse,[1] Daniel J. Westcott,[1,2] Stefan Schmollinger,[3,4] Sabeeha S. Merchant,[3,4] Krishna K. Niyogi,[1,2] **Melissa S. Roth[1,2]**

[1]University of California, Berkeley, CA [2]Howard Hughes Medical Institute, Berkeley, CA; [3]University of California, Los Angeles, CA; [4]Present address: University of California, Berkeley, CA

Project Goals: Our overarching research goal is to design and engineer high-level production of biofuel precursors in photoautotrophic cells of the unicellular green alga *Chromochloris zofingiensis*. Our strategy involves using large-scale multi-'omics systems analysis to understand and model the genomic basis for how the energy metabolism of the cell is redirected partitioning based on the carbon source. Enabled by cutting-edge synthetic biology and genome-editing tools, we will integrate the systems data in a predictive model that will guide us in the redesigning and engineering of the metabolism in *C. zofingiensis*.

Unlike the well-studied model organism *Chlamydomonas reinhardtii, C. zofingiensis* can use external glucose as a sole carbon source. Here we show that interplay of glucose with iron and nitrogen concentrations drastically changes the metabolism of this organism. In iron-depleted medium with glucose in the light, *Chromochloris zofingiensis* shuts off photosynthesis to become fully heterotrophic (1), whereas replete iron rescues photosynthesis even in the presence of glucose. Furthermore, iron and nitrogen depletion in glucose medium results in production of the valuable carotenoid astaxanthin and likely the biofuel precursor triacylglycerol. To aid the study of glucose responses and the interplay of glucose with other essential nutrients, we have designed an optimized growth medium to enhance photoautotrophic growth and compensate for increased nutrient losses during growth on glucose. This defined medium will serve as a strong reference point for time course and systems analyses of the molecular responses underlying trophic transitions and accumulation of bioproducts. Through forward genetic screens on glucose, we have also generated mutants that do not trigger expected responses to iron, glucose, or nitrogen. These include mutants related to regulation of photosynthesis by iron and glucose and mutants that accumulate astaxanthin in non-inducing conditions. In total, this work positions *C. zofingiensis* as a platform for identifying the genes underlying the nutrient regulation of photosynthesis and metabolism, which will provide new targets for engineering algae for enhanced bioproduction.

References

1. Roth MS, Gallaher SD, Westcott DJ, Iwai M, Louie KB, Mueller M, Walter A, Foflonker F, Bowen BP, Ataii NN, Song J, Chen J-H, Blaby-Haas CE, Larabell C, Auer M, Northen TR, Merchant SS, Niyogi KK (2019) Regulation of oxygenic photosynthesis during trophic transitions in the green alga *Chromochloris zofingiensis*. *Plant Cell*, in press.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under Award Number DE-SC0018301.

Genome-scale Metabolic Model of *Chromochloris zofingiensis*, an Emerging Model Organism for Sustainable Fuel Production

Alexander Metcalf[1]* (metcalf@mines.edu), Walter Prentice[1], Alex Ramsey[1], Anthony Nagygyor[1], and **Nanette Boyle**[1]

[1]Chemical and Biological Engineering, Colorado School of Mines, Golden, CO 80401, USA

Project Goals: Our overarching research goal is to design and engineer high-level production of biofuel precursors in photoautotrophic cells of the unicellular green alga *Chromochloris zofingiensis*. Our strategy involves large-scale multi-'omics systems analysis to understand the genomic basis for energy metabolism partitioning as a consequence of carbon source. Enabled by cutting-edge synthetic biology and genome-editing tools, we will integrate the systems data in a predictive model that will guide the redesign and engineering of metabolism in *C. zofingiensis*. The Boyle laboratory is tasked with developing and utilizing a genome scale metabolic network reconstruction to predict intracellular carbon fluxes which will then be compared to fluxes measured experimentally using ^{13}C-MFA.

C. zofingiensis is an emerging model system for the production of biofuels and bioproducts. It is an especially attractive system attractive because it produces astaxanthin along with a large amount of lipids. Astaxanthin is a high value product (~$7,000 per kilogram) that has uses in the pharmaceutical, nutraceutical, and cosmetic industries[1-3]. In order to investigate the metabolic capacity of this organism for both fuel and astaxanthin production, we generated a genome-scale metabolic network reconstruction. The current reconstruction includes 3522 metabolic reactions and 2880 metabolites. In order to formulate an accurate biomass formation equation, we are also measuring both the macromolecule composition of *C. zofingiensis* (DNA, RNA, protein, lipid, carbohydrate) as well as the composition of each in photoautotrophic and photoheterotrophic growth modes. Predicted carbon flux distributions for each growth mode will be presented.

To enable faster reconstruction efforts of new organisms in the future, we also developed an automated reconstruction algorithm specifically designed for photosynthetic microorganisms. This approach leverages the manual curation efforts of published genome scale network reconstructions to minimize duplicate efforts for well-characterized pathways. We will present our algorithm, Rapid Annotation of Photosynthetic Systems (RAPS), and discuss the performance of automated reconstruction compared to other automated algorithms.

References

1. Hussein, G., Sankawa, U., Goto, H., Matsumoto, K. & Watanabe, H. Astaxanthin, a Carotenoid with Potential in Human Health and Nutrition. *Journal of Natural Products* **69**, 443-449, doi:10.1021/np050354+ (2006).
2. Yuan, J. P., Peng, J., Yin, K. & Wang, J. H. Potential health-promoting effects of astaxanthin: A high-value carotenoid mostly from microalgae. *Molecular Nutrition & Food Research* **55**, 150-165, doi:doi:10.1002/mnfr.201000414 (2011).
3. Liu, J. *et al.* Chlorella zofingiensis as an Alternative Microalgal Producer of Astaxanthin: Biology and Industrial Potential. *Marine Drugs* **12**, 3487-3515, doi:10.3390/md12063487 (2014).

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0018301

A Molecular Switch for Oxygenic Photosynthesis and Metabolism in the Emerging Oleaginous Model Green Alga *Chromochloris zofingiensis*

Melissa S. Roth[1,2]* (mroth@berkeley.edu), Daniel J. Westcott[1,2,a], Sean D. Gallaher[3,a], Masakazu Iwai[1,2], Katherine B. Louie[4,5], Maria Mueller[4], Andreas Walter[4,6], Fatima Foflonker[7], Benjamin P. Bowen[4,5], Nassim N. Ataii[4], Junha Song[4], Jian-Hua Chen[4,6], Crysten E. Blaby-Haas[7], Carolyn Larabell[4,6], Manfred Auer[4], Trent R. Northen[4,5], Sabeeha S. Merchant[3,8], Krishna K. Niyogi[1,2]

[1]Howard Hughes Medical Institute, Berkeley, CA; [2]University of California, Berkeley, CA; [3]University of California, Los Angeles, CA; [4]Lawrence Berkeley National Laboratory, Berkeley, CA; [5]DOE Joint Genome Institute, Walnut Creek, CA; [6]University of California, San Francisco, CA; [7]Brookhaven National Laboratory, Upton, NY; [8]Present address: University of California, Berkeley, CA; [a]equal contribution

Project Goals: Our overarching research goal is to design and engineer high-level production of biofuel precursors in photoautotrophic cells of the unicellular green alga *Chromochloris zofingiensis*. Our strategy involves using large-scale multi-'omics systems analysis to understand and model the genomic basis for how the energy metabolism of the cell is redirected partitioning based on the carbon source. Enabled by cutting-edge synthetic biology and genome-editing tools, we will integrate the systems data in a predictive model that will guide us in the redesigning and engineering of the metabolism in *C. zofingiensis*. This presentation focuses on understanding regulation of photosynthesis and metabolism, which will enable bioengineering of microalgae for improved production of biofuels and bioproducts.

Abstract:
Microalgae have the potential to become a major source of biofuels and bioproducts without exacerbating environmental problems. Photosynthetic microbes can utilize solar energy, grow quickly, consume CO_2, and be cultivated on non-arable land. However, there are presently considerable practical limitations in the photosynthetic production of biofuels from microalgae, resulting in low productivity and high costs. Insight into regulation of photosynthesis and metabolism will enable bioengineering of microalgae to maximize production of biofuels and bioproducts.

In this study, we show that the emerging oleaginous model green alga *Chromochloris zofingiensis* has a glucose-dependent photosynthetic and metabolic switch that includes rapid reversible changes in photosynthesis, the photosynthetic apparatus, thylakoid ultrastructure and energy stores including lipids and starch (1). With glucose in the light, *C. zofingiensis* shuts off photosynthesis and accumulates large amounts of commercially relevant bioproducts including

triacylglycerols (TAGs) and the high-value nutraceutical ketocarotenoid astaxanthin, while increasing culture biomass. We used our recently published high-quality transcriptome (2) to show reversible, specific, coordinated gene expression changes in carotenoid, photosynthetic and metabolic pathways. Moreover, we used forward genetics and our chromosome-level genome assembly (2) to reveal that this photosynthetic and metabolic switch is mediated by the glycolytic enzyme hexokinase (HXK1), which likely functions as a glucose sensor that regulates algal photosynthesis, astaxanthin synthesis, and carbon metabolism (3). Our data suggest that HXK1 initiates a signaling cascade resulting in repression of photosynthetic genes. Sugars play fundamental regulatory roles in gene expression, physiology, metabolism, and growth in plants and animals, and we introduce a relatively simple, emerging model system to investigate conserved eukaryotic sugar sensing and signaling at the base of the green lineage. Advancing the basic science underlying regulatory 'switches' between growth and lipid production will likely by critical to developing economically viable algal biofuels.

References

1. Roth MS, Gallaher SD, Westcott DJ, Iwai M, Louie KB, Mueller M, Walter A, Foflonker F, Bowen BP, Ataii NN, Song J, Chen J-H, Blaby-Haas CE, Larabell C, Auer M, Northen TR, Merchant SS, Niyogi KK (2019) Regulation of oxygenic photosynthesis during trophic transitions in the green alga *Chromochloris zofingiensis*. *Plant Cell*, in press.
2. Roth MS, Cokus SJ, Gallaher SD, Walter A, Lopez D, Erickson E, Endelman B, Westcott D, Larabell C, Merchant SS, Pellegrini M, Niyogi KK (2017) Chromosome-level genome assembly and transcriptome of the green alga *Chromochloris zofingiensis* illuminates astaxanthin production. *Proc Natl Acad Sci USA* 114:E4296–E4305.
3. Roth MS, Westcott DJ, Iwai M, Niyogi KK (In revision) Hexokinase functions as a molecular switch for oxygenic photosynthesis and metabolism in a green alga. Manuscript in revision.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under Award Number DE-SC0018301.

Advancing the *Chromochloris zofingiensis* molecular toolkit

Daniel J. Westcott[1,2,*] (westcott@berkeley.edu), Krishna K. Niyogi[1,2] , **Melissa S. Roth[1,2]**

[1]Howard Hughes Medical Institute, Berkeley, CA; [2]University of California, Berkeley, CA

Project Goals:

Our overarching research goal is to design and engineer high-level production of biofuel precursors in photoautotrophic cells of the unicellular green alga *Chromochloris zofingiensis*. Our strategy involves using large-scale multi-'omics systems analysis to understand and model the genomic basis for how the energy metabolism of the cell is redirected partitioning based on the carbon source. Enabled by cutting-edge synthetic biology and genome-editing tools, we will integrate the systems data in a predictive model that will guide us in the redesigning and engineering of the metabolism in *C. zofingiensis*. This poster details recent advances our group has made in developing genetic engineering and lipid analysis tools for this emerging model organism.

Abstract:

The green alga *Chromochloris zofingiensis* has a natural ability to accumulate high amounts of energy-dense lipids and pigments in a concentrated culture. Recent and ongoing investigation of this organism has revealed key signaling and biochemical hubs in the lipid biosynthesis pathway (1,2,3). The understanding and manipulation of pathway elements require the use of modern molecular genetics tools, none of which have been optimized for this organism. Here we report on the delivery of recombinant DNA by electroporation and *Agrobacterium*-mediated transformation and on our efforts to deliver proteins by electroporation. In addition, we have adapted a relatively simple lipid extraction protocol that, coupled with TLC, provides a quick and detailed analysis of the lipid profile of *C. zofingiensis*. A pipeline combining our established UV mutagenesis protocol and TLC provides a powerful forward genetics screen for strains with aberrant lipid profiles. Of particular interest will be mutants with increased TAG accumulation or those with more efficient lipid extraction. Concurrently, we are adapting genome editing protocols for use in *C. zofingiensis* with the intention of disrupting specific genetic loci to increase lipid accumulation or facilitate extraction. Potential targets identified by RNA-Seq analysis will be discussed.

References

1. Roth MS, Cokus SJ, Gallaher SD, Walter A, Lopez D, Erickson E, Endelman B, Westcott D, Larabell C, Merchant SS, Pellegrini M, Niyogi KK (2017) Chromosome-level genome assembly and transcriptome of the green alga *Chromochloris zofingiensis* illuminates astaxanthin production. *Proc Natl Acad Sci USA* 114:E4296–E4305.

2. Roth MS, Gallaher SD, Westcott DJ, Iwai M, Louie KB, Mueller M, Walter A, Foflonker F, Bowen BP, Ataii NN, Song J, Chen J-H, Blaby-Haas CE, Larabell C, Auer M, Northen TR, Merchant SS, Niyogi KK (2019) Regulation of oxygenic photosynthesis during trophic transitions in the green alga *Chromochloris zofingiensis*. *Plant Cell*, in press.
3. Roth MS, Westcott DJ, Iwai M, Niyogi KK Hexokinase functions as a molecular switch for oxygenic photosynthesis and metabolism in a green alga. Manuscript in revision.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under Award Number DE-SC0018301.

The dynamic responses of biological soil crust communities revealed by wetting experiments coupled with metagenomics

Marc W. Van Goethem[1]* (mwvangoethem@lbl.gov), Tami L. Swenson,[1] Estelle Couradeau,[1] Danielle Goudeau,[2] Nandita Nath,[2] Rex Malmstrom,[2] and **Trent R. Northen**[1,2]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA; [2]DOE Joint Genome Institute, Walnut Creek, CA

http://www.northenlab.org/research/environmental-exometabolomics/

Project Goals: Understanding the contributions of microorganisms to soil nutrient cycling and organic matter turnover are priorities of the Department of Energy missions. Within this framework, our projects aim to resolve *in situ* microbial activity through the combination of pioneering approaches that link microbial community structure and functional attributes within critical desert biocrust communities.

Abstract:

The frequency and intensity of rainfall events in dryland ecosystems is predicted to become more variable against the milieu of a warming climate scenario. The southwestern USA is predicted to experience alterations in the timing of rainfall events with concomitant increases in soil surface temperatures. In arid lands, precipitation pulses are rare, yet their timing, intensity and duration dictate biological activity. A significant avenue of recent microbial ecology research has been to elucidate the consequences of climatic change on microbial communities and how these responses in turn affect critical ecosystem processes such as organic matter turnover. Biological soil crust communities, also referred to as biocrusts, will be some of the most impacted systems, and serve as ideal models to explore the effects of changes in rainfall.

Biocrust communities are dominant covers in deserts and occupy more than 14% of the Earth's terrestrial surface. Biocrusts are typically comprised of cyanobacteria, mosses, lichens, fungi, and heterotrophic bacteria that together colonize and stabilize the soil surface. Moreover, biocrusts provide critical functions in deserts by controlling water retention, fixing both carbon (C) and nitrogen (N), and stabilizing surface sediment. The activity of biocrusts hinges upon hydration pulses, and these moisture inputs also impart challenges to dehydrated communities that face rapidly changing osmotic pressure and surface microtopology. We sought to track the responses of biocrust community traits to a hydration event using shotgun metagenomics. We compared biocrust communities that represent a successional maturity gradient and found that the effect of wetting on microbial community structure and functional potential is greater than that of ecological succession.

We observed a dramatic increase in the proportion of *Alphaproteobacteria*, which became central to food web structure by the end of the wet-up. The upturn in gene abundance for these bacteria was linked to significant increases in their ribosomal copy number, as well as genes for sporulation and DNA replication. Thirty-seven draft population genomes of heterotrophic bacteria were reconstructed to explore their nutrient cycling and stress avoidance strategies. Our genomic exploration of abundant microorganisms such as *Bacillus*, *Brevibacillus*

and *Hymenobacter* spp. revealed many pathways for nutrient acquisition, and especially nitrogen metabolism, implicating them as central drivers of nutrient turn-over in this extreme environment. These genomes also offer evidence as to the metabolic mechanisms by which microorganisms persist in this extreme environment, whereby the accumulation of compatible solutes and capacity to transport salt from the cell appears to be critical.

Next we set out to quantify the proportion of microbial taxa that became active following a precipitation event by identifying the spectrum of active and inactive biocrust community members using labelled amino acid probing coupled with cell sorting and click chemistry. This approach highlighted clear differences between translationally-active and -inactive bacterial guilds. Overall, we found evidence that ~61% of biocrust community members become active following wetting, with a very high proportion of those belonging to the nitrogen-fixing *Alphaproteobacteria*, which increased from 42% after 4 hours of wetting, to 53% relative abundance by the end of our wet-up experiment (21 hours after wetting). This pioneering work in a model soil system provides initial evidence of the dynamic responses of biocrust microorganisms to a rainfall event, and has the potential to reveal the key drivers of nutrient cycling in arid ecosystems.

Altogether, our analyses show that most biocrust microorganisms are sensitive to sporadic wetting events, and immediately prepare for anticipated moisture loss through preservation mechanisms such as sporulation, the accumulation of compatible solutes and a switch from anabolic to catabolic metabolism. These adaptations have the potential to dramatically impact carbon fixation through the reduction in cyanobacterial abundance and the facilitation of nitrogen cycling through heterotrophic bacteria in biocrusts.

This work conducted was supported by the Office of Science Early Career Research Program, Office of Biological and Environmental Research, of the U. S. Department of Energy under Contracts No. DE-AC02-05CH11231.

Comparative genomic and activity-based analyses reveal widespread potential for direct extracellular electron transfer among diverse methane-oxidizing ANME archaea

Chadwick, Grayson,[1]* (chadwick@caltech.edu), Connor Skennerton,[1] Hang Yu,[1] Mark Ellisman[2], Gene Tyson,[3] Roland Hatzenpichler[1], Rex Malmstrom[4], Tanja Woyke[4], and **Victoria J. Orphan**[1]

[1]California Institute of Technology, Pasadena, CA; [2]University of California, San Diego, San Diego, CA; [3]University of Queensland, Brisbane, Australia; and [4]Joint Genome Institute, Walnut Creek, CA

Project Goals: The major goals of this project are to identify novel archaea and syntrophic microorganisms involved in methane cycling, examine their genomic characteristics, and test their metabolic potential and interspecies interactions using cross-disciplinary methods across a range of spatial scales. Using a combination of experimental and modeling-based approaches, we are testing potential metabolic interactions between methane-oxidizing archaea and sulfate-reducing bacteria, with a specific focus on direct extracellular electron transfer as one of the key mechanisms for energy conservation during anaerobic oxidation of methane. More broadly, by measuring and modeling spatial patterns of activity within diverse structured consortia and biofilms at the micron scale, our goal is to determine whether these patterns can provide fundamental information about the nature of the interaction and metabolic interchange.

The anaerobic oxidation of methane coupled to sulfate-reduction is a microbially mediated process requiring a syntrophic partnership between currently uncultured archaea, known as anaerobic methanotrophs (ANME) and sulfate-reducing bacteria (SRB). These sediment-hosted organisms form multi-celled aggregates comprised of one ANME and one SRB partner; and these syntrophic consortia exhibit diverse, but reproducible, spatial organization. ANME lineages are polyphyletic, representing several family-level and one order-level clades of the *Methanomicrobia*. Based on their phylogenetic position and metabolic similarity to methanogens, it is probable that the mechanism for growth by anaerobic methane oxidation occurred through horizontal gene transfer of genes involved in extracellular electron transfer and has evolved multiple times in different clades. Here we used metagenomic sequencing to reconstruct the genomes of 31 representatives from all known ANME clades to determine what separates ANME from methanogens and what differentiates ANME clades from each other. In order to better understand the functional basis for the different consortia spatial structures, we have optimized a protocol for fluorescence *in situ* hybridization (FISH) to identify the archaeal and bacterial partners and observe aggregate morphology combined with laser micro-dissection to recover and amplify the 16S ribosomal RNA genes directly from the imaged consortia, offering a direct link between aggregate structure, strain-level phylogenetic identity, and genomic content. These genomic and microscopic observations of individual ANME clades are considered in light of recent spatial modeling efforts to better understand the adaptive rational for the varying community organizations observed in nature.

Funding statement. This work was supported by the Genomic Sciences Program in the DOE Office of Science, Biological and Environmental Research DE - SC0016469. The authors declare no conflict of interest.

Reactive Transport Modeling of Microbial Processes: Cell Aggregate-scale Models and Upscaling Using Pore-scale Simulations

He, Xiaojia[1]* (xiaojia.he25@uga.edu), Heewon Jung[1], Chadwick, Grayson,[2] **Kempes, Christopher**,[4] Shi, Yimeng,[1] McGlynn, Shawn[2,3], **Orphan, Victoria**,[2] and **Meile, Christof**[1]

[1] University of Georgia, Athens, GA, USA; [2] California Institute of Technology, Pasadena, CA, USA; [3]Tokyo Institute of Technology, Tokyo, Japan; and [4]Santa Fe Institute, Santa Fe, NM, USA

https://faculty.franklin.uga.edu/cmeile/microbial-metabolism

Project Goals

The overarching project goal is to expand our understanding of the key microorganisms, metabolic strategies, and interspecies relationships involved in the formation and oxidation of methane. Our research evaluates potential mechanisms controlling anaerobic oxidation of methane (AOM) medicated by archaeal-bacterial consortia at the aggregate-scale and studies the influence of pore scale physico-chemical heterogeneity on macroscopic rate estimates in marine sediments.

At the aggregate scale, three different mechanisms for extracellular interspecies electron transfer are investigated and the results are compared to the cell-specific anabolic activity obtained by FISH-nanoSIMS. The sensitivity of model outputs towards poorly constrained model parameters is also explored to constrain model parameters for use in larger scale models. Next, we investigate the interaction between microorganisms and pore scale microenvironments, which may not be properly resolved by macroscopic rate measurements. The role of heterogeneous distribution of microorganisms on macroscopic rate estimates is explored in various flow and reaction conditions and important parameters for upscaling with improved accuracy are proposed.

Microorganisms dynamically interact with their microenvironments in nature through complex physical, chemical, and biological processes. Because of the complex and heterogeneous nature of porous media, predicting microbial activities across different scales remains challenging. We address the challenge of identifying mechanisms of syntrophic microbial interactions and upscaling microbial metabolism at the microbe-to-pore scale using reactive transport modeling approaches.

First, we simulate the activity of archaeal and bacterial cells mediating anaerobic oxidation of methane coupled with sulfate reduction at the scale of microbial aggregates. Three different mechanisms are investigated including electron transfer through the exchange of solutes such as H_2, the delivery of disulfide from methane-oxidizing archaea to bacteria for disproportionation, and direct interspecies electron transfer. We identify the mechanisms of synthrophic interactions that are consistent with both macroscopic rates and microscopic measures of anabolic activity. To that end, simulation data are compared to multi-modal image data from FISH-nanoSIMS observations, for which a machine learning approach to automate this analysis is being developed. The results

indicate electron transfer through the exchange of solutes such as H_2 is unlikely because the simulated intra-aggregate microbial activities differ from the observed distribution of nitrogen incorporation, and because oxidation rates are limited by the build-up of metabolites, inconsistent with observed rates. Instead our proposed DIET model yielded cell specific rates and archaeal activity distributions that were consistent with empirical observations, with little impact of the spatial distribution of bacterial and archaeal cells and consortium sizes. Our results demonstrate the successful integration of numerical modeling and experimental observations that improve our understandings of microbial activities, and point to direct interspecies electron transfer as a possible syntrophic mechanism (He et al. 2018).

Next, the role of heterogeneous distribution of microbial aggregates at the pore scale to upscaled microbial reaction rates is investigated under various flow and reaction kinetics conditions. Lattice-Boltzmann simulations reveal that scaling errors depend strongly on Peclet and Damkohler numbers, and to a less extent on the distribution of microbial aggregates (Jung and Meile, submitted). It is also shown that the systematic integration of macroscopic parameters improves the accuracy of upscaling biochemical reaction rates, demonstrating the importance of mechanistic understandings of upscaling physical and biochemical processes in porous media.

References

1. He, X., Chadwick, G., Kempes, C., Shi, Y., McGlynn, S., Orphan, V. and Meile, C., 2018. Microbial interactions in the anaerobic oxidation of methane: Model simulations constrained by process rates and activity patterns. *Environmental microbiology*, in press. doi.org/10.1111/1462-2920.14507
2. Jung, H., and Meile, C. Upscaling of microbially mediated reactions in porous media, submitted.

Funding statement

This work was supported by the Genomic Sciences Program in the DOE Office of Science, Biological and Environmental Research DE - SC0016469. The authors declare no conflict of interest.

Systems analysis of a fast growing N_2-fixing cyanobacterium for production of advanced biofuels and nitrogen-containing petrochemical replacement compounds

Anindita Banerjee[1]* (anindita@wustl.edu), Zi Ye,[1] John I. Hendry[2], Costas D. Maranas[2], Maciek R. Antoniewicz[3] and **Himadri B. Pakrasi**[1]

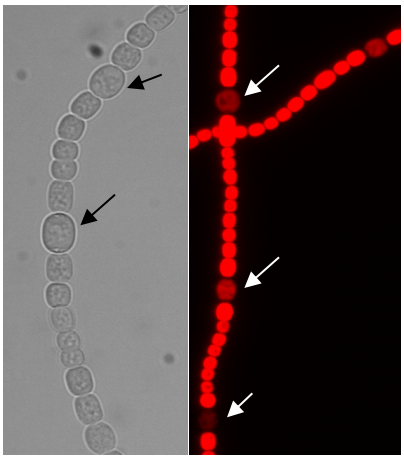
[1]Washington University in St. Louis, MO; [2]Pennsylvania State University, University Park, PA, [3]University of Delaware, Newark, DE

https://sites.wustl.edu/photosynthbio/anabaena_33047/

Project Goals:

The overall objective of this project is to use an integrated systems biology approach to develop the filamentous cyanobacterium *Anabaena* sp. PCC 33047 as a model fast-growing, photosynthetic, diazotrophic production platform. The specific goals for this project are: 1) Construct a genome-scale metabolic model and predict genetic alterations that optimally direct fixed CO_2 and N_2 into target products. 2) Apply 13C and 15N assisted metabolomics and metabolic flux analysis to dissect the metabolism of the strain. 3) Develop an efficient genetic toolkit. 4) Demonstrate production of caprolactam and valerolactam in engineered *Anabaena* 33047. 5) Establish a stable consortium between *Anabaena* 33047 and a heterotroph for cost-effective bioproduction.

Cyanobacteria are oxygenic photosynthetic prokaryotes that rely solely on atmospheric CO_2 and sunlight for growth and biomass generation. These microbes are gaining importance as platforms for bioproduction. However, the commonly used cyanobacterial production strains are relatively slow growing and high-light intolerant, prohibiting high product yields. To meet the emerging need for high-light tolerant strains with high photosynthetic CO_2 fixation rates, we are developing a fast-growing, diazotrophic cyanobacterium, *Anabaena* 33047, as a model production platform. *Anabaena* 33047 displays a doubling time of 3.8 hours under photoautotrophic and N_2-fixing condition, utilizes high light (3000µE m^{-2} s^{-1}) to convert CO_2 into biomass at a rate that is the highest known so far among oxygenic photosynthetic organisms (3.0 g L-1 day-1)[1]. However, this strain is not well studied and tools for its genetic modifications are lacking. In the current phase of our project we are focusing on developing a genome-scale metabolic model (GSM) for *Anabaena* 33047 and collecting physiological data that will be used to constrain the model. We are also working on developing a tool kit for the genetic transformation of this strain.



Images showing a high frequency of heterocysts (arrows) in the *Anabaena* 33047 filaments. The strain achieves a 2-fold higher heterocyst frequency compared to the model *Anabaena* 7120

Our GSM construction strategy seeks to consolidate the biochemical information from different pathway databases and available genome scale models to obtain a comprehensive description of the metabolism of *Anabaena* sp. ATCC 33047. We used the genome sequence of *Anabaena* sp. ATCC 33047[2] to construct draft genome scale reconstructions from three pathway databases: KEGG, MetaCyc and ModelSEED. These three draft reconstructions contained 1712,

1024 and 1521 reactions and accounted for 683, 1229 and 859 genes respectively. Additionally, based on sequence homology we constructed another draft reconstruction from the recently published model for *Anabaena* sp. PCC 7120[3]. This fourth draft network contained 964 reactions and accounted for 663 genes. Together, these four draft networks represent predicted reaction annotations for 1544 genes and the process of merging these four draft networks into a single draft network is underway. The synthesized draft network will form the basis for developing the final curated model.

We are characterizing the growth physiology of *Anabaena* 33047 and generating data that will be used to constrain genome-scale mathematical models of *Anabaena* metabolism. Specifically, we are measuring the dynamics of biomass composition during cell culture, which includes measuring macromolecular composition of cells (i.e. percentages of proteins, carbohydrates, RNA, lipids in cell dry weight), as well as quantifying nutrient uptake rates and product secretion rates. These data will then be combined with 13C-tracing studies to generate a detailed baseline map of metabolism of *Anabaena* 33047.

As part of an effort to build a genetic toolkit for *Anabaena* 33047, we have developed a conjugation system for gene transfer into the strain. We mined the genome sequence of *Anabaena* 33047 for methylases of its restriction modification system. These methylases were then cloned into a helper plasmid with the anticipation that their expression would protect the exogenous DNA from degradation by the host. Multiple gene deletion constructs for *Anabaena* 33047 were generated and used to successfully conjugate the strain. Surprisingly, unlike other filamentous cyanobacterial strains where single homologous recombination is known to be prevalent, all the colonies tested in *Anabaena* 33047 for the gene deletions were double recombinant gene replacement mutants. We are also developing systems for gene integration and CRISPR/Cpf1 mediated genome editing.

References

1. Gonzalez Lopez CV, Acien Fernandez FG, Fernandez Sevilla JM, Sanchez Fernandez JF, Ceron Garcia MC, Molina Grima E. Utilization of the cyanobacteria *Anabaena* sp. ATCC 33047 in CO2 removal processes. Bioresource technology. 2009;100(23):5904-10.
2. Pfeffer S, Brown RM. Complete Genome Sequence of the Cyanobacterium *Anabaena* sp. 33047. Genome Announcements. 2016;4(4):e00809-16.
3. Malatinszky D, Steuer R, Jones PR. A Comprehensively Curated Genome-Scale Two-Cell Model for the Heterocystous Cyanobacterium *Anabaena* sp. PCC 7120. Plant Physiol. 2017;173(1):509-23.

Funding statement: This study is being supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under Award Number DE-SC0019386

Metabolic Analysis of Cyanobacteria Carboxysome Mutant Indicates a More Flexible lux Network for Bio-manufacturing

Mary H. Abernathy[1], Jeffery J. Czajka[1*], Doug Allen[2,3], **Jeffrey C. Cameron[4,5,6]** and **Yinjie J. Tang[1]**

[1]Department of Energy, Environmental and Chemical Engineering, Washington University in St. Louis, MO 63130; [2]Donald Danforth Plant Science Center, St. Louis, MO 63132, USA; [3]United States Department of Agriculture, Agricultural Research Service, St. Louis, MO 63132, USA; [4]Department of Biochemistry, University of Colorado Boulder, Boulder, CO 80309; [5]Renewable and Sustainable Energy Institute, University of Colorado, Boulder, CO 80309; [6]National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO 80401

Presenter email: jjczajka@wustl.edu

Project goals:

1. Investigate cyanobacterial metabolic network and reveal its regulations for fast biosynthesis.
2. Elucidate effects of micro-compartmentation on cyanobacterial metabolisms.

Abstract

Cyanobacterial carboxysomes encapsulate carbonic anhydrase and ribulose-1,5-bisphosphate carboxylase/oxygenase and are key organelles that promote CO_2 concentration and fixation. Genetic deletion of the major structural proteins encoded within the ccm operon in *Synechococcus* sp. PCC 7002 (Δ*ccmKLMN*) disrupts carboxysome formation and significantly affects cell physiology. In this study, we employed both metabolite pool size analysis and isotopically nonstationary metabolic flux analysis (INST-MFA) to examine metabolic regulation in cells lacking carboxysomes. Under high CO_2 environments, the Δ*ccmKLMN* mutant had similar growth rates as the control strain and maintained a similar flux distribution through the central metabolism, with the exceptions of moderately elevated protein synthesis and photorespiration activity. Metabolite analyses indicated that the Δ*ccmKLMN* strain had larger pool sizes of pyruvate, UDPG, and aspartate as well as higher levels of secreted malate and succinate. Under photomixotrophic conditions, both the control strain and the Δ*ccmKLMN* mutant metabolized acetate and pyruvate. Provision of acetate promoted carboxysome mutant growth when light and CO_2 were insufficient. The results suggest that the Δ*ccmKLMN* mutant is able to minimize changes in fluxes (except for elevated photorespiration) and instead reorganizes its metabolism through significant changes in intracellular metabolite pool concentrations. The removal of microcompartments may loosen the flux network regulation and allow for redirection of central metabolites to competing pathways (e.g., lactate production). This study provides important insights into both metabolic regulation via enzyme compartmentation and the compensatory responses in cyanobacterial phototrophic metabolism.

Publications

1. Abernathy, et al. Cyanobacteria carboxysome mutant reveals the influence of enzyme compartmentalization on cellular metabolism and metabolic network rigidity. *Submitted*

Funding Statement

This work was partially supported by the Department of Energy, DESC0012722.

Genome-scale fluxome of the fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973

John I. Hendry[1*]-(jih11@psu.edu), Saratram Gopalakrishnan[1], Justin Ungerer[2], Yinjie J. Tang[3], Costas D. Maranas[1], **Himadri B. Pakrasi**[2]

[1]Department of Chemical Engineering, Pennsylvania State University, University Park, PA 16802, USA; [2]Department of Biology, Washington University in St. Louis, St. Louis, MO 63130, USA; [3]Department of Energy, Environment and Chemical Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA

http://www.maranasgroup.com/

Project Goals: The goal of this project is to analyze the genome-scale fluxome of *Synechococcus elongatus* UTEX 2973 (here after *Synechococcus* 2973) as part of a system level study of this fast-growing cyanobacterial strain. This would enable us to identify if there is any specific pathway utilization pattern associated with fast growth rate. The knowledge base generated from this study would help in the development of *Synechococcus* 2973 as a versatile host for a sustainable solar-based bioprocess.

Photosynthesis produces a wide variety of biochemicals from cheap and renewable raw materials such as water, sunlight and CO_2. The oxygenic phototrophs, cyanobacteria, are more pliable for genetic modification than plants and micro-algae. This makes them an ideal host for the photosynthesis-based production of industrially relevant biochemicals and biofuels. However, the slow growth rates of well-studied model strains prevented industrial application of these organisms as solar-based production platforms. The fast-growing cyanobacterium, *Synechococcus* 2973, has a short doubling of 2.1h which is comparable with the widely used industrial production host yeast (Yu et al., 2015). This makes it a potential candidate host that could sustain a solar-based bioprocess at an industrial scale. Systems level analysis of such fast-growing photoautotrophs will accelerate their development into a successful production host. Besides, such analysis would help us unlock factors that influence the growth rate of photosynthetic organisms (Mueller et al., 2017). The knowledge base, generated from such analysis, could help us further increase the growth rates of photosynthetic organisms making them more suitable for industrial production processes.

In this meta-analysis, the genome scale fluxome of *Synechococcus* 2973 was elucidated using isotopic non-stationary ^{13}C-metabolic flux analysis with experimentally measured labeling dynamics of 13 central carbon metabolites obtained from a previously reported study (Hendry et al., 2018). To begin with, we created a genome scale carbon mapping model, *i*mSyu593, using the existing mapping model for *Synechocystis* sp PCC 6803, *i*mSyn617, as the starting point. The mapping model, *i*mSyu593, traced the flow of carbons through 593 reactions encompassing central carbon metabolism, amino acid metabolism and other peripheral pathways. Flux distribution revealed that almost all (>96%) of the assimilated carbons were directed towards biomass formation. This high carbon conversion is the result of reincorporation of oxidized carbons and preferential usage of non-decarboxylating reactions such as phosphoketolase. The reincorporation of oxidized carbon compensated for the carbon loss associated with the observed higher flux through the photorespiratory C2 cycle. This allowed the organism to use the photorespiratory C2

cycle for the synthesis of glycine and serine without a significant decrease in the carbon efficiency. Unlike in other cyanobacteria, the malic enzyme flux was found to be dispensable since Pyruvate Kinase was the major source of pyruvate. Acetyl CoA was synthesized using the carbon efficient phosphoketolase pathway instead of Pyruvate dehydrogenase. These findings suggest the existence of a carbon efficient metabolism in *Synechococcus* 2973 alongside its faster growth kinetics.

References:

1. Hendry, J. I., Gopalakrishnan, S., Ungerer, J., Pakrasi, H. B., Tang, Y., Maranas, C. D., 2018. Genome-scale fluxome of *Synechococcus elongatus* UTEX 2973 using transient ^{13}C-labeling data. Plant Physiol.
2. Mueller, T. J., Ungerer, J. L., Pakrasi, H. B., Maranas, C. D., 2017. Identifying the Metabolic Differences of a Fast-Growth Phenotype in *Synechococcus* UTEX 2973. Sci Rep. 7**,** 41569.
3. Yu, J., Liberton, M., Cliften, P. F., Head, R. D., Jacobs, J. M., Smith, R. D., Koppenaal, D. W., Brand, J. J., Pakrasi, H. B., 2015. *Synechococcus elongatus* UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO_2. Sci Rep. 5**,** 8132.

Supported by funding from the U.S. Department of Energy by grant DE-SC0012722.

Syntrophic co-cultures of *Clostridium* organisms to produce higher alcohols and other C6-C8 metabolites

Kamil Charubin[1,2,*](kamilcha@udel.edu), Hannah Streett[1,2], Jonathan Otten[1,2], Charles Foster[3], Dr. Maciek Antoniewicz[1], Dr. Costas Maranas[3] and **Dr. Eleftherios Terry Papoutsakis**[1,2,]

[1] University of Delaware, Newark, DE; [2] Delaware Biotechnology Institute, Newark, DE;
[3] Pennsylvania State University, State College, PA

Project Goals: To develop syntrophic *Clostridium* co-culture systems for producing intermediate carbon-chain length metabolites (C4-C8) and their derivatives that can be used as chemicals or serve as biofuels and their precursors.

Clostridium organisms are of major importance for developing new technologies to produce biofuels and chemicals. Three major types of *Clostridium* organisms have been the focus of studies for the sustainable production of fuels and chemicals. Solventogenic clostridia are capable of utilizing a large variety of biomass-derived carbohydrates such as hexoses, pentoses, disaccharides, and hemicellulose, and can produce a good number of C2-C4 chemicals. Acetogenic clostridia can fix inorganic H_2, CO_2, and CO to generate C2 acids and alcohols. Other specialized clostridia possess diverse biosynthetic capabilities for production of a wide variety of metabolites including C4 – C8 carboxylic acids and alcohols, which could serve as commodity chemicals, biofuels, or biofuel precursors. The majority of previous work with clostridia focused on optimizing single-organisms systems for production of biochemical. In nature, microorganisms live in complex communities where syntrophic interactions result in superior resource utilization. Here, we first examined a synthetic syntrophy consisting of the solventogen *Clostridium acetobutylicum*, which converts simple and complex carbohydrates into a variety of chemicals, and the acetogen *C. ljungdahlii*, which fixes CO_2. This synthetic co-culture achieved carbon recoveries into C2-C4 alcohols almost to the limit of substrate-electron availability, with minimal H_2 and CO_2 release. The syntrophic co-culture produced robust metabolic outcomes over a broad range of starting population ratios of the two organisms. Finally, the co-culture exhibited unique direct cell-to-cell interactions and material exchange among the two microbes, which enabled unforeseen rearrangements in the metabolism of the individual species that resulted in the production of non-native metabolites, namely isopropanol and 2,3-butanediol. Next, to expand this co-culture system to include *C. kluyveri*, which can metabolite ethanol and acetate to produce C6 and C8 carboxylic acids. Both *C. acetobutylicum* and *C. ljungdahlii* produce ethanol and acetate, which makes *C. kluyveri* and ideal partner for a triple synthetic co-culture system capable to converting biomass-derived carbohydrates to C6 and C8 biochemicals.

^{13C}-based Metabolic Flux Analysis (MFA) will be used to gain insight into the regulation of cell growth and product formation pathways, and to identify metabolic bottlenecks. Currently, use of stable-isotope (e.g. ^{13}C) tracers combined with measurements of isotopic labeling by mass spectrometry represents the state-of-the-art in flux determination. After intense research and

development in past two decades, ^{13}C-based MFA methods are now widely used to probe fluxes in microbes. Metabolic fluxes will be studied using ^{13}C MFA in *C. kluyveri*, *C. acetobutylicum*, and *C. ljungdahlii* under mono- and co-culture conditions to identify key changes in metabolism of each organism.

To predict and better understand the co-culture interactions and predict steady-state organism abundances, a consortium model consisting of *C. kluyveri*, *C. acetobutylicum*, and *C. ljungdahlii* will be constructed using the SteadyCom framework. This will be done by standardizing the biomass equations and metabolite naming conventions for existing genome-scale models (GSMs), and updating each GSM using RNAseq and ^{13}C-fluxomics procured under varying experimental conditions. ^{13}C-fluxomics and RNAseq data will be used to infer regulatory events in each organism to simulate and compare transient monoculture and co-culture population dynamics. Computational strain design algorithms, such as OptKnock, OptForce, will be then customized for consortium models and combined with the SteadyCom framework to identify genetic perturbation targets in the clostridia consortium that lead to the overproduction of C6 and C8 biochemicals.

SUPPORTED by the U.S. Department of Energy (Award No. DE-SC0019155).

Environmental Drivers of the North American *Populus* Microbiome

Michael E. Van Nuland (mvannula@stanford.edu),[1]* Caroline C. Daws,[1] and **Kabir G. Peay**[1]

[1]Department of Biology, Stanford University, Stanford, CA 94305

https://mykophile.com/plant-microbe-interactions/mutualistic-niche/

Project Goals: Plant-microbe symbioses play a critical role in the health and nutrition of all plants, and especially so for *Populus* tree species targeted for bioenergy feedstock programs. The primary goal of this project is to test quantitative predictions about the climatic, edaphic, and historical variables that best control *Populus*-microbe symbioses and their impacts on ecosystem function. Identifying the factors that influence the distribution of different *Populus* species and their symbiotic microbial communities (e.g., rhizosphere bacteria, mycorrhizal fungi) has practical implications for human management of forest systems and sustainable production of biofeedstocks that depend on these associations. Results from this project will identify areas where *Populus*-symbiont relationships may be especially critical for maintaining natural forest productivity in marginal habitats, indicating a potential probiotic source for tree plantations under future environmental conditions.

Microbial communities form symbiotic relationships with plants above- and belowground, but we lack a general understanding of the biogeographic drivers of such relationships. This is important because the distributions of host-associated symbionts (i.e., fungal mutualists) are likely to vary across the host species' range and affect their ability to tolerate certain abiotic pressures. This creates a long-term disadvantage to sustainably manage biofeedstock programs and accurately forecast ecosystem productivity as plant-microbiome interactions will likely be disrupted or altered by environmental change. However, examining how microbial communities vary throughout a host species' range can reveal the environmental factors most important for shaping plant-microbiome symbioses, thus refining our ability to predict the dynamics of such relationships under future conditions. We sequenced bacterial and fungal communities in leaf, root, and soil material from five ecologically and economically important *Populus* species across the United States (*P. angustifolia*, *P. deltoides*, *P. fremontii*, *P. tremuloides*, and *P. trichocarpa*). Here, we asked two main questions: (1) how do microbial communities associated with *Populus* leaves, roots, and soils differ? (2) what are the major environmental gradients that structure microbial community composition within and among *Populus* species?

We find differences in foliar, root, and soil microbial community composition among the five *Populus* species, with the greatest divergence between species from disparate habitats. Moreover, the environmental gradients identified as the most important for driving microbial community turnover differed by sample type (leaves, roots, and soil) and among *Populus* species. This indicates that the climatic and edaphic factors important for shaping *Populus* microbial communities (and potentially their impact on plant performance and functioning) varies

depending on how they interact with the host (i.e., within or surrounding plant tissues) and respond to abiotic stress. Species differed in the level of belowground fungal symbioses, with *P. fremontii* having the least (mean = 5.4% ± 3) and *P. tremuloides* having the greatest (mean = 28.7% ± 3.5) root colonization by ectomycorrhizal fungi. We also found higher total numbers of ectomycorrhizal taxa from *Populus* species with larger range sizes, suggesting a positive link between the diversity of fungal mutualists and plant host distributions on the landscape. The dominant environmental gradients corresponding with microbial community turnover differed among *Populus* species, indicating that the role of abiotic effects on plant-microbiome symbioses may be species-specific. Finally, analysis of the most common taxa found across all *Populus* species shows variation in the connectivity and modularity of co-occurrence networks, suggesting there are differences in how community members interact or respond to the same environmental pressures. Collectively, these results demonstrate the key environmental drivers of the *Populus* microbiome, which can be used to inform management decisions and potential planning of *Populus* biofeedstock programs based on the diversity and biogeography of their natural microbial symbionts.

This project is funded by an Early Career Award (DE-SC0016097) from the Office of Biological and Environmental Research in the DOE Office of Science.

Enhanced Resistance Pines for Improved Renewable Biofuel and Chemical Production

Gary F. Peter[1], gfpeter@ufl.edu

[1]University of Florida, Gainesville, FL

Project Goals: Our goal is to genetically increase constitutive terpene defenses of loblolly and slash pine to enhance protection against pests and pathogens and at the same time expand terpene supplies for renewable biofuels and chemicals.

Abstract: Today, the southeastern U.S. hosts the world's largest biomass supply chain, annually delivering 17% of global wood products, more than any other country. This well-developed regional supply chain supports southern pine genetic improvement, seedling production and planting, silviculture, harvesting, and transportation annually delivering ~250 million tons of pine wood to integrated manufacturing facilities. In the SE U.S., 39 million acres of land not suited for food production are planted with genetically improved loblolly and slash pine seedlings selected and managed for fast growth and high wood yields. The SE also houses the U.S. pine chemicals industry the oldest and one of the largest renewable hydrocarbon chemical industries with favorable cost-competitiveness with petroleum derived feedstocks. Our focus is on increasing constitutive terpene production to enhance loblolly and slash pine resistance to pests and pathogens. Enhanced resistance in these commercial species is critical to protect against widespread losses as biotic pressures increase due to global warming, landuse change and introduced exotic organisms.

Increasing pine terpenes also is aligned well with the needs of the developing bioeconomy. Today, commercial scale collection of pine terpenes occurs from live trees by tapping, stumps by solvent based steam extraction, and from pulp mills as co-products. US pulp mills recover ~900,000 tonnes y^{-1} of terpenes and fatty acids supporting specialty chemical biorefineries that compete in markets with petroleum derived feedstocks supporting our concept that biofuels from pine terpenes could be profitable without subsidy if supply was increased. Proven technologies exist to efficiently convert pine terpenes to biofuels: a 30 million gallon y^{-1} bio-refinery produces renewable diesel from pine terpenes and fatty acids and pine monoterpenes can be efficiently dimerized to produce a replacement for JP10, the highest density jet fuel. Higher wood terpene content will increase the yield of bioenergy per unit mass in the rapidly developing wood pellet industry. Pine terpene supply is limited by the relatively low average wood terpene content.

Pine terpenes evolved as a primary chemical and physical defense system and are a main component of a durable, quantitative defense mechanism against pests and pathogens. The terpene defense traits are under genetic control and can be improved by breeding and genetic engineering. Our goal is to genetically increase constitutive terpene defenses of loblolly and slash pine to enhance protection against pests and pathogens and at the same time expand terpene supplies for renewable biofuels and chemicals. Objective one will integrate existing and new genome wide association genetic results with RNA expression, QTL mapping, and allele

frequency information in known high oleoresin flow selections and our breeding populations to discover and validate loblolly and slash pine alleles/genes that are important for resistance. Objective two will use information from objective one to accelerate breeding for increased resistance in loblolly and slash pine through marker assisted introgression, and will develop and test genomic selection models to accelerate breeding of resistant slash pine.

Funding: U.S. Department of Energy – USDA Biomass Feedstocks DE-SC0 019099

Taxon-Specific Growth and Mortality Rates of Bacteria in Soil Following the Rewetting of a Seasonally Dried Grassland

Steven J. Blazewicz (blazewicz1@llnl.gov)[1*], Erin E. Nuccio[1], Benjamin J. Koch[2], **Bruce A. Hungate**[2], Ember Morrissey[3], Alexa Nicholas[4], Eoin L. Brodie[5]**, Mary K. Firestone**[4], **Jennifer Pett-Ridge**[1]

[1]Lawrence Livermore National Laboratory; [2]Northern Arizona University, Flagstaff, AZ; [3]West Virginia University, Morgantown; [4] University of California, Berkeley, CA; [5]Lawrence Berkeley National Laboratory, Berkeley, CA

Project Goals: Microorganisms play key roles in soil carbon turnover and stabilization of persistent organic matter via their metabolic activities, cellular biochemistry, and extracellular products. Microbial residues are the primary ingredients in soil organic matter (SOM), a pool critical to Earth's soil health and climate. We hypothesize that microbial cellular-chemistry, functional potential, and ecophysiology fundamentally shape soil carbon persistence, and we are characterizing this via stable isotope probing (SIP) of genome-resolved metagenomes. We focus on soil moisture as a 'master controller' of microbial activity and mortality, since altered precipitation regimes are predicted across the temperate U.S. *Our SFA's ultimate goal is to determine how microbial soil ecophysiology, population dynamics, and microbe-mineral-organic matter interactions regulate the persistence of microbial residues under changing moisture regimes.*

Microbial activity increases after rewetting dry soil, resulting in a pulse of carbon mineralization and nutrient availability. This phenomenon is of particular interest because predicted changes in precipitation patterns in Mediterranean ecosystems could profoundly change soil C dynamics and nutrient availability. While there has been much interest in the response of indigenous communities to wet-up of dry soil, to date no work has identified the microorganisms in semi-arid soils that grow in response to soil wet-up. We used heavy water ($H_2{}^{18}O$) DNA quantitative stable isotope probing coupled with high throughput sequencing of bacterial 16S rRNA genes to characterize taxonomic and phylogenetic composition of bacteria following the rewetting of a seasonally dried California annual grassland soil. Bacterial growth and mortality were detected at all time points throughout the incubation (3, 24, 72, 168 h), with patterns of sequential growth observable at the phylum and order levels. Of the 37 phyla detected in the pre-wet community, growth was found in 18 phyla, while mortality was measured in 26 phyla. Rapid growth and mortality rates were measurable within 3 hours of wet-up but had contrasting characteristics; growth at 3 h was dominated by a select few taxa found in the Proteobacteria and Firmicutes phyla, whereas mortality was taxonomically widespread. Sequential growth patterns observed at the phylum and order level suggest that an ecologically coherent response occurred at a high taxonomic level with members of particular groups employing different life strategies in response to rewetting.

To determine the effect of reduced Spring rainfall on newly fixed plant C persistence, genome resolvable ecophysiological traits of growing microbes, and phage-driven bacterial mortality, we conducted a second wet-up experiment using soils from an ongoing field experiment at the Hopland Research and Extension Center (HREC) in Hopland, California. Precipitation on these

plots has been manipulated so that half the plots receive a 50% reduction of the 65-year rain fall average, and the other half receive the full average rainfall amount. A $^{13}CO_2$ plant-labeling event was conducted in the Spring of 2018 to trace the pathways and persistence of C as it was fixed by plants, delivered belowground in the form of exudates and fresh roots, consumed by the various residents of the soil during the growing season, and persisted after plant senescence and the hot dry summer. Isotopically labeled soils were collected at the end of summer, shortly before the first seasonal rainfall event. These soils were subjected to a simulated precipitation event using $H_2{}^{18}O$ in the laboratory to quantify and characterize C efflux, microbial growth and mortality, and phage dynamics driven by the rapid change in water potential using a combination of SIP-targeted metagenomics, virome sequencing and imaging, and ^{13}C flux measurements.

Funding Statement
The initial wet up study was supported by a National Science Foundation Graduate Research Fellowship and a NSF Doctoral Dissertation Improvement Grant (DEB-1011093) (to SJB). Data analyses and the second wet up study were conducted as part of the LLNL Soil Microbiome Scientific Focus Area, funded by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under FWP SCW1632, "Microbes Persist: Systems Biology of the Soil Microbiome". Work at LLNL was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 and Lawrence Berkeley National Lab under contract DE-AC02-05CH11231.

Quantitative Stable Isotope Probing to Measure Microbial Growth and Activity

Megan Foley*[1] (mmf289@nau.edu), Erin E Nuccio[2], Ashley Campbell[2], **Steven J Blazewicz[2], Jennifer Pett-Ridge[2] and Bruce Hungate**[1]

[1]Center of Ecosystem Science and Society and Biological Sciences, Northern Arizona University, Flagstaff AZ, [2]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore CA

Project Goals: Microorganisms play key roles in soil carbon turnover and stabilization of persistent organic matter via their metabolic activities, cellular biochemistry, and extracellular products. Microbial residues are the primary ingredients in soil organic matter (SOM), a pool critical to Earth's soil health and climate. We hypothesize that microbial cellular-chemistry, functional potential, and ecophysiology fundamentally shape soil carbon persistence, and we are characterizing this via stable isotope probing (SIP) of genome-resolved metagenomes. We focus on soil moisture as a 'master controller' of microbial activity and mortality, since altered precipitation regimes are predicted across the temperate U.S. *Our SFA's ultimate goal is to determine how microbial soil ecophysiology, population dynamics, and microbe-mineral-organic matter interactions regulate the persistence of microbial residues under changing moisture regimes.*

Quantitative stable isotope probing, or qSIP, measures the concentration of rare, stable, isotopes in nucleic acids with taxon-specific resolution, enabling quantitative measurements of population-specific, *in-situ* growth (DNA) and activity (RNA) in complex microbial communities[1]. qSIP is promising for its potential to resolve taxon-specific element fluxes, which is one way to link microbial biodiversity with ecosystem processes[2]. The technique relies on stable isotope tracers whose concentrations are resolved by physical density separation. Here, we highlight recent advances and new frontiers for DNA-qSIP, including new developments in sample processing and mathematical analysis.

Addition of Isotope Tracers

qSIP begins with the addition of a tracer enriched in a rare, stable, isotope, paired with the addition of a non-labeled control, otherwise chemically identical to the tracer. For a qSIP experiment with ^{18}O-labeled water, for example, the tracer sample would receive water at 97 atom % ^{18}O (the highest enrichment typically available), and a paired control sample would receive water with natural abundance ^{18}O composition (approximately 0.2 atom % ^{18}O). While the vast majority of soil experiments have added tracers in laboratory incubations, applications to soils in the field may also be possible.

Sequencing and qPCR

Amplicon sequencing measures taxon-specific relative abundance of the target marker gene in each density fraction. A broad-coverage qPCR technique measures the total (pan-bacterial) copies of the marker gene in every fraction. Together, relative abundance data from sequencing and quantification of the marker gene allow computation of the absolute abundance of each taxon's gene copies in across fractions. We are working to expand the qSIP pipeline to enable analysis of SIP-targeted metagenomes in order to characterize the ecophysiological potential of

in-situ microorganisms. In these experiments, we are using ^{18}O and ^{13}C to target the metagenomes of specific microbial populations: $H_2{}^{18}O$ to target growing, dormant, and dying organisms; and ^{13}C organic matter (exudates, necromass, litter) to target specific heterotrophic populations.

High-throughput SIP (HT-SIP)

Recently, LLNL partnered with the Join Genome Institute to build a semi-automated SIP-metagenomics pipeline that uses robotics to speed processing of SIP samples by an order of magnitude, making well-replicated SIP experiments far more tractable. HT-SIP can be fully integrated with quantitative SIP *(*qSIP*)* coupled to metagenomic or amplicon data.

Modelling Incorporation of the Rare Isotope into DNA

The observed difference in DNA density, calculated as a weighted average, between the labeled and control treatments reflects isotopic assimilation and can be used to quantify the concentration of the rare isotope in DNA. For any given taxon, the molecular weight of DNA in the unlabeled control is estimated from the relationship between GC content and buoyant density. The theoretical heavy weight of fully-labeled DNA can be calculated based on the expected increase in weight from each additional neutron if 100% replacement by the rare isotope is expected. The molecular weight in the labeled treatment is estimated based on the molecular weight of unlabeled DNA and the density shift between treatments due to isotopic incorporation alone. Knowing the background fractional abundance of the rare isotope in DNA, the difference in molecular weight between the labeled and control treatments, relative to the theoretical maximal change in molecular weight, can be used to quantify atom fraction excess of the rare isotope.

References

1. Koch, B. J. et al. Estimating taxon-specific population dynamics in diverse microbial communities. Ecosphere **9**, e02090 (2018).
2. Hungate, B. A. et al. Quantitative Microbial Ecology through Stable Isotope Probing. Appl. Environ. Microbiol. **81**, 7570–7581 (2015).

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Number SCW1632 to the Lawrence Livermore National Laboratory, and a subcontract to the Northern Arizona University.

Linking Microbial Growth to Edaphic Variation Across a Mediterranean-Grassland Precipitation Gradient Using Shotgun-Metagenomic qSIP

Alex Greenlon*[1] (agreenlon@berkeley.edu), Karis McFarlane[2], Dinesh Adhikari[2], **Steve Blazewicz[2]**, **Mary Firestone[1]**, **Jill Banfield[1]**, **Jennifer Pett-Ridge[2]**

[1]Department of Environmental Science, Policy, and Management, University of California, Berkeley, [2]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA

Project Goals: Microorganisms play key roles in soil carbon turnover and stabilization of persistent organic matter via their metabolic activities, cellular biochemistry, and extracellular products. Microbial residues are the primary ingredients in soil organic matter (SOM), a pool critical to Earth's soil health and climate. We hypothesize that microbial cellular-chemistry, functional potential, and ecophysiology fundamentally shape soil carbon persistence, and we are characterizing this via stable isotope probing (SIP) of genome-resolved metagenomes. We focus on soil moisture as a 'master controller' of microbial activity and mortality, since altered precipitation regimes are predicted across the temperate U.S. *Our SFA's ultimate goal is to determine how microbial soil ecophysiology, population dynamics, and microbe-mineral-organic matter interactions regulate the persistence of microbial residues under changing moisture regimes.*

Recent evidence suggests that persistent soil organic matter (SOM) is largely derived from microbial sources, but the microbial-community level processes that mediate SOM formation and stabilization—and the edaphic factors that control these processes—are largely unknown. To elucidate mechanistic links between microbial ecophysiology, soil variables, and SOM formation, our SFA team is using stable-isotope labeling techniques to synthesize microbial ecology through metagenome sequencing with chemical characterization of SOM under varying mineralogical and precipitation conditions.

Here we present our initial characterization of soils from three Mediterranean grassland sites which are the primary field locations for our SFA. In addition to soil chemistry and mineralogy, we assessed the microbial communities active during the California winter, when water is least limiting to microbial growth. Triplicate soil cores were collected during the 2018 wet season from three sites that span a natural precipitation and geographic gradient in CA, with overlapping mineralogical composition at differing degrees of weathering. X-ray diffraction (XRD) suggests these sites vary primarily in clay mineral content, particularly montmorillonite. We measured soil CO_2 respiration after a 6-day incubation and found the highest median respiration rates at the moderate-precipitation site. Radiocarbon dating indicates the age of soil carbon increases with precipitation, but that $\Delta^{14}C$ of respired CO_2 is stable, and close to modern across the sites, suggesting microbial populations in all locations preferentially metabolize younger carbon sources. We applied ^{13}C nuclear magnetic resonance (NMR) spectroscopy to determine the composition of SOC from each site, and found high levels of aromatic compounds (likely plant-derived) at the highest-precipitation site, and higher lipid and carbohydrate signatures (likely microbial) at the moderate-precipitation site.

To link soil properties to microbial processes, we are developing an approach for shotgun-metagenomic quantitative stable isotope probing (qSIP). Triplicate soil samples from each site were incubated for 8 days with either natural abundance ^{16}O-H_2O or 97% atom-fraction excess ^{18}O-H_2O. DNA from each incubation was separated by ultracentrifugation into nine density fractions. Total community DNA from each density fraction for each isotope treatment and soil replicate, as well as unfractionated DNA from each incubation was used for metagenomic library preparation and sequenced on the Illumina NovaSeq platform to an average depth of 8 Gbp per fraction for each sample and treatment (180 metagenomes). We have implemented a novel approach to metagenomic assembly, separately assembling sliding-windows of adjacent density fractions and evaluating whether density fractionation improves metagenome-assembly contiguity and binning relative to assembling all fractions together, estimating that sliding windows result in approximately 40% more genome bins, with more bins at higher genome completeness. From these nine initial soil samples we predict recovery of upwards of 2,000 genome bins, with 800 at high completeness. We have quantified levels of microbial activity during incubation as shifts in density in ^{18}O-treatment libraries for individual metagenome-assembled genome bins and taxonomically-defined markers in unbinned scaffolds. In this poster we describe a statistical framework for linking the functional potential of active and inactive genome bins to mineralogical, environmental and molecular attributes of sampled soils. This work details the geographic variation in active microbial communities across the three sites that underpin our SFA and explores the utility of shotgun-metagenome sequencing to qSIP, laying the groundwork for experiments that will mechanistically link microbial ecophysiology at each site to soil moisture variation, rhizosphere carbon utilization and SOM persistence.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Number SCW1632 to the Lawrence Livermore National Laboratory, and a subcontract to the University of California, Berkeley.

Microbes Persist: Building a KBase Foundation for Viral Ecogenomics In Soil

Jeffrey A. Kimbrel*[1] (kimbrel1@llnl.gov), Benjamin Bolduc[2], **Matthew B. Sullivan[2,3]**, **Jennifer Pett-Ridge[1]**

[1]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA; [2]Ohio State University, Department of Microbiology, Columbus OH; [3]Ohio State University, Department of Civil, Environmental and Geodetic Engineering, Columbus OH

Project Goals: Microorganisms play key roles in soil carbon turnover and stabilization of persistent organic matter via their metabolic activities, cellular biochemistry, and extracellular products. Microbial residues are the primary ingredients in soil organic matter (SOM), a pool critical to Earth's soil health and climate. We hypothesize that microbial cellular-chemistry, functional potential, and ecophysiology fundamentally shape soil carbon persistence, and we are characterizing this via stable isotope probing (SIP) of genome-resolved metagenomes. We focus on soil moisture as a 'master controller' of microbial activity and mortality, since altered precipitation regimes are predicted across the temperate U.S. *Our SFA's ultimate goal is to determine how microbial soil ecophysiology, population dynamics, and microbe-mineral-organic matter interactions regulate the persistence of microbial residues under changing moisture regimes.*

The LLNL Soil Microbiome SFA is focused on the microbial biochemistry, functional potential and physiology of the soil microbiome. To measure and model the functioning of uncultivated soil communities, we are targeting specific ecophysiological 'traits' via analysis of genomes reconstructed from soil (informed by stable isotope tracing), as well as the role of phage in population turnover. Our approach to characterizing the role of viruses on microbial communities aims to identify, characterize and ecologically contextualize viruses in large-scale sequence datasets, all within the KBase environment. While metagenomic analytical platforms are now becoming widely available for microbial metagenomes, viral metagenomic analyses are critically different because:

1) Viruses lack universal genetic markers (they share no single gene), making taxonomic classification complicated with no standard available
2) Viral hosts must be predicted *in silico* to be able to evaluate virus impact on the ecosystem
3) Viral genomes can be extrachromosomal (episomal) or integrated (endogenised) into host genomes, which can confound 'viral' inferences if the genome 'termini' are not well identified
4) Coding sequences in viral genomes are governed by different rules that render prokaryote-centric structural and functional annotations less accurate

Currently, KBase, like many other platforms, does not contain opportunities for identification or exploration of viral and phage genomes. Our SFA team is currently adapting two viral workflow suites (iVirus and PhATE) into the KBase ecosystem, bringing these novel resources to the DOE research community and providing a foundation for future viral ecogenomic informatics development.

iVirus represents a collection of community and Sullivan Lab-based tools and datasets designed to enable cutting-edge viral ecology research, i.e., establishing and studying the patterns and processes that impact viral genes and genomes. This includes establishing 'bins' or 'units' that

represent functional sequence space (protein clusters), species- (viral populations) or genera- (viral clusters in network space) level taxonomy, and virus-tuned analytical tools to query these units once established. The incorporation of the iVirus suite of tools into KBase has already begun--with the vContact viral taxonomy tool (available now)—and more tools (such as identifying viral contigs in metagenomes) are coming soon. A challenge is that KBase uses a completely complementary relational database that is object-based instead of file-based (as in CyVerse where iVirus was 'born'). This requires significant time for establishing new objects and their relationships within the KBase environment. While this has slowed development time in the interim, as momentum builds, we expect to benefit for further downstream development.

The Phage Annotation Toolkit and Evaluator (PhATE) is a complementary LLNL-developed workflow for structural and functional viral genome annotation. PhATE begins by running multiple bacterial gene callers (GeneMarkS, Glimmer, Prodigal) and PHANOTATE, a phage-centric gene caller, and provides summary statistics for the resulting genes. Gene sequences (nucleotide and protein) are next BLASTed against several databases, including NCBI virus genomes and proteins, PhAnToMe, pVOGs, KEGG virus proteins, Swissprot, and NR. For KEGG virus proteins, annotation tags for each hit are queried and recorded (e.g., Pfam, taxonomy, uniprot identifiers). pVOG group identifiers are extracted from each top hit and used to construct an alignment-ready FASTA file data type containing the phage peptide of interest plus all of the members of a given pVOG group. In KBase, all annotations will be combined for each phage peptide and summarized in multiple downloadable output formats, including summary figures for visualization.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Number SCW1632 to the Lawrence Livermore National Laboratory, and a subcontract to the Ohio State University.

The Impacts of Redox Periodicity on Microbial Community Structure and Carbon Transformations in a Wet Tropical Forest Soil

Ashley Campbell[1], Amrita Bhattacharyya[2], Jeffrey A. Kimbrel[1], Gareth Trubl[1], Nicole Didonato[3], Allison Thompson[3], Rosalie Chu[3], Yang Lin[4], Whendee Silver[4], Peter Nico[2], **Jennifer Pett-Ridge[*1] (pettridge2@llnl.gov)**

[1]Lawrence Livermore National Laboratory, Livermore CA, [2]Lawrence Berkeley National Laboratory, Berkeley CA, [3]Pacific Northwest National Laboratory, Richland WA, [4]University of California Berkeley, Berkeley CA

https://pls.llnl.gov/people/staff-bios/nacs/pettridge-j

Project Goals: This Early Career research examines the genomic potential and expression of tropical soil microorganisms as they experience shifts in soil temperature, moisture, depth and oxygen availability. Associated fluctuations in redox potential are proximal controls of mineral-organic matter interactions in humid, tropical soils. By tracking the degradation and fate of organic ^{13}C labeled compounds during shifts in soil redox status, this work will improve our understanding of microbial metabolic flexibility, and how microbial processes affect the fate of organic carbon in wet tropical systems. The mechanistic understanding produced by this research will also improve the predictive capacity of mathematical models that forecast future tropical soil carbon balance.

Wet tropical soils can alternate frequently between fully oxygenated and anaerobic conditions, constraining both the metabolism of tropical soil microorganisms, and the mineral-organic matter relationships that regulate many aspects of soil C cycling. Using a 44 day redox manipulation experiment with soils from the Luquillo Experimental Forest, Puerto Rico, we examined patterns of tropical soil microorganisms and metabolites when soils were exposed to different redox regimes - static oxic, static anoxic, high frequency redox fluctuation (4 days oxic, 4 days anoxic), or low frequency redox fluctuation (8 days oxic, 4 days anoxic). Replicate microcosms were harvested throughout the incubation to measure the impact of redox condition on microbial community structure and activity, organic matter turnover, and soil chemistry. An addition of ^{13}C enriched plant biomass allowed us to distinguish decomposition of fresh liter vs native organic matter, and conduct Stable Isotope Probing (SIP) to identify the responsible microorganisms. Recently we finished sequencing 16S rRNA libraries for 1127 SIP fractions, and metagenomic DNA from 88 SIP fractions, yielding over 22 billion reads and over 3.3 trillion base pairs. Individual metagenome assemblies produced over 6,000 genome bins (MAGs), and co-assemblies are under way to produce more high-quality MAGs. Virus-specific informatics recovered >30,000 viruses clustering into 6,123 viral populations (vOTUs), with ~1.82 vOTUs/Gbp of metagenome, and 466 vOTUs $\geq$10 kb. Viral richness was highest in the oxic samples and decreased by oxic dynamics (oxic>low frequency>high frequency>anoxic).

Our bulk 16S data shows the bacterial and fungal community composition in the two fluctuating redox treatments was indistinguishable from the native soil community, while those from the static redox conditions were distinct, suggesting communities in these soils are highly adapted to fluctuating redox conditions. Using differential abundance analysis, we found that fluctuating

redox enriched for relatively more bacterial and fungal taxa as compared to the static redox conditions. However, the anoxic treatment exhibited a distinct iron-cycling microbial structure compared to the other treatments. While gross soil respiration was slightly higher in static oxic soils, respiration derived from added litter was highest in static anoxic soils, suggesting that decomposition of preexisting SOM was limited by O_2 availability in the anoxic treatment. This is supported by distinct molecular composition of soil metabolites for each treatment (measured by FTICR-MS). Together, these results suggest that dynamic environmental conditions influence microbial community assembly and biogeochemistry in ways that could not be predicted based on extrapolation from static systems. Our results also point toward a microbial community that is highly resilient to dynamic redox conditions and show that distinct C compounds are differentially processed under varying redox conditions, likely because of their varying bioavailability (driven by mineral-OM dynamics) and/or shifted microbial metabolic capacity.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Early Career Research Program Award Number SCW1478 to J. Pett-Ridge at Lawrence Livermore National Laboratory. Work was performed at Lawrence Berkeley National Lab under contract DE-AC02-05CH11231 and at Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

Sensitivity Analysis and Methodological Improvements for Quantitative Stable Isotope Probing

Ella T. Sieradzki*[1] (ellasiera@berkeley.edu), Benjamin J. Koch[2], Rebecca Mau[2], Michaela Hayer[2], Paul Dijkstra[2], Egbert Schwartz[2], Kirsten Hofmockel[3], Sheryl Bell[3], Ember Morrissey[4], **Steven Blazewicz[5],** Erin Nuccio[5], Rex Malmstrom[6], **Jill Banfield[1], Mary K. Firestone[1], Bruce Hungate[2]** and **Jennifer Pett-Ridge**[5]

[1]University of California, Berkeley, CA; [2]Northern Arizona University, Flagstaff, AZ; [3]Pacific Northwest National Laboratory, Richland, WA; [4]West Virginia University, Morgantown WV; [5]Lawrence Livermore National Laboratory, Livermore, CA; [6]Joint Genome Institute, Walnut Creek, CA

Project Goals: Microorganisms play key roles in soil carbon turnover and stabilization of persistent organic matter via their metabolic activities, cellular biochemistry, and extracellular products. Microbial residues are the primary ingredients in soil organic matter (SOM), a pool critical to Earth's soil health and climate. We hypothesize that microbial cellular-chemistry, functional potential, and ecophysiology fundamentally shape soil carbon persistence, and we are characterizing this via stable isotope probing (SIP) of genome-resolved metagenomes. We focus on soil moisture as a 'master controller' of microbial activity and mortality, since altered precipitation regimes are predicted across the temperate U.S. ***Our SFA's ultimate goal is to determine how microbial soil ecophysiology, population dynamics, and microbe-mineral-organic matter interactions regulate the persistence of microbial residues under changing moisture regimes.***

Stable Isotope Probing (SIP) is a highly effective method for linking diversity to functionality within naturally-occurring communities. Microbial substrates enriched in a stable isotope are added to a microbial community and organisms that utilize that substrate will present an increase in DNA buoyant density due to the incorporation of the stable isotope into their DNA. Quantitative SIP (qSIP) enhances the conclusions we can draw from SIP by adding statistical methods to determine the atom % enrichment of those organisms using the density shift in their DNA buoyant density, allowing us to estimate their effect on substrate assimilation or *in situ* growth rate.

Although the SIP approach for environmental studies was developed in the early 2000's[1] there has been very little assessment of reproducibility and sensitivity. As SIP becomes more quantitative and statistically robust, and transitions towards the use of genomes from metagenomes rather than marker genes, we can now take advantage of new and well-replicated studies in order to improve the sensitivity and detection limit of the method[2,3].

Using three existing datasets reflecting utilization of different labeled carbon substrates and labeled water in various environments we assessed two key aspects of qSIP variability: (1) using buoyant density to predict the GC content of a genome and (2) the effect of density gradient resolution on detection of density shifts. We confirmed that the correlation between GC content and buoyant density using pure cultures varies by up to 0.004 g/ml, which stems from method limitations. In addition, the mean buoyant density of a genome within a community has a 95% confidence interval of ±0.005 g/ml. We have been unable to improve this sensitivity by automation as opposed to manual processing but feel that it could potentially be improved via

use of an internal standard. While there has long been a presumption that collecting more density fractions might lead to more precise results, our analysis suggests otherwise, and that the ability to detect a density shift remains comparable while decreasing the resolution of gradient fractionation (i.e. fewer fractions) as long as the individual fraction size is < 0.01 g/ml (roughly 9 fractions collected per 5 ml gradient). This is an important consideration, since collecting fewer fractions renders qSIP less labor-intensive as well as less computationally complex, which has important implications for the transition to SIP-metagenomics. Even a fraction size threshold of ~0.024 g/ml (three fractions) was sufficient to quantify ^{13}C enrichment in half of all metagenome-assembled genomes above a conservatively estimated detection limit of 26 atom percent excess.

Our analysis is a first step towards enabling genome-informed SIP-metagenomics, in which instead of amplifying a marker gene from each fraction we are using shotgun sequencing and genomes assembled from metagenomes to detect DNA density shifts of entire genomes. Critically, this approach will enable us to move beyond simply identifying active organisms, and instead will identify their full metabolic potential.

References

1. Radajewski, S., Ineson, P., Parekh, N. R., & Murrell, J. C. (2000). Stable-isotope probing as a tool in microbial ecology. *Nature*, *403*(6770), 646.
2. Hungate, B. A., Mau, R. L., Schwartz, E., Caporaso, J. G., Dijkstra, P., van Gestel, N., ... & Morrissey, E. (2015). Quantitative microbial ecology through stable isotope probing. *Applied and Environmental Microbiology*, AEM-02280.
3. Youngblut, N. D., Barnett, S. E., & Buckley, D. H. (2018). SIPSim: A modeling toolkit to predict accuracy and aid design of DNA-SIP experiments. *Frontiers in Microbiology*, *9*, 570.

This research is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award Number SCW1632 to the Lawrence Livermore National Laboratory, and subcontracts to the University of California, Berkeley, and Northern Arizona University.

Solving Multi-disciplinary, Multi-Scale, Interactive Data Management for Heterogeneous Users: The IsoGenieDB

Suzanne Hodgkins[1†], Benjamin Bolduc[1†], Matthew Sullivan[1*] (sullivan.948@osu.edu), Ruth Varner[2], The IsoGenie Team[3], and **Scott Saleska**[4], **Virginia Rich**[1*] (rich.270@osu.edu)

[1]Ohio State University, Columbus; [2]University of New Hampshire, Durham; [3]10 co-investigator institutions; [4]University of Arizona, Tucson. [†]These authors contributed equally.

https://isogenie-db.asc.ohio-state.edu/

Project Goals: The objective of the IsoGenie3 Project is to discover how microbial communities mediate the fate of carbon in thawing permafrost landscapes under climate change. We are engaged in a systems approach integrating (a) microbial and viral ecology, (b) organic chemistry and stable and radiocarbon isotopes, and (c) state-of-the-art modeling, across an interconnected system of thawing permafrost and post-glacial lakes in Arctic Sweden. To facilitate interdisciplinary collaboration in the IsoGenie Project, and to ensure broader community access to its near-decade of system-scale datasets, the IsoGenie Database (IsoGenieDB) integrates these datasets into a cohesive graph database framework. IsoGenieDB is not only a data repository for the project, providing both public and private access, but it also allows custom querying to explore connections between different datasets. The code base for building the database is open access and can be adapted by other projects to suit their own unique needs.

Organization of disparate environmental data types isn't a new challenge, with storage systems designed for this purpose including ESS-Dive, NEON, LTER, Pangaea, DataDryad, DataONE, etc. These systems are powerful and have their respective strengths, but they do not generally integrate data into frameworks that can be queried in a unified, detailed way. Doing so would require assimilation of data from different sources, which may exist in different formats and involve differences in calibration methods, into a single framework. This framework's structure would ideally reflect that of the system itself, in that the relationships connecting research sites, sampling locations, and different assay methods are recapitulated in the data structure. Such a structure would move beyond data organization, to facilitate exploration of ecological, organismal, and physicochemical interactions occurring both horizontally (between different systems at the same scale) and vertically (across scales).

We sought to address this data integration challenge for the IsoGenie Project by developing the IsoGenie Database (IsoGenieDB), a novel data management and exploration platform. IsoGenieDB has been built as a Neo4j graph database (the same platform used by eBay, Cisco, NASA, and numerous others, helping ensure its reliable longevity) that leverages the inherent relationships within the data to build the basic framework of the database. The graph data structure is a natural mimic of biological and conceptual relationships. The fundamental design of the database follows a property graph model, where *nodes* serve as the primary unit of

organization. Nodes can have *labels*, which serve as a high-level means of categorizing nodes for fast access and classification, while data is stored within the nodes as a set of key:value *properties*. These properties can contain any numeric or text-based information (e.g., time, temperature, pH), as well as links to flat files that store non-text information. Nodes are connected to other nodes through *relationships* (also known as edges), which store information about how the two nodes are related.

Most of the nodes in IsoGenieDB are organized hierarchically by location and data type, starting with a "root" node representing the whole of Stordalen Mire. This root node is then connected by branching relationships to other nodes representing increasingly specific temporal and/or spatial resolution, and finally to data nodes for each data type existing at any given point in time or space. While this base structure is tree-like, the flexibility of the graph database format allows other conceptual patterns to be traced with other relationship structures. These include the storage of file-level metadata (e.g. file links, version, and quality information) in dedicated metadata nodes with relationships to each of their corresponding data nodes, as well as more complex networks of relationships, such as those describing data processing pipelines.

The code base for building the graph database, written in Python, is designed to be easily re-usable for new datasets with only minimal data cleaning of the contributed data files. This allows for automated building of node networks that illustrate common conceptual patterns in ecological research. For instance, within the base tree structure, each child node automatically inherits identifying metadata properties from its parent or "upstream" node (e.g., a node for a soil sample is automatically labeled with the core name and sampling date). This automation has the additional benefit that common node types are given consistent labels regardless of dataset origin; this allows for greater ease and flexibility of querying the database.

A front-end web portal, accessed at https://isogenie-db.asc.ohio-state.edu/, provides data access for non-coding project members and the general public. This web interface includes general information about the IsoGenie project, a Data Downloads page with metadata and downloads of source data files, a map interface for georeferenced data (e.g. core locations and remote sensing images), a tagged photo repository, and a dynamic querying page.

The flexibility of the IsoGenieDB codebase, and availability of its construction scripts and web portal, allows the database to be adapted and modified by other projects. One of these is the A2A-DB, a database for the NASA-funded Archaea to Atmosphere (A2A) project (PI: Ruth Varner, co-I on the IsoGenie Project), whose goal is to use modeling and remote sensing to upscale carbon cycle feedback findings from five Arctic peatlands (including that examined by IsoGenie) to the pan-Arctic. The A2A-DB is in development but will include a public-facing side like the IsoGenie-DB, and is designed to be further expandable to house data from other related studies.

This study was funded by the Genomic Science Program of the United States Department of Energy Office of Biological and Environmental Research, grants DE-SC0010580 and DE-SC001644. Additional funding for the expanded A2A-DB was provided by the NASA Interdisciplinary Studies in Earth Science program (Award # NNX17AK10G).

Genome-resolved Insights Into Permafrost Thaw Processes

Gene Tyson[1], Ben Woodcroft[1], Caitlin Singleton[1], Miguel Martinez[2], Dylan Cronin[2], The IsoGenie Project Team[3], Eoin Brodie[4], Ruth Varner[5], Jeff Chanton[6], Patrick Crill[7], Matt Sullivan[2], **Scott Saleska[8], and Virginia Rich*[2]** (rich.270@osu.edu)

[1]Australian Centre for Ecogenomics, University of Queensland, Brisbane; [2]The Ohio State University, Columbus; [3]10 IsoGenie Project Institutions; [4]Berkeley Labs, Berkeley, California; [5]University of New Hampshire, Durham; [6]Florida State University, Tallahassee; [7]Stockholm University, Stockholm; [8]University of Arizona, Tucson.

http://isogenie.osu.edu

Project Goals: The objective of the IsoGenie3 Project is to discover how microbial communities mediate the fate of carbon in thawing permafrost landscapes under climate change. We are engaged in a systems approach integrating (a) molecular microbial and viral ecology, (b) molecular organic chemistry and stable and radiocarbon isotopes, and (c) state-of-the-art modeling, across an interconnected system of thawing permafrost and post-glacial lakes in Arctic Sweden.

Our IsoGenie Project examines microbial carbon cycling and climate feedbacks along a natural permafrost thaw gradient in Arctic Sweden, where permafrost-underlain palsas sink into partially-thawed bogs and fully-thawed fens, within a mosaic landscape of interconnected post-glacial lakes. Using an ecosystems approach with nearly a decade of high-resolution characterization of microbes, organic matter and carbon gas isofluxes, our team has mapped how thaw's simultaneous impacts on hydrology and vegetation lead to different microbial communities and processes.

In the seasonally-thawed active layer, microbial diversity and cell numbers increase overall from palsa to bog to fen, with different 'type' communities present at each stage of thaw. Metagenomic recovery of genomes for ~60% of the mire's microbial genera has illuminated phylogenetic and functional novelty, bringing genomic representation to many phyla and clades previously lacking any (including the globally relevant methanotroph lineages USCα), connecting new functionality to known lineages (such as the encoded potential for methanotrophy in the Hyphomicrobiaceae), and expanding the phylogenetic breadth of biogeochemically relevant enzymes (such as the presence of an active 'fungal' pathway for xylose degradation in bacterial MAGs). Beyond novelty, the MAGs also revealed thaw-associated shifts in the web of carbon degradation that fuels microbial carbon gas emissions. The dominant microbial lifestyle encoded by the mire's MAGs was the breakdown of polysaccharides, which are the primary source of both carbon and energy, as inorganic terminal electron acceptors are rare. While the dissolved organic matter itself becomes more labile and less diverse in the fully-thawed habitat (see companion geochemistry poster, Wilson et al.), the microbiota encoding its breakdown becomes more diverse, shifting from dominance in the MAGs of genes for complex polysaccharide degradation towards that of simpler polysaccharides, along with fermentation, acetogenesis, and methanogenesis. Concomitantly, the lineages encoding these processes become more diverse.

Methane emissions provide a potent positive feedback to climate warming from these sites, and are proximately controlled by methane cyclers, whose presence, composition, and activity are dramatically impacted by thaw. Methanogens are absent from the dry palsas, and become more abundant, diverse, and active from the bog to the fen, in parallel with increasing methane emissions. High-affinity methanotrophs in the palsa drive slight atmospheric methane uptake, then peak in abundance in the bog and mitigate emissions (as inferred by isotopes), before becoming more diverse and physiologically active in the fen, but playing a lesser role in mitigating emissions. Demonstrating the importance of more explicit microbial representation in models, a single novel genus of methanogens is the best predictor - better than typical environmental drivers or various guild-level aggregations of the methanogens - of the bog methane's isotopic signature (a characteristic used to partition sources of atmospheric CH_4) over multiple years and methods.

Below the seasonally-thawed active layer at the depth of the permafrost itself, thaw also increases microbial diversity, and drives development of a community distinct from those of both the initial permafrost and the overlying seasonally-thawed 'active layer'. Of particular note is a deeply-branching lineage related to the phylum Caldiserica (novel at the level of phylum or class by 16S or genome-based taxonomies, respectively), which rises to ~60% of the community (by metagenome recruitment) in the permafrost itself, has been seen at similarly

high abundances at another Mire nearby, and is present at other global high-carbon sites. Via inferences from 7 metagenome-assembled genomes, this lineage is a fermenter of carbohydrates and possibly amino acids, that can use labile plant compounds and peptides, and encodes adaptations to low temperature. Meta- transcriptome and -proteome analyses indicate that this lineage is active after thaw, but its role as the dominant member of the permafrost microbiome is not yet known.

As microbes shift with thaw, so do viruses (see companion virus poster, Sun et al.), which infect a range of carbon-cycling hosts including methanogens and methanotrophs, based of *in silico* virus-host linkages. Viral diversity loosely tracks microbial diversity across the thaw stages, peaking in the fen. The mire's viruses also carry at least one active 'host' gene responsible for carbon substrate transformation.

In the lakes, which are responsible for half the system's methane emissions, and are the main portal for 'old carbon' loss (whereas mire's methane is from recently fixed 'new' carbon), differences in resident methanogens' temperature sensitivity highlights the importance of understanding acclimation to accurately predict future fluxes.

IsoGenie's MAGs and extensive multi-disciplinary data are housed within a neo4j graph database with a map-based user interface (at http://isogenie.osu.edu; see companion database poster, Hodgkins et al.). By connecting these datasets, and the arising ecosystem meta-omics insights, to a hierarchy of models (including explicit representation of MAG-derived information), IsoGenie is endeavoring to improve predictions of climate feedbacks from these rapidly thawing systems.

This study was funded by the Genomic Science Program of the United States Department of Energy Office of Biological and Environmental Research, grants DE-SC0010580 and DE-SC0016440.

Role of Viruses in Carbon Processing Along a Permafrost Thaw Gradient

Christine L. Sun[1], Matthew B. Sullivan[1]* (sullivan.948@osu.edu), Joanne Emerson[1^], Gareth Trubl[1&], Dylan Cronin[1], Ahmed Zayed[1], Ben Woodcroft[2], Suzanne Hodgkins[1], Yueh-Fen Li[1], Rachel Wilson[3], The IsoGenie Project Coordinators[1], Jeff Chanton[3], Patrick Crill[4], Gene Tyson[2], Scott Saleska[5], **Virginia Rich**[1]

[1]The Ohio State University, Columbus; [2]Australian Centre for Ecogenomics, University of Queensland, Brisbane [3]Florida State University, Tallahassee; [4]Stockholm University, Stockholm; [5]University of Arizona, Tucson; [^]Current address: University of California at Davis, Davis; [&]Current address: Lawrence Livermore National Labs, Livermore, California.

http://isogenie.osu.edu

Project Goals: The objective of the IsoGenie3 Project is to discover how microbial communities mediate the fate of carbon in thawing permafrost landscapes under climate change. We are engaged in a systems approach integrating (a) molecular microbial and viral ecology, (b) molecular organic chemistry and stable and radiocarbon isotopes, and (c) state-of-the-art modeling, across an interconnected system of thawing permafrost and post-glacial lakes in Arctic Sweden.

Permafrost accounts for roughly 30%-50% of global soil carbon (C). Due to elevated temperatures resulting from climate change, permafrost is thawing. It is not fully understood how thawing permafrost will impact carbon dioxide or methane release, which in turn could accelerate climate change. Our prior work has demonstrated that microbial communities play a critical role the C cycle in these permafrost soils[1,2,3] as has also been revealed in other permafrost systems[4]. Less well-studied is the role that viruses infecting these microbes might play in C cycling, though viruses are abundant across diverse soils, including permafrost[5]. If marine systems are a guide, then these abundant viruses are likely to impact carbon cycling by controlling microbial communities via predation, transferring genes from one host to another, and metabolically reprogramming their host cells via regulatory take-over and directly encoding auxiliary metabolic genes (AMGs). While relatively well-studied in marine systems, information about viruses in soils like permafrost is only now emerging.

We previously employed two approaches to explore viral roles along a permafrost thaw gradient in northern Sweden. First, we optimized viral capture methods to be able to directly study quantitatively-amplified metagenomes derived from the viral particles that could be resuspended from soils. These efforts have led to increasingly efficient viral capture methods[5] and sequence datasets that revealed ~50 viral populations from 7 viral metagenomes, which were predominantly novel (species and genera), distinctly distributed along the thaw gradient, and contained 30 AMGs that likely directly impact C metabolism and soil organic matter degradation[6]. Second, we optimized identification of viral contigs from >200 bulk soil metagenomes along the permafrost thaw gradient collected in 2010-2012[7]. Whole genomes and large genome fragments from nearly 2,000 viral populations were identified, of which ~60% of were determined to be active, and also placed into ecological context to reveal soil-type preferences, variations in lineage-specific virus-host pressures, and that most key C cyclers in the system are infected by viruses. Functional analyses of one of the detected virus-encoded glycoside hydrolase genes further

demonstrated that these AMGs are functional, and offer a clear example of how viruses can directly impact C degradation, here converting complex C to simpler C forms.

Building upon these prior successes, we are now analyzing the viral populations present in an expanded bulk soil metagenomic dataset that extends the permafrost thaw time series out to include the years 2013-2017. The resulting sequencing reads were assembled into contigs using an updated assembly algorithm (metaSPADES), and viruses were identified using an updated suite of approaches (VirSorter and DeepVirFinder) optimized from our work with marine metagenomes. These efforts alone have already revealed thousands of viral populations, though co-assemblies and genome binning are also being evaluated to increase viral population detection. These viral populations, combined with already-available reference genomes, will be used to further our understanding of the role and impact of viruses in permafrost.

References:

1. Woodcroft, B. J., Singleton, C. M., Boyd, J. A., Evans, P. N., Emerson, J. B., Zayed, A. A. F., Hoelzle, R. D., Lamberton, T. O., McCalley, C. K., Hodgkins, S. B., Wilson, R. M., Purvine, S. O., Nicora, C. D., Li, C., Frolking, S., Chanton, J. P., Crill, P. M., Saleska, S. R., Rich, V. I., & Tyson, G. W. (2018). Genome-centric view of carbon processing in thawing permafrost. Nature, 560(7716), 49–54.
2. Singleton, C. M., McCalley, C. K., Woodcroft, B. J., Boyd, J. A., Evans, P. N., Hodgkins, S. B., Chanton, J. P., Frolking, S., Crill, P. M., Saleska, S. R., Rich, V. I., & Tyson, G. W. (2018). Methanotrophy across a natural permafrost thaw environment. The ISME Journal, 1–15.
3. McCalley, C. K., Woodcroft, B. J., Hodgkins, S. B., Wehr, R. A., Kim, E.-H., Mondav, R., … Saleska, S. R. (2014). Methane dynamics regulated by microbial community response to permafrost thaw. Nature, 514(7523), 478–481.
4. Schuur, E. A. G., McGuire, A. D. , Schädel, C. , Grosse, G., Harden, J. W. ,Hayes, D. J., Hugelius, G., Koven, C. D. , Kuhry, P. , Lawrence, D. M., Natali, S. M. , Olefeldt, D., Romanovsky, V. E., Schaefer, K. , Turetsky, M. R. , Treatm C. C., Vonk, J. E. (2015) Climate change and the permafrost carbon feedback. Nature 520, 171–179.
5. Trubl, G, Solonenko, N., Chittick, L., Solonenko, S., Rich, V.I., & Sullivan, M.B. 2016. Optimization of viral resuspension methods for carbon-rich soils along a permafrost thaw gradient. *PeerJ.* 4:e1999.
6. Trubl, G., Jang, H. B., Roux, S., Emerson, J., Solonenko, N., Vik, D., Solden, L., Ellenbogen, J., Runyon, A., Bolduc, B., Woodcroft, B., Saleska, S., Tyson, G., Wrighton, K., Sullivan, M.B. & Rich, V. 2018. Soil viruses are underexplored players in ecosystem carbon processing. *mSystems. 3.5 (2018): e00076-18.*
7. Emerson, J.B., Roux, S., Brum, J.R., Bolduc, B., Woodcroft, B.J, Jang, H.B., Singleton, C.M., Solden, L.M., Naas, A.E., Boyd, J.A., Hodgkins, S.B., Wilson, R.M., Trubl, G., Li, C., Frolking, S., Pope, P.B., Wrighton, K.C., Crill, P.M., Chanton, J.P., Saleska, S.R., Tyson, G.W., Rich, V.I. & Sullivan, M.B. 2018. Host-linked soil viral ecology along a permafrost thaw gradient. *Nature Microbiology*. 3:870–880.

This study was funded by the Genomic Science Program of the United States Department of Energy Office of Biological and Environmental Research, grants DE-SC0010580 and DE-SC0016440, as well as a Gordon and Betty Moore Foundation Award (#3790) to enable the bulk of the virus-specific work.

Integrating Metabolomics and Proteomics Reveals Biotic and Abiotic Controls on CO_2 and CH_4 Production in Peatlands

Rachel M. Wilson[1*] (rmwilson@fsu.edu), Malak M. Tfaily[2], Yueh-Fen Li[3], Alexandra Cory[1], Ben Woodcroft[4], Suzanne Hodgkins[3], Patrick Crill[6], Gene Tyson[4], Scott Saleska[2], **Jeff Chanton[1],** and **Virginia Rich[3]**

[1]Florida State University, Tallahassee; [2]University of Arizona, Tucson; [3]The Ohio State University, Columbus; [4]Australian Centre for Ecogenomics, University of Queensland, Brisbane; [6]Stockholm University, Stockholm

http://isogenie.osu.edu

Project Goals: The objective of the IsoGenie3 Project is to discover how microbial communities mediate the fate of carbon in thawing permafrost landscapes under climate change. We are engaged in a systems approach integrating (a) molecular microbial and viral ecology, (b) molecular organic chemistry and stable and radiocarbon isotopes, and (c) state-of-the-art modeling, across an interconnected system of thawing permafrost and post-glacial lakes in Arctic Sweden.

Arctic peatlands contain a large reservoir of organic C (OC) that may become more available for microbial decomposition upon permafrost thaw and peat warming. Microbial decomposition of peat OC results in the production of CO_2 and CH_4 which may be emitted to the atmosphere, exacerbating climate warming. A decade of radiocarbon dating in non-permafrost and discontinuous permafrost peatlands suggests that CO_2 and CH_4 are produced from dissolved phase OC that is much younger than the solid peat. Even when CO_2 and CH_4 production are observed to increase in response to warming, radiocarbon evidence suggests that peat OC is not mobilized to fuel this increased production. To determine what is fueling the enhanced production in thawing peatlands, we have utilized a suite of complementary environmental metabolomics and proteomics datasets. This approach has highlighted the previously unrecognized importance of abiotic reactions for producing CO_2 in peat. Sterilized incubations revealed the abiotic reaction of N-compounds which can subsequently undergo decarboxylation to produce CO_2, and may also immobilize N through the creation of condensed N-containing rings. Intermediates of biotic reactions, such as de/nitrification, have been shown to participate in similar abiotic transformations suggesting that the biotic and abiotic cycles are intricately linked and may influence CO_2 and CH_4 production in peatlands in complex ways. While abiotic reactions are likely to increase in response to warming (kinetic control), the response of biotic reactions to climate change is complicated by the response of the organisms to warmer temperatures, both their current enzymatic temperature sensitivity and the relative importance of acclimation, adaptation, and assembly processes in changing community-level enzymatic behavior, and therefore may prove more difficult to predict without further characterization of the transformations and interaction with the abiotic cycle.

This study was funded by the Genomic Science Program of the United States Department of Energy Office of Biological and Environmental Research, grants DE-SC0010580 and DE-SC0016440.

Introgression of Novel Disease Resistance Genes from *Miscanthus* into Energycane

Alexander E. Lipka[1], Marcus O. Olatoye[1], Nicholas R. LaBonte,[1] Jack C. Comstock[2], Sushma Sood[3], Vanessa S. Gordon[3], Per McCord[3], Keiji Motomura[4], Ken-Taro Sekine[4], Toshihiko Yamada[5], and **Erik J. Sacks**[1*] (esacks@illinois.edu)

[1]University of Illinois, Urbana, IL; [2]Department of Plant Pathology, University of Florida, Gainsville, FL ; [3]USDA-ARS Sugarcane Field Station, Canal Point, FL; [4]University of the Ryukyus, Okinawa, Japan; [5]Hokkaido University, Sapporo, Japan

Project Goals: Our long-term goal is to improve energycane productivity and sustainability by providing resistance to key diseases with novel genes from *Miscanthus*.

1) **In miscane BC_1 populations (sugarcane x (sugarcane x *Miscanthus*)), identify molecular markers associated with novel genes from *Miscanthus* that confer resistance to at least two out of four of the following economically important diseases of sugarcane: ratoon stunt, yellow leaf, orange rust, and smut.**
2) **Compare effectiveness of different molecular marker analysis methods for selecting disease resistance alleles in miscane backcross populations. In particular, compare the pseudo-testcross QTL mapping strategy with genomic selection.**
3) **Screen germplasm collections of *M. sinensis* and *M. sacchariflorus* for resistance to ratoon stunt, yellow leaf, orange rust, and smut to confirm that *Miscanthus* is uniformly immune to these diseases as prior data suggest or to quantify genetic variation for resistance if not all accessions are resistant.**

Sugarcane is among the world's leading bioenergy crops. Modern sugarcane cultivars are derived from a relatively small set of founder genotypes, which has contributed to cultivar susceptibility to diseases. Modern sugarcane cultivars originated in the late 1800's when cultivars of *Saccharum officinarum* with high sugar yield potential but disease-susceptibility were crossed with the undomesticated, non-sugar producing but disease-resistant *S. spontaneum*, then backcrossed to *S. officinarum* to recover sugar yield. Recent efforts to improve sugarcane for disease resistance, pest resistance, and abiotic stress tolerance have continued to rely primarily on introgressions from *S. spontaneum*. Similarly, energycanes (sugarcanes bred specifically for energy) are also bred from modern sugarcane cultivars and *S. spontaneum*.

Yet, diseases are the primary constraints to cane productivity in commercial fields. Sugarcane is a perennial crop but a planting's commercial productivity in the U.S. is typically limited to 3-5 years because disease pressure reduces yield and decreases stand via plant death each year. Moreover, the utility of a successful cane cultivar for commercial production is typically ended by the emergence of a virulent strain of a common pathogen, resulting in reduced yields. Though many diseases affect sugarcane, the following four are of especially great concern (Rice, 2007): 1) ratoon stunt (bacterium *Leifsonia xyli* subsp. *xyli*.); 2) sugarcane yellow leaf (virus, Sugar Cane Yellow Leaf Virus (SCYLV); 3) orange rust (fungus *Puccinia kuehnii*); and, 4) smut

(fungus *Sporisorium scitamineum*). Arguably, *Miscanthus* would be a better source of genes for improving sugarcane than *S. spontaneum* because the former is highly resistant to diseases and pests, is more broadly adapted to diverse environments, and is more genetically distant from *S. officinarum* (thus providing more novel alleles). Previously, we have obtained more than a dozen F_1 hybrids between sugarcane and *Miscanthus*.

We conducted disease screenings on miscane F_1 progeny, and core collections of *M. sinensis* and *M. sacchariflorus*. For the most part, our hypotheses of ubiquitous resistance in *Miscanthus* were confirmed. For smut, all of the 66 *Miscanthus* genotypes tested were fully resistant, and for orange rust all but one of the *Miscanthus* genotypes tested were fully resistant. For ratoon stunt disease, only two *Miscanthus* genotypes were susceptible and these were much less susceptible that the sugarcane positive control. Response of *Miscanthus* to SCYLV was more variable than for the other diseases tested, with 14/31 *M. sacchariflorus* and 24/35 *M. sinensis* fully resistant, but the remainder were partially to fully susceptible. Of six miscane genotypes tested, one was fully resistant to all four diseases, six were fully resistant to orange rust and smut, four were resistant to ratoon stunt, and three were resistant to SCYLV.

Backcross progeny (BC_1F_1) were obtained for two populations (n = 210 and n = 30) and planted in the field at Canal Point, FL in spring of 2018. In 2019, the BC_1F_1 progeny will be phenotyped for disease-resistance and genotyped. Both QTL selection and genomic selection will be conducted on the BC_1F_1 and the relative effectiveness of these methods compared by evaluating BC_2F_1 progeny of the selected BC_1F_1s crossed to a susceptible cane. To increase the size of the smaller population, we have made additional crosses in the greenhouse to obtain more BC_1F_1 seed, and we are also self-pollinating the BC_1F_1s in the field.

To predict the relative efficiencies of QTL selection and genomic selection under different genetic architectures, we conducted simulation studies in F_1 and subsequent BC_1 populations, using marker data and linkage maps from previous studies of *Miscanthus* and sugarcane. For traits with four or fewer QTLs, we observed that the performance of QTL selection was comparable and sometimes superior to genomic selection. In contrast, as the number of simulated QTL increased, all four GS models that were evaluated outperformed QTL selection. Thus, we expect genomic selection to be preferable to QTL selection for introgressing genetic sources of horizontal disease resistance from *Miscanthus* to energycane, whereas QTL selection remains a suitable option for introgressing vertical disease resistance.

Because disease susceptibility is a major limitation for cane production, cultivar durability, and sustainability, the introgression of resistance genes from *Miscanthus* is expected to increase the economic and environmental benefits of energycane while reducing costs and risks. Such benefits should promote further investment by industry in energycane.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0016264.

Next-generation Synthetic Biology Technologies for Controlling Metabolism in Clostridia

Daniel Cetnar[1], Michael Koepke[2], Steve Brown[2], **Howard Salis[1]*** (salis@psu.edu)

[1]Pennsylvania State University, University Park, PA; [2]LanzaTech Inc., Skokie, Illinois

https://salislab.net https://www.lanzatech.com

Project Goals: We are harnessing next-generation Synthetic Biology technologies for rationally controlling metabolic flows in *Clostridium autoethanogenum*, an organism capable of converting syngas feedstock into valuable chemical products.

Clostridium autoethanogenum is non-model organism capable of converting low-cost C1 feedstock (e.g. carbon dioxide and carbon monoxide from waste gas streams) into valuable chemical products. Here, we leverage recently developed Synthetic Biology technologies to improve our ability to rationally control metabolic flows in *Clostridium autoethanogenum* with the overall goal of reducing byproduct formation, diversifying the bioprocess' product portfolio, and improving production titers. These technologies include: (1) 4600 highly non-repetitive promoters with well-characterized transcription rates (across a 1.4-million range) that can all be used simultaneously without triggering genetic instability; (2) rationally designed operons expressing multi-enzyme pathways with optimized expression levels for maximal productivities; and (3) Extra Long sgRNA Arrays co-expressing 20+ sgRNAs to knock-down the expression of enzymes responsible for byproduct formation. In each case, we previously developed and experimentally validated automated algorithms to design these genetic parts and systems to achieve desired functionalities, while minimizing sources of genetic instability (e.g. the Operon Calculator, ELSA Calculator, and the Non-Repetitive Parts Calculator). The first application of these technologies is the design and characterization of over 10000 highly non-repetitive AT-rich promoters for use in *Clostridia* and other organisms with AT-rich genomes, leveraging oligopool synthesis and massively parallel reporter assays.

This project is funded by the Department of Energy, Biological Systems Science Division, SC-23.2, in the Office of Biological and Environmental Research (DE-SC0019090).

Title: Leaf Carbon and Nitrogen Isotope Composition in Diverse Sorghum Lines Under Differential Water and Nitrogen Treatments

Asaph B. Cousins[1]* (acousins@wsu.edu), Balasaheb V. Sonawane[1], Max Braud[2], Indrajit Kumar[2], Rajiv Parvathaneni[2], Andrea Eveland[2], Kira Veley[2], Jeffrey Berry[2], Rebecca Bart[2], Varsha Pathare[1], Stephanie Futrell[4], Ellen Marsh[4], Yen Ning Chai[4], Peng Liu[3], Cody Creech[4], Ismail Dweikat[4], Stephen Kresovich[5], Allyn Pella[4], Ellen Marsh[4], Peng Wang[4], **Daniel P. Schachtman**[4]

[1]Washington State University; [2]Donald Danforth Plant Science Center; [3]Iowa State University; [4]University of Nebraska-Lincoln; [5]Clemson University

http://sorghumsysbio.org

Project Goals:
The *overall project goal* is to establish a foundational, systems-level understanding of plant, microbial, and environmental interactions that will lead to strategies for enhancing growth and sustainability of sorghum through genetic and microbial adaptations to water and nitrogen limited environments.

The ***specific objectives*** of the research presented here are to:
1. Conduct phenotypic characterizations of a diverse panel of sorghum genotypes to define photosynthetic and isotope response under water and nitrogen limited environments.
2. Test sorghum genotype by water and nitrogen limited environment interactions in both controlled environment and field growth conditions.
3. Determine if measurements of leaf carbon and nitrogen isotope composition can screen for differences in water and nitrogen use efficiency in diverse sorghum genotypes.

Towards achieving our project goals, we have conducted several phenotyper, greenhouse and field experiments. Our initial phenotyper experiment in 2016 was to screen 30 diverse sorghum lines under controlled environment growth conditions in the Bellweather Phenotyping System at the Danforth Center. This population included 18 energy, 2 grain and 10 sweet sorghum lines. Under the phenotyping system two separate experiments were conducted using a random block design to study the growth, photosynthetic and stable isotope response of this diverse sorghum panel to changes in nitrogen and water availability. Additionally, whole plant nitrogen and water use efficiency were estimated from the phenotyping data. Towards the end of both experiments the upper most fully expanded leaf from individual plant was used for gas exchange measurements with a LI-6400XT open gas exchange system (Li-COR Biosciences, Inc. Lincoln, NE). A portion of the same leaf was sampled for nitrogen and carbon isotope composition ($\delta^{15}N$ and $\delta^{13}C_{leaf}$, respectively).

For these controlled environment experiment data will be presented on differences in nitrogen and water use efficiency, rates of CO_2 assimilation, stomatal conductance, leakiness, intrinsic TE_i and photosynthetic nitrogen use efficiency (PNUE). Measurements of total leaf C/N content and $\delta^{13}C_{leaf}$ and $\delta^{15}N$ will also be presented.

We also collected leaf samples for $\delta^{15}N$ and $\delta^{13}C_{leaf}$ analysis from field experiments in Nebraska in both 2016 and 2017. The western Nebraska field site (Scottsbluff) was used to study response to water stress conditions and the eastern location (Central City) was used to characterize response to low nitrogen. Initial characterization of these data sets suggests significant variation between the sorghum genotypes in their responses to nitrogen and water availability.

To follow up on this previous work, we used a greenhouse experiment to determine the relationship of whole plant water use efficiency (WUE_{plant}) with intrinsic transpiration efficiency (TE_i), defined as the rate of CO_2 assimilation (A_{net}) relative to water loss *via* stomatal conductance (g_s), and $\delta^{13}C_{leaf}$ in sorghum. The TE_i has been considered as a major component of WUE_{plant}. Further, theoretical models in C_4 plants have demonstrated that $\delta^{13}C_{leaf}$ is related to TE_i, when efficiency of CO_2 concentrating mechanism (leakiness) remains constant. Accordingly, $\delta^{13}C_{leaf}$ has been proposed as a high-throughput phenotyping tool for TE_i in C_4 plants. However, there is inadequate information about how leakiness responds to water stress and therefore limits the application of $\delta^{13}C_{leaf}$ for TE_i and thereby WUE_{plant} screening in C_4 crops. The aim of these experiments was to determine response of leakiness to short- or long-term water stress and to revisit the relationship of $\delta^{13}C_{leaf}$ with TE_i and WUE_{plant} in the C_4 bioenergy sorghum line (Grassl). Our results demonstrated that the leakiness is not responsive to short- or long-term water stress. Yet, $\delta^{13}C_{leaf}$ was uncorrelated with TE_i under short- and long-term water stress conditions, whereas $\delta^{13}C_{leaf}$ showed a significant negative relationship with WUE_{plant}. This suggests that in contrast to WUE_{plant}, the steady-state measurements of TE_i do not capture time-integrated responses to water stress. The fact that leakiness is not responsive under water stress suggests that the time-integrated signal of $\delta^{13}C_{leaf}$ can be used as a phenotyping tool for $\delta^{13}C_{leaf}$ in this bioenergy sorghum.

To scale up our results on the relationships between $\delta^{15}N$ and $\delta^{13}C_{leaf}$ with leaf nitrogen and water use efficiency, respectively, we conducted a nitrogen limitation and water stress experiment with three diverse genotypes at Greenhouse Innovation Complex Phenotyper Facility, UNL. The goal of this work was to establish mechanistic link between $\delta^{15}N$ and nitrogen use efficiency (NUE) and scale $\delta^{13}C_{leaf}$ from additional sorghum lines to WUE_{plant}. Our results suggested significant variation between the sorghum genotypes and their responses to nitrogen uptake and use efficiency as well as differences in WUE_{plant}.

Future directions

Leaf level and whole plant traits will be assessed across genotypes in response to both changes in nitrogen and water availability. This information will be analyzed in comparison to field grown material to help identify and select for genomic traits and potentially elite lines for enhanced nitrogen and water use efficiency in sorghum.

Funding statement

This project is funded by the DOE BER Sustainable Bioenergy Research Program, Award DE-SC0014395, and was also supported by DOE JGI Community Science Program.

Defining the Functional Genome Associated with Enhanced Water and Nitrogen Use Efficiency in Bioenergy Sorghum

Andrea L. Eveland[1*] (aeveland@danforthcenter.org), C. Maxwell Braud,[1] Rajiv Parvathaneni,[1] Indrajit Kumar,[1] Balasaheb V. Sonawane[2], Scott Lee[1], Philip Ozersky[1], Ellen Marsh[3], Todd C. Mockler[1], Asaph B. Cousins[2] and Daniel Schachtman[3]

[1] Donald Danforth Plant Science Center, St. Louis, MO; [2] Washington State University, Pullman, WA; [3] University of Nebraska, Lincoln, NB

http://sorghumsysbio.org

Project Goals:

- **Overall project objective: Establish a foundational, systems-level understanding of plant, microbial, and environmental interactions that will lead to translational strategies to enhance growth and sustainability of sorghum through improved genetic and microbial adaptations to water and nutrient limited environments.**
- **Perform controlled environment, high-throughput phenotyping at multiple scales to investigate responses of diverse sorghum lines to low water and low nitrogen.**
- **Generate RNA-seq-based transcript profiles and genome-wide accessible chromatin maps for developing leaf and root tissues across genotypes in response to water and N stresses.**
- **Integrate transcriptional, epigenomic, and physiological signatures with existing genetics and genomics resources to prioritize genetic loci for functional validation.**

The overall goal of this work is to establish a functional genomics resource that can be leveraged for investigating effects of low water and nutrient inputs on growth and development of three diverse sorghum accessions; BTx623 (reference line, NAM parent, EMS population), Grassl (bioenergy, NAM parent, EMS population) and PI_510757 (high NUE, WUE and large biomass, NAM parent). These lines are representatives of three diverse races of sorghum (Kafir, Caudatum and Durra, respectively) and thus will serve as founders for deeper explorations of the genetic diversity present in sorghum. Here, we investigate morphological, physiological, and molecular responses of sorghum diversity to low water and low nitrogen (N) using controlled environment phenotyping. Transcriptome profiles analyzed across emerging leaf and root samples were aligned with respective chromatin accessibility generated using ATAC-seq. Together with the exhaustive imaging data collected, these genomics data are being integrated to identify physiological, transcriptional, and/or epigenetic signatures associated with enhanced water and N use efficiency. An integrated resource that utilizes the same founder lines (i.e. three diverse sorghum lines), tissues and developmental time series will **i)** improve resolution of our gene regulatory networks, **ii)** provide a foundation for a sorghum pan-genome ENCODE-like project to identify regulatory regions across sorghum races and **iii)** identify targets for engineering enhanced biomass, NUE and drought tolerance.

Storage, Integration and Dissemination of data for a multi-disciplinary project.

Philip Ozersky[1] (pozersky@danforthcenter.org), Jeff Berry[1], Todd Mockler[1], and **Daniel Schachtman**[2]

[1]Danforth Plant Science Center, St. Louis, MO; [2]University of Nebraska, Lincoln;

Website for the project: www.sorghumsysbio.org

Overall Project Goals:

- Establish a foundational, systems-level understanding of plant, microbial, and environmental interactions that will lead to translational strategies to enhance growth and sustainability of sorghum through improved genetic and microbial adaptations to water and nutrient-limited environments.

Project Objectives:

- Conduct deep census surveys of root microbiomes concurrent with phenotypic characterizations of a diverse panel of sorghum genotypes across multiple years to define the microbes associated with the most productive lines under drought and low nitrogen conditions.
- Associate systems-level genotypic, microbial, and environmental effects with improved sorghum performance using robust statistical approaches.
- Develop culture collections of sorghum root/leaf associated microbes that recapitulate root-enriched sequences defined in the census.
- Perform controlled environment experiments for in-depth characterization and hypothesis testing of $G_{sorghum}$ x $G_{microbe}$ x E interactions.
- Validate physiological mechanisms, map genetic loci for stress tolerance, and determine the persistence of optimal microbial strains under greenhouse and field conditions.

This multiple disciplinary project spans a wide array of data types, ranging from field-grown plant measurements to indoor greenhouse phenotypic analysis and images, and from metabolomic analysis to metagenomic sequences and OTU counts. Implementation of a website, database, and interface allows all project members to access data, coordinate efforts and merge datasets for more powerful uses of the data as a whole. We have developed a fluid database to store the data and designed a Shiny app to allow project members to mine and parse data for downstream analysis. We have designed group-specific Shiny apps to produce "on the fly" analyses. Our poster will demonstrate the website, database structure, interface utility and tools produced for this project.

This project is funded by the DOE BER Sustainable Bioenergy Research Program, Award DE-SC0014395, and was also supported by DOE JGI Community Science Program.

The Metabolome of Early Season Sorghum Plant Tissue is Predictive of End of Season Biomass

Sheflin, Amy[1]; Prenni, Jessica E.[1]* (jprenni@colostate.edu); Kresovich, Stephen[2]; Dweikat, Ismail[3]; Marsh, Ellen[3]; and **Schachtman, Daniel**[3]

[1]Colorado State University, Fort Collins, CO; [2]Clemson University, Clemson, SC; [3]University of Nebraska – Lincoln, Lincoln, NE

https://sorghumsysbio.org/

Project Goals:

The overall goal of this project is to establish a foundational, systems-level understanding of plant, microbial, and environmental interactions that will lead to translational strategies to enhance growth and sustainability of sorghum through improved genetic and microbial adaptations to water and nutrient limited environments. Specific objectives that apply to this poster include: (1) Phenotypic characterizations of a diverse panel of sorghum genotypes across multiple years to define the most productive lines under drought and low nitrogen conditions. (2) Associate systems-level metabolomic and environmental effects with improved sorghum performance using robust statistical approaches.

Abstract:

To compete in the biofuel energy market, cellulosic feedstocks will need to be high yielding and carbon neutral or negative while requiring low inputs. To avoid competition with existing food production systems, these crops will also need to be grown on marginal lands. This will require the introduction of novel traits to increase abiotic stresses tolerance associated with marginal soils. Thus, efforts to maximize biomass production requires an understanding of both the genetic and environmental influences on plant growth. Metabolomics is an analytical approach that enables comprehensive profiling of plant metabolism at a given point in time, providing valuable information about how genetic diversity combines with the environment to affect phenotypic variation in plants. While often removed from the visible phenotype, metabolomics allows for a more functional interpretation of plant response to the environment at the molecular level, facilitating investigation of variation within the species, how metabolism affects the phenotype and discovery of new metabolic pathways that influence yield traits. This information is particularly useful when interpreting complex phenotypes such as adaptations to nutrient stress and can be used to select for phenotypes with the metabolic adaptations that maximize yield in marginal environments.

Importantly, the chemical variation in plant tissue can be associated with yield and other traits to create a *chemical profile predictive of phenotypic outcomes*. Furthermore, in the context of breeding efforts, metabolomics can provide valuable data about the adjustment of metabolism in hybrids in comparison to inbreds and has been suggested to contain mechanistic insights about how hybrids boost performance. While predictive modeling for yield phenotypes based on genomic resources has proven highly successful in multiple species, the models are less effective when environmental stresses are a factor. Metabolomics can capture the plant phenotype resulting from the genotype-environment interaction and has emerged as an effective molecular platform for predicting agronomic traits in multiple crops species.

Here, we present the generation of highly accurate predictive models for end of season biomass in field grown sorghum based on metabolite profiles of early season (29 days after planting) root and leaf tissue. Predictive models trained with data from plants grown under abiotic stress (nitrogen deprivation or drought) were more accurate for predicting biomass in plants grown under stress than models trained with data from controls. This result demonstrates the advantage of metabolomics to capture the unique phenotypes resulting from the environmental interaction. Furthermore, depending on the type of abiotic stress, the metabolome data from different tissues (e.g. roots or leaves) was most effective for training the predictive model of the given environment. For example, for plants grown under nitrogen stress, the most accurate models were trained on metabolite data from leaf tissue acquired using a gas chromatography mass spectrometry platform (GC-MS). Evaluation of the metabolites driving this model revealed biologically relevant molecules such as asparagine, aspartic acid, fructose, and maltose. Importantly, these compounds are not readily detectable by reverse phase liquid chromatography (LC-MS) approaches, highlighting the additional key consideration of analytical platform(s) in generating the most appropriate metabolome data for each stress condition. Finally, we also present the potential of the predictive models across fields (controls only). Ultimately, the results presented here demonstrate the effectiveness of metabolite profiles trained on young plants to accurately predict end of season sorghum biomass. In the future, larger scale studies could employ multiple genotypes and inbreds to evaluate the predictive capacity of the metabolome for biomass yield under environmental stress enabling enhanced and accelerated breeding for maximal biomass accumulation in low nutrient environments.

Funding statement. This research was supported by the Office of Science (BER), U.S. Department of Energy (DE-SC0014395).

Traits for Development of Sustainable Energy Sorghum Biofuel Feedstock Production in Marginal Environments

Daniel Schachtman[1] (Daniel.schachtman@unl.edu), Zachary Brenton[2], Cody Creech[1], Ismail Dweikat[1], Stephanie Futrell[1], Yufeng Ge[1], Stephen Kresovich[2], Jiating Li[1], Ellen Marsh[1], Yeyin Shi[1], Hao Zhang[1]

[1] University of Nebraska, Lincoln, NE; [2]Clemson University, SC

Website for project: https://sorghumsysbio.org/

Overall project goals:

- Establish a foundational, systems-level understanding of plant, microbial, and environmental interactions that will lead to translational strategies to enhance growth and sustainability of sorghum through improved genetic and microbial adaptations to water and nutrient limited environments.

Specific objectives addressed by this poster:

- Phenotypic characterizations of a diverse panel of sorghum genotypes across multiple years to define the most productive lines under drought and low nitrogen conditions.
- Associate genotypic and environmental effects with improved sorghum performance using robust statistical approaches.

This sustainable systems project involves both plant genetics and studies of the soil microbial communities associated with sorghum. Part of this group is focused on understanding the relationships between soil microbes, abiotic stress and sorghum genotypes using culture independent methods and by culturing sorghum associated bacteria. Another major area of the grant is the development of energy sorghum germplasm that is responsive to soil microbes and remains highly productive even under conditions of water deficit and low nitrogen.

This poster will provide an update on the characterization of one recombinant inbred population (RIL) segregating for the ability to maintain growth under low nitrogen conditions and the screening of the bioenergy association panel (BAP) [3] under water deficit conditions.

Field work conducted in 2016 and 2017 evaluated 24 – 30 sorghum lines including founder lines of a Nested Association Mapping panel (NAM) [1]. Under sufficient and low nitrogen conditions, lines were evaluated for biomass and specific traits using a hyperspectral radiometer. Drone flights also provided data for biomass [2]. Sufficient variation for growth under low nitrogen conditions was identified in 2017. In 2018 progeny from the parents Grassl (recurrent female parent) X Rio were evaluated for growth on low nitrogen soils and possible gene mapping.

These NAM founder lines were also grown under water limited conditions in 2017, but the variation was not sufficient enough to allow for the use any of the existing RIL populations for further studies. Therefore the (BAP) [3] in which more variation for tolerance to drought was expected was evaluated under water-deficit in 2018.

A primary objective of the work in the 2018 field season for both the low nitrogen and water deficit studies was to measure the relative biomass at the end of the season. However additional information on traits was collected and focus was placed on developing high throughput methods for evaluating germplasm. These traits will complement biomass measurements and be used to understand the physiological mechanisms underlying growth in these marginal environments.

Advanced phenotyping methods as well as brute force approaches, were used to collect data on the RIL population leaf nitrogen, chlorophyll, specific leaf area and whole plant nitrogen use efficiency using a hyperspectral radiometer in 250 progeny of the NAM population from the Grassl X Rio cross. Those results will be presented.

Under drought conditions the BAP was evaluated for traits that have been previously studied in grain sorghum but not energy sorghum that potentially impact drought tolerance. These traits include leaf wax composition [4], canopy temperature, leaf angle [5], specific leaf area and osmoregulation. Those preliminary results will also be presented.

References

1. Boyles, R.E., Z.W. Brenton, and S. Kresovich, *Genetic and genomic resources of sorghum to connect genotype with phenotype in contrasting environments.* Plant J, 2019. **97**(1): p. 19-39.
2. Li, J., Y. Shi, A. Veeranampalayam-Sivakumar, and D.P. Schachtman, *Elucidating sorghum biomass, nitrogen and chlorophyll contents with spectral and morphological traits derived from unmanned aircraft system.* Frontiers in Plant Sciences, 2018. **online/in press**.
3. Brenton, Z.W., E.A. Cooper, M.T. Myers, R.E. Boyles, N. Shakoor, K.J. Zielinski, B.L. Rauh, W.C. Bridges, G.P. Morris, and S. Kresovich, *A Genomic Resource for the Development, Improvement, and Exploitation of Sorghum for Bioenergy.* Genetics, 2016. **204**(1): p. 21-33.
4. Awika, H.O., D.B. Hays, J.E. Mullet, W.L. Rooney, and B.D. Weers, *QTL mapping and loci dissection for leaf epicuticular wax load and canopy temperature depression and their association with QTL for staygreen in Sorghum bicolor under stress.* Euphytica, 2017. **213**:207 - 218
5. Truong, S.K., R.F. McCormick, W.L. Rooney, and J.E. Mullet, *Harnessing Genetic Variation in Leaf Angle to Increase Productivity of Sorghum bicolor.* Genetics, 2015. **201**(3): p. 1229-1238.

This project is funded by the DOE BER Sustainable Bioenergy Research Program, Award DE-SC0014395, and was also supported by DOE JGI Community Science Program.

Sorghum root microbiome dynamics under nutrient-limited and drought conditions

Susannah G. Tringe (sgtringe@lbl.gov),[1*] Dawn Chiniquy,[1] Kyle Hartman,[1] Jinglie Zhou,[1] and **Daniel Schachtman**[2]

[1]Joint Genome Institute, Lawrence Berkeley National Laboratory, Walnut Creek, CA;
[2]University of Nebraska, Lincoln, NE

http://www.sorghumsysbio.org

Project Goals:

We aim to establish a foundational, systems-level understanding of plant, microbial, and environmental interactions that will lead to translational strategies to enhance growth and sustainability of sorghum through improved genetic and microbial adaptations to water and nutrient limited environments. In working towards this goal, we are conducting deep census surveys of root microbiomes concurrent with phenotypic characterizations of a diverse panel of sorghum genotypes across multiple years to define the microbes associated with the most productive lines under drought and low nitrogen conditions.

Abstract:

Sorghum bicolor is a genetically diverse crop cultivated for a variety of agronomic uses, including grain, sugar, and energy production. However, cultivation of energy sorghum for biofuel production will require the use of marginal lands with potentially low nutrient availability and/or periods of water stress. All plants growing in soil harbor diverse communities of microbes that inhabit the areas in, on, and around their roots. Selected members of these microbial communities can provide benefits to their plant hosts, including direct growth promotion and conferring tolerances to abiotic and biotic stress. To examine the effects of nutrient and water stress on soil and root microbial communities and explore a possible microbial solution to increase the nutrient use efficiency and resilience to water stress in sorghum, we are utilizing 16S rRNA sequencing to survey the bacteria communities in replicate soil, rhizosphere, and root samples collected from ~30 different sorghum genotypes grown under different nitrogen (N; high/low) and water (watered/drought) treatments over multiple growing seasons at two sites in Nebraska.

In a small-scale pilot experiment in 2015, we collected ~200 soil, rhizosphere, and root samples from 10 different sorghum genotypes grown under high or low N conditions. Diversity analyses of rhizosphere samples collected early in the growing season show a significantly lower Shannon diversity compared to those collected later in the season. Investigation of taxonomic profiles from the early season rhizosphere samples reveal a dominance of bacteria from the genus *Pseudomonas.* In a second, larger scale field experiment in 2016, we sequenced ~ 4100 samples from all genotypes cultivated under the different N and water treatments. Additionally, we sampled three selected genotypes at four time points throughout the growing season to monitor changes in the

bacteria communities over time. Results from the analysis of the large genotype panel suggest a strong effect of sorghum genotype on bacterial community diversity in rhizosphere and root samples. We also observe a decrease in Shannon diversity in rhizosphere samples at one early time point, which again appears to be driven by a marked increase in the relative abundance of *Pseudomonas*. Most recently, we completed the sequencing of ~3200 samples collected from the 2017 field season. Consistent with the results from 2015 and 2016, initial analyses reveal a marked increase in the relative abundance of *Pseudomonas* bacteria in rhizosphere samples early in the growing season. Additionally, ordination analyses suggest water stress may drive differences in bacteria community composition in rhizosphere and root samples. Our ongoing work focuses on statistical confirmation of our current observations, quantification of experimental treatment effects, and integrating our findings with the metabolomic and phenotypic data being generated with project collaborators.

This project was funded by the DOE BER Sustainable Bioenergy Research Program, Award DESC0014395, and was also supported by DOE JGI Community Science Program. The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Gene Regulatory Networks Enabling Fungi to Selectively Extract Sugars from Lignocellulose

Jesus Castano[1*], **Jonathan S. Schilling[1]** (schillin@umn.edu), Claudia Schmidt-Dannert[1], Jiwei Zhang[1], Claire Anderson[1], David Hibbett[2], Igor Grigoriev[3], Young-Mo Kim[4]

[1]University of Minnesota, Saint Paul, [2]Clark University, [3]Joint Genome Institute (JGI) [4]Pacific Northwest National Laboratory (PNNL)

URL: http://schillinglab.cfans.umn.edu

Project Goals: Fungi dominate the biological decomposition of wood and other lignocellulosic plant tissues in nature. These saprotrophs offer us a proven model for making energy, sustainably, from biomass. They also offer those with commercial interests a range of pathways for unlocking sugars embedded in lignin. Their strategies range from 'white rot' mechanisms that remove lignin to gain access to polysaccharides to 'brown rot' mechanisms that selectively extract sugars, leaving most lignin behind. This metabolic diversity could be harnessed, industrially, but research has generally been focused more toward white rot delignification pathways. White rot fungi can unsheath polysaccharides by selectively removing lignin, a capacity that historically attracted interest for the potential to extract intact fibers for papermaking. Modern bioenergy schemes, however, do not aspire for intact fibers - instead, the goal is to depolymerize polysaccharides to release fermentable sugars (saccharification), saving lignin as a co-product, if possible. This is a better fit for the carbohydrate-selective pathways of brown rot fungi, but our grasp of fungal brown rot metabolism lags behind what we know about white rot.

Our collaborative project is aligned to address these gaps, with the **goal** of producing an integrated regulatory model for brown rot. Our proposed objectives insure stand-alone advances, but will also synergize to push ideas forward in a systems context.

Objective 1 is to identify fungal gene regulation patterns that distinguish brown rot fungi from fungi with other decay modes (e.g., white rot). We plan to compare fungi among relevant lineages but with varied carbohydrate-selectivities. We will culture these strains on solid wood wafers, spatially mapping gene expression and then overlaying fungal/wood metabolite patterns to enable temporally-resolved functional genomics. These maps can isolate patterns unique to brown rot and can target characterization.

Objective 2 focuses on characterization, starting with a short list generated in an earlier transcriptomics study, and progressively adding objective 1 gene targets. We plan to use routine single-/multi-cellular *in vitro* transformation pipelines, but will complement this with efforts to develop a brown rot transformation system, enabling *in vivo* manipulations (e.g., Crispr-Cas9).

Objective 3 is to use metabolomics to map metabolite-expression feedback over time, providing networks of gene regulation. This approach promises to advance our understanding of this unique brown rot strategy, beyond current ROS-centric models toward a systems view.

Abstract:

Certain filamentous fungi are uniquely able to deconstruct lignocellulose, and their poorly understood mechanisms have potential biofuels applications. A key hindrance to harnessing these fungal mechanisms has been their spatial complexity. Our past work has shown that differentiated networks of hyphae that penetrate wood are not metabolically uniform, with critical reactions occurring near the hyphal front. Standard omics analyses of these fungi from artificial media or from colonized wood ground en masse fail to distinguish expression of key gene products occurring in localized regions along growing hyphae.

Our focus for this research is specifically on brown rot fungi, a more recently evolved decay fungal group (relative to white rot) that circumvents the lignin barrier to extract sugars from lignocellulose. The genetic basis for how this capacity evolved away from white rot multiple times remains unknown, despite the modern options to align the compare brown rot and white rot fungal genomes. Our new collaboration aims to focus omics techniques to map and integrate expression over networks of wood-degrading fungal hyphae *in planta*. Wood-degrading fungal genomes are an emerging resource, particularly brown rot functional types. We recently optimized a thin-section wood set-up that can finely resolve reaction zones along an advancing mycelium. Within these zones, we can employ deep omics approaches without the typical noise of whole-sample homogenization. By co-localizing gene expression, secretions, and wood modifications, we can prioritize the genes most useful for application, as well as understanding the underlying regulation of a globally important decomposition mechanism – brown rot.

Our goal is to discover which genes are differentially up-regulated across the mycelia of wood-degrading fungi *in planta*, particularly at the leading edge of wood decomposition where reactive oxygen species (ROS) are deployed. To do this, we need to compare global expression profiles among mycelial regions. To map this wood-fungal interaction, we must match gene expression patterns with the extracellular secretome and with physiochemical wood modifications. Given this potential for substrate-fungus feedback, we will cross-check genes using separate clade representatives for brown rot fungi alongside their white rot ancestors, harnessing the JGI MycoCosm portal and several key resources and expertise at JGI and the Pacific Northwest National Laboratory.

References:

Castaño, J.D., Zhang, J., Anderson, C.E., and **Schilling, J.S.*** (2018) Oxidative damage control as a brown rot fungus attacks wood using oxygen radicals. Applied and Environmental Microbiology 84(22):e01937-18

Castaño, J., Zhang, J., Schilling, J.S.* (2018) Evaluation of colorimetric assays for determination of H_2O_2 in planta during fungal decomposition. Journal of Microbiological Methods 145:10-13.

Funding statement

This work was supported by the U.S. Department of Energy Office of Science Grant DE-SC0019357 from the Office of Biological and Ecological Research (BER) to J.S.S., by user Facility grant #48607 at the Environmental Molecular Sciences Laboratory of Pacific Northwest National Laboratory (to J.S.S.), and by Fulbright Colombia (to J.C.).

The Metabolic Basis of Low Biosynthetic Efficiency in an Oilseed

Lisa Carey[1,] Lisa Carey*[1] (careyl@msu.edu), Teresa J. Clark[1], Rahul R. Deshpande[1], Jean-Christophe Cocuron[1], Emily Rustad[1], Yair Shachar-Hill[1], Danny Schnell[1]

[1]Michigan State University East Lansing, MI

Project goals:

The practical of this project is to increase the seed yield of *Camelina sativa* by engineering increases in the metabolic efficiency with which its seeds convert sugars and amino acids supplied by the plant into storage products. To accomplish this, we have used Metabolic Flux Analysis to understand the basis of metabolic efficiency in this oilseed crop.

Abstract:

Camelina sativa, a member of the Brassicaceae, has tremendous potential for oil production. It is resilient to limited water and fertilizer supplies, has a short maturation time, and is not widely used for food production compared to other oilseeds (eg. Soybean, Canola). However, its seed yields are currently lower than those needed for agricultural success as a biofuel/chemical feedstock crop in the US. Developing Camelina embryos have a carbon conversion efficiency (CCE) lower than sunflower and less than half that of Canola embryos. This is unusual because in green seed plants, such as camelina, light improves CCE and generates ATP and reductant for the developing embryo. CCE derives from central metabolic fluxes, for whose quantification steady state metabolic flux analysis (MFA) has proven effective in other systems. We applied MFA to quantitatively map internal fluxes in developing Camelina embryos under high, normal and no light, which demonstrated that high fluxes through the Oxidative Pentose Phosphate are responsible for Camelina seeds' low CCE.

This project was funded by DOE

Transcriptional Regulatory Network Analysis to Increase Yields in Camelina

Fabio Gomez[1], Erich Grotewold*[1] (grotewol@msu.edu) Steve Lundback[1], Eric Maina[1], Danny Schnell[1]

[1]Michigan State University East Lansing, MI

Project goal:

Although metabolic regulation plays a key role in controlling carbon fixation, transport and allocation in plants, genetic control at the systems-level is superimposed on metabolism and thereby constitutes a key element controlling vegetative growth and seed development. To address the genetic controls that regulate seed and oil production, we will investigate and the transcriptional regulatory networks that impact seed oil production.

Abstract:

Understanding control of gene expression in oil producing crops is vital towards engineering a more economically feasible crop for biofuel production. To address the genetic controls that regulate seed and oil production, we are investigating and engineering the transcriptional regulatory networks that impact seed oil production. Much of what is known regarding the regulation of seed oil genes is derived from studies in Arabidopsis, a close relative of Camelina. Using this information, we have developed a gene regulatory knowledge base for Camelina using the published and publicly available Camelina RNA-seq datasets that will contain various analytical and computational tools with a friendly user interface. The database also includes transcription factors and co-regulators classified into different families as well as lipid gene annotations and homologs in Camelina. This will serve as a useful resource for the research community for accelerated discovery in elucidating regulatory mechanisms essential to their needs.

This work was funded by DOE BER

Enhancing chloroplast-mitochondrial metabolic networks to improve photosynthesis and crop yields

Bibin Paulose[1], Emily Pawlowski[1], Thomas D. Sharkey[1], and **Danny J. Schnell[1*]** (schnelld@msu.edu)

[1]Michigan State University, East Lansing, MI, USA

https://groecamelina.natsci.msu.edu

Project Goals:

Increased photosynthetic efficiency has the potential to address the predicted shortfalls in agricultural production by the middle of this century. Photosynthesis is highly sensitive to varying environmental conditions, including temperature, water and nutrient availability. Metabolic interactions between chloroplasts and mitochondria impact photosynthesis by redistributing and decreasing excess reducing equivalents and optimizing CO_2 fixation via photorespiratory metabolism. We studied components of the carbon concentrating mechanisms (CCMs) of aquatic algae to gain insight into the chloroplast-mitochondrial metabolic networks that function to optimize photosynthesis under limiting environmental conditions. We describe improved carbon assimilation, increased water and nitrogen use efficiency, and higher seed yields in the oil seed crop, *Camelina sativa*, expressing LIP36, a mitochondrial CCM component of *Chlamydomonas reinhardtti*. Our results suggest that LIP36 participates in redistributing metabolic intermediates to balance photosynthetic and photorespiratory metabolism, thereby maintaining plant productivity under non-ideal growth conditions.

Abstract:

The projected increases in global population growth, and the desire to transition to a sustainable, bio-based economy are resulting in ever increasing demands on agricultural productivity. One major limitation to increasing crop yields is the inefficiency of photochemical conversion of light to fixed carbon during photosynthesis. For example, the oxygenation of ribulose-1,5-bisphosphate (RuBP) catalyzed by Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) competes with CO_2 fixation and leads to a significant energy cost, generation of harmful reactive oxygen species, and the net loss of fixed carbon to the plant (Walker et al., 2016). Furthermore, the toxic products of oxygenation must be recycled via photorespiration, accounting for a >40% decrease in the efficiency of carbon fixation in C_3 photosynthesis, the pathway of carbon fixation used by the majority of crops and about 85% of all terrestrial plant species (Walker et al., 2016). Thus, strategies to balance the flux through photosynthetic and photorespiratory metabolism, while maintaining resilience to changing environmental conditions have the potential to broadly impact crop yields.

Chloroplasts and mitochondria serve complementary functions in plant growth through their primary roles in photosynthesis and respiration, respectively. In addition to their participation in energy generation, photosynthesis and respiration act in concert as components of an integrated network of metabolism to maintain cellular redox balance, optimize photosynthetic carbon capture, coordinate carbon and nitrogen metabolism, and mediate specific aspects of cellular stress signaling (Igamberdiev et al., 2018). As a consequence, knowledge of the interdependence and

coordinate control of chloroplast and mitochondrial metabolism is required to understand how plants optimize growth under both normal and stress conditions. In turn, these studies will facilitate strategies aimed at optimizing plant resilience and agricultural yields in the face of changing environmental conditions.

The primary example of chloroplast-mitochondrial metabolic integration is photorespiration. The conversion of glycine to serine in mitochondria by the combined action of glycine decarboxylase and serine hydroxymethyltransferase is a required step in recycling the toxic product of the oxygenation reaction catalyzed by ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco)(Bauwe et al., 2010) via photorespiratory metabolism. Photorespiration is estimated to account for upwards of 40% of flux through central carbon metabolism in C_3 photosynthesis, the pathway used by the majority of crops and about 85% of all terrestrial plant species (Ehleringer et al., 1991;Walker et al., 2016;Bauwe et al., 2010). Although photorespiration is often viewed as a non-productive process, photorespiratory metabolism is essential in plants, and higher photosynthetic rates and biomass production are observed by increasing the levels of glycine decarboxylase activity, demonstrating the positive link between mitochondrial photorespiratory metabolism, photosynthesis, and growth (Timm et al., 2015;Timm et al., 2012).

Photosynthesis and CO_2 assimilation also are affected by mutations that alter the activity of select enzymes in the tricarboxylic acid (TCA) cycle or that impact complex I of the mitochondrial electron transport chain (Carrari et al., 2003;Araújo et al., 2011;Nunes-Nesi et al., 2007;Nunes-Nesi et al., 2005;Bartoli et al., 2005;Dutilleul et al., 2003;Gandin et al., 2012;Chai et al., 2010), implicating specific aspects of respiratory metabolism in maintaining photosynthesis. Conversely, alterations in photosynthetic cyclic electron flow (Dang et al., 2014;Larosa et al., 2018) or NADH utilization in chloroplasts (Wu et al., 2015) result in significant changes to mitochondrial metabolic activity. The physiological role of these links is not fully understood, but has been attributed to an essential contribution of mitochondria to controlling the levels and distribution of reducing equivalents across cellular compartments and providing TCA cycle intermediates as carbon skeletons for anabolic metabolism and nutrient assimilation (Araújo et al., 2011;Nunes-Nesi et al., 2008).

Algal carbon concentrating mechanisms (CCMs)(Giordano et al., 2005) provide an excellent system to investigate cellular responses to the limitations imposed on photosynthesis and associated metabolism by environmental conditions. Many algae are regularly subjected to CO_2 limitation because of their aquatic habitats, and in response, have evolved low-CO_2 inducible CCMs to concentrate CO_2 at the Rubisco active site, and reconfigure metabolic networks to maintain CO_2 fixation and growth (Meyer et al., 2013). In *Chlamydomonas reinhardtii* nearly one-third of all nonoverlapping genes in the genome undergo significant changes in expression when the CCMs are induced under low-CO_2 conditions, and more than a thousand genes are significantly upregulated (Brueggeman et al., 2012;Meyer et al., 2013;Zhu et al., 2010). Two of the most highly induced genes of the chlamydomonas CCM encode low-CO_2-inducible proteins with a molecular mass of 36 kD (LIP36)(Moroney et al., 1991;Ramazanov et al., 1993). The LIP36 proteins, LIP36G1 and LIP36G2, are 96% identical and appear to be functionally redundant. They are predicted to be members of the mitochondrial carrier protein family, and LIP36-GFP fusion constructs localize to mitochondria in chlamydomonas and transgenic tobacco (Atkinson et al., 2016). Suppression of LIP36 expression by RNA interference demonstrated that they are essential for the growth of chlamydomonas under low CO_2 conditions (Pollock et al., 2004), consistent with

an essential role for mitochondrial metabolism in optimizing photosynthesis under limiting environmental conditions.

To gain insight into the role of LIP36 proteins, we expressed LIP36G1 in the oilseed crop plant, *Camelina sativa* (camelina). We demonstrate that expression of LIP36G1 significantly increased CO_2 assimilation and seed yields in camelina relative to control plants under limiting environmental conditions. Metabolic analysis suggests that LIP36 has a role in modifying mitochondrial metabolism during photosynthesis to optimize the balance between photosynthesis and photorespiration, and to provide intermediates for anabolic metabolism. The LIP36 lines also exhibited increased water use efficiency due to decreased stomatal conductance, demonstrating that it is possible to break the trade-off between water use efficiency and yield. Our results provide new insight into the interdependence of organellar metabolic networks in photosynthetic metabolism and illustrate the potential positive effects of engineering components of the algal CCM into vascular plants on crop productivity.

References

Araújo, W.L., et al., Plant, Cell & Environment, 2011. **35**(1): p. 1-21.
Araújo, W.L., et al., The Plant Cell, 2011. **23**(2): p. 600.
Atkinson, N., et al., Plant Biotechnol J, 2016. **14**(5): p. 1302-15.
Bartoli, C.G., et al., Journal of Experimental Botany, 2005. **56**(415): p. 1269-1276.
Bauwe, H., M. Hagemann, and A.R. Fernie, Trends in Plant Science, 2010. **15**(6): p. 330-336.
Brueggeman, A.J., et al., Plant Cell, 2012. **24**(5): p. 1860-75.
Carrari, F., et al., Plant physiology, 2003. **133**(3): p. 1322-1335.
Chai, T.-T., et al., Plant physiology, 2010. **152**(3): p. 1638-1649.
Dang, K.-V., et al., The Plant Cell, 2014. **26**(7): p. 3036.
Dutilleul, C., et al., Plant physiology, 2003. **131**(1): p. 264-275.
Ehleringer, J.R., et al., Trends Ecol Evol, 1991. **6**(3): p. 95-9.
Gandin, A., et al., Plant and Cell Physiology, 2012. **53**(9): p. 1627-1637.
Giordano, M., J. Beardall, and J.A. Raven, Annu Rev Plant Biol, 2005. **56**: p. 99-131.
Igamberdiev, A.U. and N.V. Bykova, Free Radic Biol Med, 2018. **122**: p. 74-85.
Larosa, V., et al., Plant Physiology, 2018. **176**(3): p. 2305.
Meyer, M. and H. Griffiths, J Exp Bot, 2013. **64**(3): p. 769-86.
Moroney, J.V. and C.B. Mason, Can J Bot, 1991. **69**(5): p. 1017-1024.
Nunes-Nesi, A., et al., The Plant Journal, 2007. **50**(6): p. 1093-1106.
Nunes-Nesi, A., et al., Plant physiology, 2005. **137**(2): p. 611-622.
Nunes-Nesi, A., et al.,. Journal of Experimental Botany, 2008. **59**(7): p. 1675-1684.
Pollock, S.V., et al., Plant Mol Biol, 2004. **56**(1): p. 125-32.
Ramazanov, Z., et al., Plant Physiol, 1993. **101**(4): p. 1195-1199.
Timm, S., et al., FEBS Letters, 2012. **586**(20): p. 3692-3697.
Timm, S., et al.,. The Plant Cell, 2015. **27**(7): p. 1968.
Walker, B.J., et al., Annu Rev Plant Biol, 2016. **67**: p. 107-29.
Wu, J., et al., Cell Research, 2015. **25**: p. 621.
Zhu, X.G., S.P. Long, and D.R. Ort, Annu Rev Plant Biol, 2010. **61**: p. 235-61.

This work is funded by grant #DE-SC0018269 from the Department of Energy BER program.

Strategies to Increase Photosynthetic Efficiency, CO_2 Fixation and Resource Allocation in Crops.

Heike Sederoff[1], Nathan Wilson[1], Brianne Edwards[1] and Danny Schnell[2]

[1]North Carolina State University, Dept. Plant and Microbial Biology
[2]Michigan State University, Department of Plant Biology

https://groecamelina.natsci.msu.edu

Project Goals:

One of the major limitations in photosynthetic carbon fixation occurs when O_2 competes with CO_2 at the active site of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco), resulting in the oxygenation of ribulose-1,5-bisphosphate (RuBP). This reduces CO_2 fixation and leads to the non-productive photorespiratory pathway at a significant energy cost and net loss of fixed carbon. Photorespiration accounts for a 30-40% decrease in the efficiency of carbon fixation in C3 photosynthesis, the pathway of carbon fixation used by the majority of crops, including Camelina, and about 85% of all terrestrial plant species. Our team is developing several distinct, but complementary approaches to address the limitations of CO_2 fixation in source tissues and increase seed yields in the model oilseed crop, Camelina sativa.

Abstract:

To increase the efficiency of biofuel crops to produce oil from photosynthetic CO_2 fixation, we designed and implemented several metabolic pathway modifications into our model oil seed crop Camelina sativa. One of the major limitations of photosynthetic CO_2 fixation is due to low activity and specificity of the CO_2-fixing enzyme Ribulose bisphosphate Carboxylase/Oxygenase (RuBisCO) and the CO_2 conductivity of the membrane systems surrounding RuBisCO. Other bottlenecks are due to the export control of assimilated carbon in form of sucrose from the photosynthetic leaf to the seeds and other sink tissues and the allocation of sucrose in seeds to increase the oil content. We used genetic engineering to a) increase the CO_2 conductivity of the leaf tissue by expression of a CO_2 transporter; b) expression of a synthetic CO_2 fixation cycle based on bacterial enzymes; c) expression of a photorespiratory bypass to reduce CO_2 loss; and d) reduction of a cell wall invertase inhibitor to reduce the limitation of sucrose phloem loading and unloading. While each modification (a-d) was successful in generating overall yield increases, integration of those traits show yield increases that are more dependent on environmental conditions like photoperiod and light intensity. This research will enable us to understand mechanisms underlying the integration of genotype and environment to optimize the use of genetic engineered traits in different bioenergy crops and growth regions.

This work was funded by ARPAe, BETO and BER #DE-SC0018269.

Advancing Field Pennycress as a New Oilseed Biofuels Feedstock that does not Require New Land Commitments

John Sedbrook[1]* (jcsedbr@ilstu.edu), Winthrop Phippen,[2] John Ralph,[3] and David Marks[4]

[1]Illinois State University, Normal, Il; [2]Western Illinois University, Macomb, Il; [3]University of Wisconsin-Madison, WI, [4]University of Minnesota, Minneapolis, MN

http://www.wiu.edu/pennycress/ http://cbs.umn.edu/marks-lab/home
https://about.illinoisstate.edu/jcsedbr/Pages/Research.aspx

Project Goals: This project aims to genetically improve the agronomic traits of Field Pennycress (*Thlaspi arvense* L.; pennycress) for its use as a profitable oilseed winter cover crop grown throughout the U.S. Midwest. We have identified large numbers of EMS-induced pennycress mutant lines exhibiting a variety of improved agronomic traits. We are also successfully and routinely employing CRISPR genome editing to generate pennycress mutants to test hypotheses about mutational effects on agronomic traits and to identify alleles to be used commercially. The traits on which we are focused for this funding are 1) Harvestable seed yields of at least 1,500 lbs/acre; 2) Reduced sinigrin (glucosinolate) to below the regulatory limit; 3) Reduced seed coat fiber to improve the seed meal nutritional value 4) Shortened time to maturity to consistently allow pennycress harvest in time to plant full-season soybeans.

Pennycress (*Thlapsi arvense* L.) is an emerging oilseed crop closely related to rapeseed canola and Arabidopsis that holds considerable agronomic and economic potential in producing seed oil to be used as a liquid biofuels feedstock. Pennycress possesses a unique combination of attributes including extreme cold tolerance, rapid growth, over-wintering growth habit, and a natural ability to produce economically-relevant amounts of seeds high in oil and protein. Pennycress could generate billions of liters of oil annually throughout the U.S. Midwest without displacing food crops or requiring land use changes. For example, pennycress can be grown on the ~35 million acres of U.S. Midwest farmland rotating each year between corn and soybeans. Much of this land otherwise lays fallow, resulting in nutrients loss into streams and soil erosion – two urgent problems which pennycress can help mitigate as a winter cover crop. This poster will highlight our efforts to 1) reduce seed glucosinolate content to make the oil and meal edible/palatable (we are characterizing multiple mutations in 15 glucosinolate biosynthetic and related regulatory genes), 2) reduce seed coat fiber content to allocate more metabolite to oil and to improve the nutritional value of the meal for use in animal feeds (mutations in 11genes controlling seed coat condensed tannins and acid detergent fiber (ADF) production have been identified and are being characterized), and 3) develop earlier maturing lines to avoid delays in planting soybeans, thereby enhancing farmer adoption (19 early flowering mutants have been confirmed for inheritance of earliness traits; we are focusing our efforts on three lines that produce seedpods 7-10 days earlier than wild type).

This material is based upon work that is supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, under award numbers 2014-67009-22305 and 2018-67009-27374.

Model exploration of carbon feedbacks between microalgae and heterotrophic bacteria

Yiwei Cheng[1] (yiweicheng@lbl.gov), Ty J. Samo[2], Xavier Mayali[2], Nicholas J. Bouskill[1] and Eoin L. Brodie[1]

[1]Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA; [2]Physical and Life Sciences, Lawrence Livermore National Laboratory, Livermore, CA

https://bio-sfa.llnl.gov/

Project Goals: The LLNL Bioenergy SFA seeks to support sustainable and predictable bioenergy crop production through a community systems biology understanding of microbial consortia that are closely associated with bioenergy-relevant crops. We focus on host-microbial interactions in algal ponds and perennial grasses, with the goal of understanding and predicting the system-scale consequences of these interactions for biomass productivity and robustness, the balance of resources, and the functionality of surrounding microbial communities. Our approach integrates 'omics measurements with quantitative isotope tracing, characterization of metabolites and biophysical factors, genome-enabled metabolic modeling, and trait-based representations of complex multi-trophic biological communities, to characterize the microscale impacts of single cells on system scale processes.

Eukaryotic microalgae and prokaryotic cyanobacteria (often collectively described as algae) have been proposed as a promising commercially viable feedstock for biofuels and other bioproducts. Open pond systems offer the most feasible and cost-effective way of scaling up biofuel production. However, productivity can be limited by fluctuations in environmental conditions, as well as competitive and trophic (e.g. prey-predator) interactions. While some interactions can negatively impact algal productivity, certain interactions can enhance productivity or mediate factors limiting productivity such as nutrient limitation. A major interaction between algae and bacteria is widely observed, where algae fix inorganic carbon (C) and release dissolved organic C for bacterial consumption. Under certain conditions, an exchange occurs, whereby bacteria release resources that may enhance microalgal growth including dissolved inorganic C (DIC) or nitrogen (N), phosphorus (P), iron, and vitamins. Similarly, bacteria may enhance algal growth through the reduction of inhibitory by-products such as oxygen.

In this work, *Phaeodactylum tricornutum*, a biofuel relevant microalgal species, was grown in laboratory co-cultures with and without a heterotrophic bacterium under cyclical light conditions. *P. tricornutum* grown in co-culture showed higher biomass yield than the axenic culture by 10%. Coupled with earlier experimental findings (Mouget et al., 1995; Samo et al., 2018), we proposed a positive feedback loop between algae and associated heterotrophic bacteria that enhances productivity of the system. Specifically, we hypothesize that algal exudation of fixed dissolved organic C (DOC) fuels growth of heterotrophic bacteria, leading to DOC re-mineralization and increased DIC availability for the algae, relieving limitation of carbon fixation. The strength of this interaction is dependent on environmental and biological factors such as exchange of CO_2 across the air/water interface, combined with algal and bacterial traits. Studies to date have yet to focus on the synergistic effects of these factors on productivity.

To test the carbon (DOC-DIC) feedback hypothesis, we developed a dynamic energy budget model of the algae and interacting heterotrophic bacteria. Model simulations were benchmarked against experimental biomass carbon data. A trait based modeling approach was then utilized to explore productivity under scenarios representing interacting gas exchange efficiencies, bacterial-algal respiration traits, and N or P uptake rates. Simulated exchange of CO_2 and O_2 across the air/water interface demonstrates how mass transfer efficiency impacts dissolved CO_2 (DIC) and O_2 (DO) concentrations. In well-mixed systems with rapid atmospheric exchange, DIC and DO concentrations are in equilibrium with atmospheric CO_2 and O_2. Under such a scenario, with no DIC and DOC accumulation, no enhanced productivity is observed due to algal-bacterial interactions. Inefficient atmospheric exchange however leads to accumulation of DIC and DOC, allowing DIC and DOC exchange between the algae and the bacteria, enhancing productivity. Model representation of the effect of O_2 affinity for RubisCO shows that DIC:DO ratios control the kinetics of RuBisCO and therefore, bacterial production of DIC or reduction of O_2 should enhance the C fixation rate, strengthening the algal-bacterial carbon feedback loop. This interaction is optimal when heterotrophic bacteria have traits that confer high respiration rates but with lower uptake or requirement for N and P than algae. Overall, our results illustrate that algal-bacterial interactions can enhance algal productivity and that that gas exchange efficiencies interact with algal and bacterial traits that can be selected to strengthen the carbon feedback loop.

This work was performed under the auspices of the U.S. Department of Energy at Lawrence Livermore National Laboratory under Contract DE-AC52- 07NA27344 and supported by the Genome Sciences Program of the Office of Biological and Environmental Research under the LLNL Biofuels SFA, FWP SCW1039. Work at LBNL was performed under the auspices of U.S. Department of Energy Contract No. DE-AC02-05CH11231.

References

Mouget, J., Dakhama, A., Lavoie, M. C., de la Noue, J. (1995). Algal growth enhancement by bacteria: Is consumption of photosynthesis oxygen involved? *FEMS Microbiology Ecology*, 18, 35-44.

Samo, T. J., Kimbrel, J. A., Nilson, D. J., Pett-Ridge, J., Weber, P. K., Mayali, X. (2018). Attachment between heterotrophic bacteria and microalgae influences symbiotic microscale interactions. *Environmental Microbiology*, 20(12), 4385-4400.

Improving Draft Metabolic Models in KBase: Tools for Importing, Comparing and Merging Metabolic Annotations

Patrik D'haeseleer[1] (dhaeseleer2@llnl.gov), Jeffrey A. Kimbrel[1], Janaka N. Edirisinghe[2], Felipe Liu[2], James Jeffryes[2], Ali Navid[1], Chris Henry[2], and Rhona K. Stuart[1]

[1]Lawrence Livermore National Laboratory, Livermore CA, USA
[2]Argonne National Laboratory, Lemont IL, USA

https://bio-sfa.llnl.gov/

Project Goals: The LLNL Bioenergy SFA seeks to support sustainable and predictable bioenergy crop production through a community systems biology understanding of microbial consortia that are closely associated with bioenergy-relevant crops. We focus on host-microbial interactions in algal ponds and perennial grasses, with the goal of understanding and predicting the system-scale consequences of these interactions for biomass productivity and robustness, the balance of resources, and the functionality of surrounding microbial communities. Our approach integrates 'omics measurements with quantitative isotope tracing, characterization of metabolites and biophysical factors, genome-enabled metabolic modeling, and trait-based representations of complex multi-trophic biological communities, to characterize the microscale impacts of single cells on system scale processes.

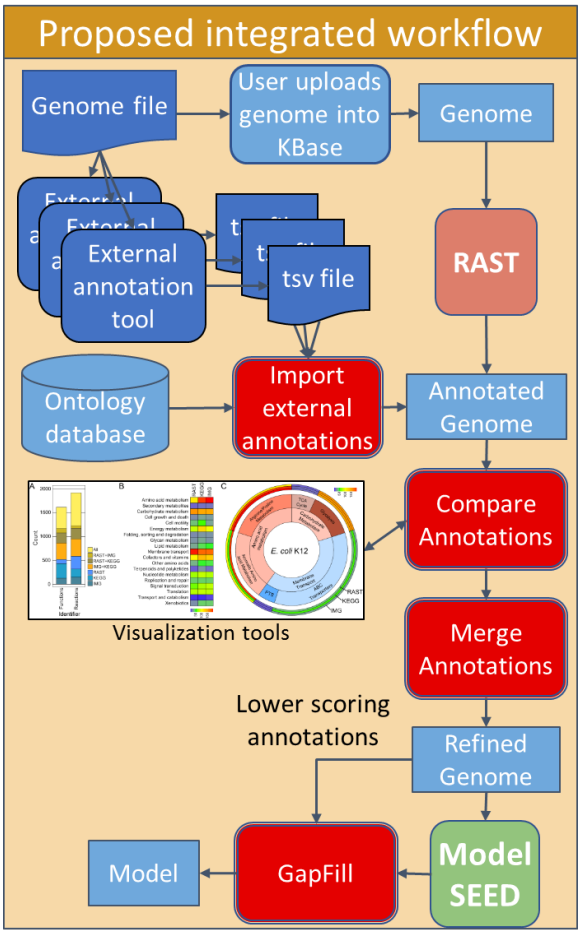
Metabolic pathway analysis, and especially metabolic modeling, is one of the cornerstones of modern Systems Biology, because it allows us to go straight from sequence data to gene functions to an understanding of how whole biological systems function. Metabolic modeling is a critical part of our LLNL Biofuels SFA, which investigates metabolic interactions in bioenergy-relevant microbial communities. Accurate metabolic models require well-annotated genomes. Unfortunately, assigning functional annotations to genes is an imperfect science, and 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 previous research, we have shown that single metabolic annotation tools such as RAST or KEGG tend to be incomplete and inconsistent, and that merging annotation from multiple sources can drastically increase the number of genes and metabolic reactions included in metabolic models (Griesemer *et al*, BMC Genomics 2018). Our comprehensive approach added on average 40% more reactions, 3-8 times more substrate-specific transporters, and 37% more metabolic genes, compared to annotation using only a single tool. These results are even more pronounced for pathways outside of the core carbon metabolism, and for bacterial species that are phylogenetically distant from well-studied model organisms.

The DOE Systems Biology Knowledgebase (KBase) contains a suite of powerful Apps for building genome-scale metabolic models. ModelSEED, originally developed by our collaborator

Chris Henry, is the central Flux Balance Analysis model building App in KBase, and one of the most popular metabolic modeling tools for generating draft models because of its accessibility, ease of use and quality. However, it currently only supports metabolic annotations produced by the Annotate Microbial Genome App, based on RAST (Rapid Annotations using Subsystems Technology). This means that so far it has been impossible for researchers who may prefer to use other high quality annotation tools such as KEGG or even JGI's IMG platform to import their annotations into KBase, let alone merge annotations from multiple sources.

We are currently developing a set of KBase Apps to allow users to upload functional annotations from popular third-party annotation tools, compare and merge them, and use them for metabolic modeling. (1) An Import App (close to completion) will allow user to upload a simple tab-separated file with annotation data in the form of EC numbers, KEGG and MetaCyc reactions identifiers. (2) A Compare App will allow the user to compare metabolic annotations from different sources, by mapping all of them to the ModelSEED reaction database. (3) The Merge App will provide the user with a simple yet flexible scoring mechanism to select a preferred set of annotations from among the full set of functional identifiers mapped to each gene in the genome. (4) Finally, we will also assist the KBase metabolic modeling team to make modifications in the existing ModelSEED App and the Gap Filling tool, to enable users to build models from the merged highest-confidence annotations, and prioritize the remaining lower-scoring annotations for gapfilling.



This work was performed under the auspices of the U.S. Department of Energy at Lawrence Livermore National Laboratory under Contract DE-AC52- 07NA27344 and supported by the Genome Sciences Program of the Office of Biological and Environmental Research under the LLNL Biofuels SFA, FWP SCW1039.

Changes to the elemental composition of the green alga *Chlamydomonas reinhardtii* based on variations in cultivation conditions

Colleen Hui[1,2*] (hui@chem.ucla.edu), Stefan S. Schmollinger[1], Sean D. Gallaher[1], **Peter K. Weber**[2], **Sabeeha S. Merchant**[1]

[1]University of California, Los Angeles, Los Angeles, CA; [2]Lawrence Livermore National Laboratory, Livermore, CA

https://bio-sfa.llnl.gov/

Project Goals: The LLNL Biofuels SFA is developing advanced methods to support sustainable microalgae fuel production. In this part of the project we are working on characterizing metal metabolism in microalgae through imaging. Metals are cofactors for essential enzymes and the ability to localize and quantify them will lead to a better understanding of how microalgae metal metabolism works. Because metals occur at parts per million and lower levels in some cells, they can be difficult to image by standard methods. Here we use high spatial resolution secondary ion mass spectrometry with a Cameca NanoSIMS 50 to image metals in embedded and sectioned cells. The NanoSIMS data are correlated with other methods to provide a more complete understanding of metal metabolism in these microalgae.

Trace metals are critical for many essential reactions in all living organisms. Among the most important metals are copper (Cu), iron (Fe), and manganese (Mn), which all serve as cofactors to enable catalysis and redox chemistry. For decades, the unicellular green alga *Chlamydomonas reinhardtii* has been used as a reference organism for understanding eukaryotic metal homeostasis because it retains many genes that were present in the common ancestor to both plants and animals, and it is amenable to classical and molecular genetic manipulation. The elemental composition in Chlamydomonas cells can be altered by nutrient availabilities in the environment (e.g. trace metal abundances, carbon/nitrogen supply). In addition, growth parameters such as cell density, carbon/energy source, soil/water pH and aeration are known to influence the metabolism of the alga, which might affect the trace metal quota. To understand the potential impact of such growth parameters on metal homeostasis in Chlamydomonas, we have completed a systematic analysis of cellular Cu, Fe, and Mn content as a function of six different cultivation variables using ICP-MS/MS. These variables include 1) cell density/sampling time, 2) starting medium pH, 3) photon flux density, 4) vessel size, 5) culture volume, and 6) shaker

speed. Chlamydomonas cells accumulate Cu and Fe, but not Mn, as a function of time, especially during the stationary phase. Alkaline medium (pH 8.5) under photoheterotrophic growth condition induced hyperaccumulation of Fe in cells up to 10- to 20- fold higher than cells grown in the neutral pH medium. Other parameters like aeration and photon flux density have an impact on Cu, Fe, and Mn quota, but these effects are smaller compared to the influence of pH on Fe accumulation. Our results suggest that the regulation of individual trace metals (Cu, Fe, and Mn) in Chlamydomonas may each be specific in response to different environmental stimuli. Furthermore, this analysis demonstrates that cell density/timing of sampling and medium pH must be firmly controlled during experimentation in order to ensure a consistent elemental makeup of the cells, and thereby yielding reproducible and reliable outcomes. Meanwhile, because the accumulation of Fe in alkaline condition was so striking, we further investigated this aspect in more detail. To examine the effects of medium pH on cell physiology, we have undertaken a comparative transcriptome analysis of cells grown in media at pH 8.5 versus pH 7.0. We observed that 2523 genes (~14%) are differentially regulated between cultures grown at alkaline vs. neutral pH. We also visualized Fe distribution in cells sampled at alkaline and neutral pH by nanoSIMS. Fe was evenly distributed within the cell when grown in neutral pH media. In comparison, imaging of cell sections sampled at pH 8.5 showed distinctive hotspots of Fe colocalized with calcium and phosphorus. This suggests that the excess Fe was stored in acidocalcisomes, a lysosome-related organelle.

This research was supported by the LLNL Biofuels Scientific Focus Area and an award to the Merchant Lab, funded by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under FWP SCW1039 and DE-FG02-04ER15529. Work at LLNL was performed under the auspices of the U.S. Department of Energy at Lawrence Livermore National Laboratory under Contract DE-AC52- 07NA27344 and the Livermore Graduate Scholar Program.

Chemical and Proteomic Profiling of Organic Compounds for Detecting Algal Interactions with Grazers and Commensalistic Bacteria

Kristen L. Reese [1,2]; Carolyn L. Fisher [3]; Matthew W. Moorman [4]; A. Daniel Jones [2]; Matthias Frank [1]; Miriam Windler[5], Alfred Spormann[5], Mary Lipton[6], Lindsey Anderson[6], Sam Purvine[6], Carrie D. Nicora[6], Patrik D'haeseleer[1], Xavier Mayali[1], and **Rhona Stuart[1]**, **Todd W. Lane (twlane@sandia.gov)[3]**

[1]Lawrence Livermore National Laboratory, Livermore, CA; [2]Michigan State University, East Lansing, MI; [3]Sandia National Laboratories, Livermore, CA; [4]Sandia National Laboratories, Albuquerque, NM; [5]Stanford University, Stanford, CA; [6]Pacific Northwest National Laboratory, Richland, WA

Project Goals: The LLNL Bioenergy SFA seeks to support sustainable and predictable bioenergy crop production through a community systems biology understanding of microbial consortia that are closely associated with bioenergy-relevant crops. We focus on host-microbial interactions in algal ponds and perennial grasses, with the goal of understanding and predicting the system-scale consequences of these interactions for biomass productivity and robustness, the balance of resources, and the functionality of surrounding microbial communities. Our approach integrates 'omics measurements with quantitative isotope tracing, characterization of metabolites and biophysical factors, genome-enabled metabolic modeling, and trait-based representations of complex multi-trophic biological communities, to characterize the microscale impacts of single cells on system scale processes. https://bio-sfa.llnl.gov/

Biotic interactions in algal ponds are both affected by and influence the chemical environment surrounding algal cells. As part of our SFA, we aim to disentangle various components of these chemical signals in order to better understand the mechanisms of how biotic interactions influence algal productivity and physiology. Here, we present two studies focused on different aspects—the first highlights the role of trophic interactions, identifying volatiles produced in algal-rotifer interactions, while the second examines algal-bacterial interactions at the molecular level in a model algal species.

Chemical compounds, in particular those defined as volatile, are released by all organisms, and in addition to being byproducts of metabolism, can play roles as signaling molecules, both within and between species. We are interested in detecting, quantifying, and identifying VOCs capable of differentiating microalgal predators. The presented work uses solid-phase microextraction (SPME)-GC-MS to analyze volatiles emitted before and after predator infection in microalgal strains that are currently important to the biofuels community. *Microchloropsis salina* was grown in six- 15 L cultures of chemically-defined, enriched artificial seawater media with high purity gas. VOC profiles were sampled from all cultures after 24-48 hour propagations using 65 μm polydimethylsiloxane-divinylbenzene SPME fibers with 60 minute exposures, and 70 eV electron ionization on a quadrupole mass analyzer. After 48 hours of growth, the marine rotifer, *Brachionus plicatilis*, was added at high density to two of the cultures. VOC profiles and algal counts were monitored for an additional 2-4 timepoints after

inoculation. Including our controls we had four conditions to assist with identification of a rotifer infection: 1*) M. salina*-only cultures, 2) *B. plicatilis*-infected *M. salina* cultures, 3) enriched seawater controls, and 4) SPME fiber blanks. Volatiles were attributed to each condition if meeting the following conservative requirements: 1) presence in <66% of replicates at a timepoint, 2) greater than 10x abundance of corresponding abundance in media-only control. Subsequent characterization, alignment, and preliminary identifications were made through analysis of deconvoluted experimental spectra with comparisons to the NIST14 spectrum database (Match>70%) and retention index matching.

One main finding was that the addition of *B. plicatilis* to healthy cultures of *M. salina* produced an abundance of tentatively identified carotenoids, such as trans β-ionone and 4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-2-butanone, which may indicate algal lysis. Concurrent to VOC analysis, daily measurements of algal density confirmed that rotifer-inoculated cultures displayed decreased live algae counts relative to the uninoculated controls. The peak areas of the carotenoid biomarkers were observed to increase when rotifers had consumed greater amounts of algae. Application of multivariate statistical analyses to these datasets are in-progress. Ongoing experiments include work with the closely-related algal species *Nannochloropsis gaditana* and *Nannochloropsis oceanica*, additional rotifer and chytrid predators, and evaluating biomarker targets in complex, open-environment microalgal systems. Our work aims to increase the breadth and depth of reported algal and rotifer-specific VOCs, providing a tool to better define the chemical environment of microalgal ponds.

In parallel, we are also investigating a model green algae, *Chlamydomonas reinhardtii,* and its metabolic interactions with an actinobacterium, *Arthrobacter* sp. P2b. In order to gain a mechanistic understanding of the physiological effects on algal metabolism caused by chemicals released by heterotrophic bacteria, we are using a simple model system of one bacterial mutualist to ultimately understand the complex chemical and biotic interactions occurring in ponds at the molecular and gene level. We observed that both co-culturing with P2b, and P2b cell-free spent media enhanced algal chlorophyll content, biomass, and cell size, suggesting a beneficial or commensal interaction between the two species. In order to determine putative gene pathways involved in *C. reinhardtii* increased biomass, we compared global protein expression between co-cultures and monocultures of each species, as well as the algal monoculture incubated with the P2b cell-free spent media, at two time points. We find that several factors may be involved in the effect of P2b on *C. reinhardtii* physiology, influencing pathways related to cell cycle regulation.

This work was performed under the auspices of the U.S. Department of Energy at Lawrence Livermore National Laboratory under Contract DE-AC52- 07NA27344 and supported by the Genome Sciences Program of the Office of Biological and Environmental Research under the LLNL Biofuels SFA, FWP SCW1039. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This poster describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. A portion of the

research was performed at the Environmental Molecular Sciences Laboratory (EMSL), a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research.

Characterization of microalgal-bacterial symbioses at the microscale to examine association-dependent remineralization processes at the single-cell level

Ty Samo[1] (samo1@llnl.gov), Mary Lipton[2], Sam Purvine[2], Carrie D. Nicora, Alice Dohnalkova[2], Scott Lea[2], Patrik D'haeseleer[1], Peter K. Weber[1], Rhona K. Stuart[1], **Xavier Mayali[1]**

[1]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA; [2]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA

https://bio-sfa.llnl.gov/

Project Goals: The LLNL Bioenergy SFA seeks to support sustainable and predictable bioenergy crop production through a community systems biology understanding of microbial consortia that are closely associated with bioenergy-relevant crops. We focus on host-microbial interactions in algal ponds and perennial grasses, with the goal of understanding and predicting the system-scale consequences of these interactions for biomass productivity and robustness, the balance of resources, and the functionality of surrounding microbial communities. Our approach integrates 'omics measurements with quantitative isotope tracing, characterization of metabolites and biophysical factors, genome-enabled metabolic modeling, and trait-based representations of complex multi-trophic biological communities, to characterize the microscale impacts of single cells on system scale processes.

As part of our investigations into the effects of algal-associated bacteria on biomass production (Samo et al. 2018), we have continued studying the interactions between the model diatom *Phaeodactylum tricornutum* CCMP2561 and individual bacterial strains isolated from outdoor raceways, including strains that are numerically abundant and can be found as phycosphere-associated. Our overarching goals are to: 1) examine the role of physical interactions (i.e., attachment) in the remineralization of excreted dissolved organic matter as well as 2) leverage and develop a suite of new approaches to quantify and monitor the underpinnings of these associations.

To quantify the remineralization of *P. tricornutum* organic exudates by bacteria and subsequent transfer back to algal cells as inorganic C and N, we employed a high-resolution stable isotope imaging approach (nanoSIP; Pett-Ridge and Weber 2012). In a previous experiment performed with *P. tricornutum* and a *Marinobacter* strain, we found that the bacteria incorporated ^{15}N-labeled exudates and transferred significantly more ^{15}N-labeled compounds to the algal host compared to the algae grown axenically. Here, we collected isotope-labeled exudates released from axenic algae grown in ^{13}C-bicarbonate ($NaH^{13}CO_3$) and ^{15}N-nitrate ($Na^{15}NO_3$) using solid phase extraction, fed these exudates to individual co-cultures of *P. tricornutum* and 16 bacterial genera, and incubated for 24 h. We examine the rate of transfer of both C and N from bacteria to algae, providing relative benefit metrics of each bacterial genus to algal nutrient requirements. Additionally, in collaboration with PNNL, we are beginning to investigate how attachment of one of these bacterial isolates, *Marinobacter* sp. 3-2 (*Marinobacter*), to *P. tricornutum* cells

progresses over time and facilitates the exchange of compounds modulating these symbioses (Fig. 1).

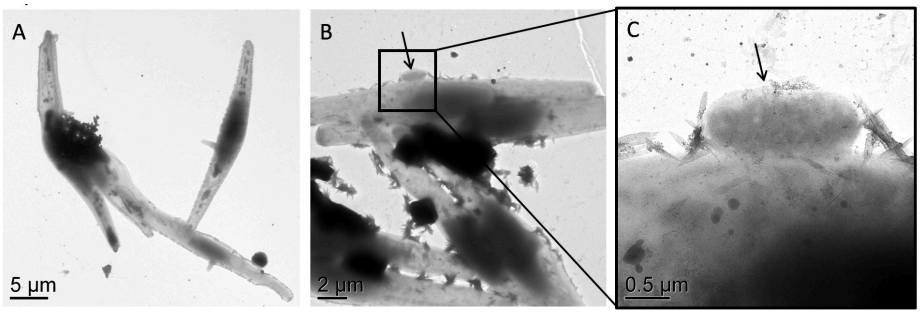


Figure 1. CryoTEM images depicting close associations between P. tricornutum *and* Marinobacter *sp. 3-2. A) View of four individual* P. tricornutum *cells. B) Enhanced view of a single* Marinobacter *cell attached to the surface of* P. tricornutum*, with C) depiciting a highly zoomed acquisition of the bacterium and its orientation on the algal cell.*

In order to better understand *P. tricornutum*-bacterial interactions at the molecular and genetic level, and identify putative genes involved in beneficial interactions, we also conducted global protein expression profiling. We compared *P. tricornutum* axenic cultures with two *P. tricornutum*-bacterial co-cultures, both during mid-exponential growth. The two bacterial strains have distinct interactions with *P. tricornutum*. The first includes the above *Marinobacter* isolate, which is capable of attachment, and enhances growth and single cell carbon fixation. The second, *Rhodobacteriaceae* sp. PT6CLA, does not attach to *P. tricornutum* nor provide any significant growth or carbon fixation enhancement. The two bacterial strains also exhibit distinct growth patterns in co-culture, with *Marinobacter* maintaining low, consistent abundance relative to *P. tricornutum*, and PT6CLA increasing exponentially with *P. tricornutum* growth. Proteomic profiling comparing these co-cultures with each other and the axenic culture revealed a number of putative pathways involved in algal-bacterial interactions, highlighting the multifaceted nature of the interaction.

Our findings have implications for how we understand nutrient cycling at the microscale and highlight the influence of pairwise interactions between heterotrophic bacterial symbionts and their photosynthetic algal hosts on bulk biogeochemical cycling. By combining multiple methods, we find that molecular, biogeochemical, and physiospatial factors are all critical to our understanding of algal-bacterial interactions at the microscale and inform efforts to extrapolate these observations to the community scale using numerical and conceptual models.

References

Pett-Ridge, J., and P. K. Weber. 2012. NanoSIP: NanoSIMS Applications for Microbial Biology, p. 375–408. *In* Microbial Systems Biology. Humana Press, Totowa, NJ.

Samo, T. J., J. A. Kimbrel, D. J. Nilson, J. Pett-Ridge, P. K. Weber, and X. Mayali. 2018. Attachment between heterotrophic bacteria and microalgae influences symbiotic microscale interactions. Environmental Microbiology **20**: 4385–4400. doi:10.1111/1462-2920.14357

This work was performed under the auspices of the U.S. Department of Energy at Lawrence Livermore National Laboratory under Contract DE-AC52- 07NA27344 and supported by the Genome Sciences Program of the Office of Biological and Environmental Research under the LLNL Biofuels SFA, FWP SCW1039. A portion of the research was performed at the Environmental Molecular Sciences Laboratory (EMSL), a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research.

Multitrophic considerations critical to understanding algal pond ecosystems

Christopher S. Ward[1] (ward56@llnl.gov), Jeffrey A. Kimbrel[1], Haifeng Geng[2], Anthony Siccardi[3], Laura T. Carney[4], Todd W. Lane[2], **Rhona K. Stuart[1]**, **Xavier Mayali[1]**

[1]Physical and Life Sciences, Lawrence Livermore National Laboratory, Livermore, CA; [2]Department of Systems Biology, Sandia National Laboratories, Livermore, CA; [3]Department of Biology, Georgia Southern University, Statesboro, GA; [4]Agricultural Science and Molecular Ecology, Heliae Development LLC, Gilbert AZ

Project Goals: The LLNL Bioenergy SFA seeks to support sustainable and predictable bioenergy crop production through a community systems biology understanding of microbial consortia that are closely associated with bioenergy-relevant crops. We focus on host-microbial interactions in algal ponds and perennial grasses, with the goal of understanding and predicting the system-scale consequences of these interactions for biomass productivity and robustness, the balance of resources, and the functionality of surrounding microbial communities. Our approach integrates 'omics measurements with quantitative isotope tracing, characterization of metabolites and biophysical factors, genome-enabled metabolic modeling, and trait-based representations of complex multi-trophic biological communities, to characterize the microscale impacts of single cells on system scale processes. https://bio-sfa.llnl.gov/

Considerable effort has been dedicated towards improving production of algal biomass for renewable bioenergy. There is a growing recognition that biotic factors – *i.e.*, interactions between algae and their microbiome, which includes microorganisms from all three kingdoms – can significantly affect algal productivity and offer tangible areas for improvement besides algal strain development and abiotic factors. In three highlighted studies of mass algal cultivation facilities described below, we explore several key biotic factors, namely functional attributes of bacteria associated with algal cultures and potential trophic interactions with competitors, predators and parasites. These findings highlight the importance of various aspects of algal pond ecology in algal culture outcomes and offer insights into unique ecology-driven cultivation improvements.

Algal mesocosms are dynamic systems in which the associated microbial communities change with concurrent changes in algal physiology and population density. To interrogate the functional capabilities of bacterial communities associated with algal cultures, six replicate outdoor mesocosms with *Microchloropsis salina* cultures were sampled for metagenomes at 3 weeks (Early) and 5 weeks (Late) after inoculation. Both time points included bacterial functions involved in lipid/nucleotide metabolism, as well as production of vitamins. Global functional gene differences indicated an increase in functional capability to degrade carbohydrates and to produce secondary metabolites, with a decrease in genes for signal transduction and quorum sensing. Binning the metagenomes resulted in 10 high-quality reconstructed genomes (MAGs)

with a diverse suite of functional potentials. Several MAG genomes code for functions that suggest they may acquire energy phototrophically (via bacteriochlorophyll and photosynthetic gene clusters). We continue to investigate how key microbiome functions, *i.e.* carbon capture/remineralization and secondary metabolite production, may aid algal photosynthesis and growth.

In order to further expand our knowledge on controls of pond dynamics, in a separate study, we examined the progression of both prokaryotic (16S rRNA) and eukaryotic (18S rRNA) community members in mesocosms seeded with either *M. salina* or *Phaeodactylum tricornutum* over a typical production timeline with multiple algal harvestings. Initially, the *P. tricornutum* ponds had variable taxa in each replicate pond but converged towards Alphaproteobacteria and eventually Flavobacteria over time. Conversely, the *M. salina* ponds were much less variable, having primarily Deltaproteobacteria throughout the experiment. Although each pond was initially seeded with only a single algal species, results show that several ponds experienced contamination from a neighboring algal species over the course of the experiment. In the *P. tricornutum* ponds, a pathogenic *Kordia* bacterial species bloom likely led to the premature crash in all replicate mesocosms. In the *M. salina* mesocosms, however, potentially pathogenic *Pseudobacteriovorax* species were abundant, yet no algal crash was observed. In all ponds, predatory ciliates were prevalent, with the most abundant families distinct to the host algal strain. These findings indicate that in algal mesocosms, multitrophic interactions are widespread and, though often overlooked, top-down controls may be key factors in pond dynamics.

In order to gain a better understanding of the role of a particularly prominent algal parasite in algal ponds, we conducted a time course sampling of algal microbiomes in *Haematococcus pluvialis* raceways frequently infected by chytrid fungi. The bacterial component of the pond microbiome displayed strong successional dynamics driven by changes both in culture age and algal physiology. Chytrid ecology was also associated with algal processes, with high pond infectivity only in cyst stages. Culture-based assays revealed a previously unrecognized non-infective chemoorganotrophic lifestyle. To extend these findings to the field, we developed a fungal group-specific qPCR method to assess abundances of chytrids and blastoclads. qPCR results confirmed that samples with active chytrid infections had high chytrid and/or blastoclad abundances, while also revealing that in some samples where no active chytrid infection was detected, abundances were still high – approximately 12-25% of those samples from infected ponds. This validates that the noninfective chytrid lifestyle is relevant to chytrid dynamics in algal ponds and suggests that abundant noninfective chytrids may make additional contributions to carbon and nutrient cycling processes in algal systems.

This work was performed under the auspices of the U.S. Department of Energy at Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 and supported by the Genome Sciences Program of the Office of Biological and Environmental Research under the LLNL Biofuels SFA, FWP SCW1039.

Novel Microbial Routes to Synthesize Industrially Significant Precursor Compounds

F. Robert Tabita[1*] (tabita.1@osu.edu), Justin A. North[1], Kelly C. Wrighton[2], Adrienne Narrowe[2], and William R. Cannon[3]

[1]The Ohio State University, Columbus, OH; [2]Colorado State University, Fort Collins, CO; [3]Pacific Northwest National Laboratory, Richland, WA

Project Goals

Ethylene is the most widely employed organic precursor compound in industry. Using recently discovered efficient microbial anaerobic ethylene synthetic processes, the potential to impact ethylene formation is tenable using plentiful lignocellulose and/or CO_2 feedstocks. The overall long-term objective is to develop an industrially compatible microbial process to synthesize ethylene in high yields. Based on the discovery of a novel and genetically regulated anaerobic pathway to produce high levels of ethylene (the DHAP ethylene pathway), the following specific aims will be addressed:

1. Fully probe the catalytic potential of all enzymes of the DHAP ethylene pathway and determine the regulatory mechanism of DHAP-ethylene pathway gene expression. Model the thermodynamics and kinetics of ethylene synthetic pathways.
2. Discover effective and active ethylene enzymes encoded in cultured and uncultured organisms from anoxic environments.
3. Construct a modular set of optimized genes (from Aims 1 and 2) on a DNA fragment containing specific regulatory elements that will allow high level gene expression in model organisms that have been flux optimized.

Abstract

While studying the role of RubisCO and the RubisCO-like Protein (RLP) in sulfur metabolism, we discovered novel pathways for the metabolism of the key compound 5-metthylthioadenosine (MTA). We discovered distinct anaerobic MTA pathways in *Rhodospirillum rubrum* and the related organism *Rhodopseudomonas palustris* (1). **During the course of elucidating these novel anaerobic MTA metabolic pathways, we found that one of these routes resulted in the production of copious quantities of ethylene (1) (Fig. 1A), the first reported anaerobic route to produce ethylene, involving novel genes and enzymes.** Ethylene synthesis was shown to be highly regulated, with significantly high concentration produced under optimum conditions (~ 100 µmol/L/culture OD) (1) using either CO_2 or organic carbon growth substrates. A search of publically available genomic databases for other organisms that possess the newly discovered genes encoding anaerobic ethylene synthesis enzymes indicated this pathway was widespread in genomes from multiple phyla of industrially relevant Proteobacteria, Firmicutes (including lignocellulose degrading strains, Fig.1B) and a few archaea (1).

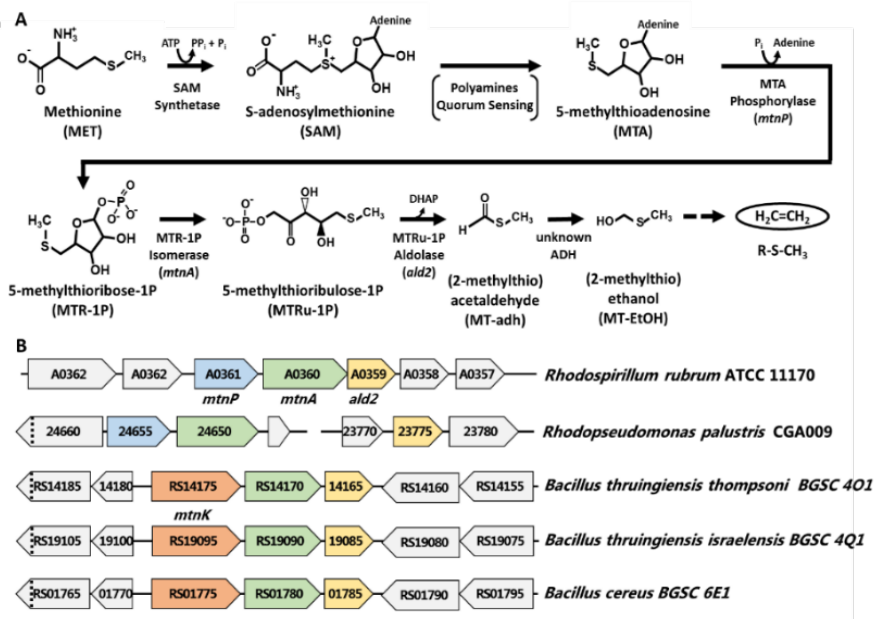


Fig. 1. Anaerobic ethylene pathway, enzymes, and genes. (A) Anaerobic ethylene-forming pathway from *R. rubrum* and *R. palustris* (from ref. 1). (B) Alignment of *R. rubrum* and *R. palustris* gene cluster Rru_A0359 – Rru_A0361 for bioethylene production with genes from representative lignocellulose degrading *Bacillus* spp. (Blue) *mtnP*; (Red) *mtnK*, 5-methylthioribose kinase, which in conjunction with *mtnN* (MTA nucleosidase) replaces *mtnP* in *Bacillus sp.*; (Green) *mtnA*; (Yellow) *ald2*.

Based on the discovery of a novel and genetically regulated anaerobic pathway to produce high levels of ethylene (1) (Fig. 1), the following specific aims will be addressed to enhance ethylene production:

<u>Specific Aim 1</u>: Analysis and Characterization of Ethylene Biosynthetic Pathways (Tabita, Cannon). In this aim, we will characterize the ethylene biosynthetic pathway with respect to its (A) biochemistry, (B) regulation and (C) thermodynamics & kinetics.

<u>Specific Aim 2</u>: Discover effective and active ethylene enzymes encoded in cultured and uncultured organisms from anoxic environments (Wrighton).

<u>Specific Aim 3</u>: Construct a modular set of optimized genes (from Aims 1 and 2) on a DNA fragment containing specific regulatory elements that will allow high level gene expression in model organisms that have been flux optimized (Tabita, Wrighton, Cannon)).

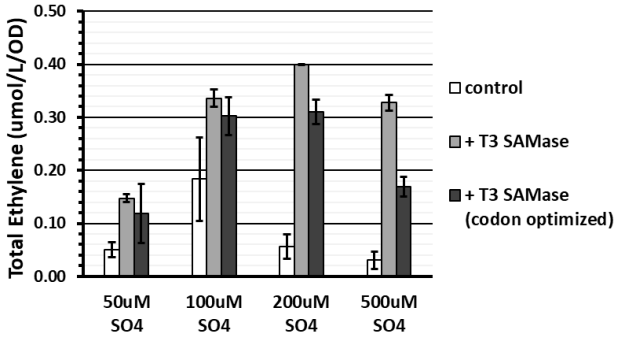


Fig. 2. Total ethylene produced by *R. palustris* when grown anaerobically in 0.5 % ethanol / 0.1 % bicarbonate minimal media supplemented with the indicated amount of ammonium sulfate. When either the native T3 coliphage SAM hydrolase (T3 SAMase) or the codon optimized gene for *R. palustris* was expressed, total ethylene production was enhanced up to 10-fold and suppression by non-limiting sulfate levels (≥ 200 µM) was overcome.

Recent Progress/Current Studies:

Aim 1: (A) Enhanced ethylene synthesis: Under conditions where cellular MTA formation is minimal, subsequent ethylene production is low. To circumvent this rate limiting formation of MTA, we have employed a novel viral enzyme, SAM hydrolase, from the T3 coliphage. This highly active enzyme (kcat ~ 300 / s) directly catalyzes SAM conversion to MTA and homoserine, and it is naturally used by the coliphage to evade the *E. coli* SAM-dependent restriction complex. When heterologously synthesized in *R. palustris*, the SAM hydrolase enzyme not

only enhanced ethylene production by up to 10-fold due to the specific conversion of native cellular SAM to MTA, it also overcame the regulatory effects of non-limiting extracellular sulfate levels on ethylene production (Fig. 2).

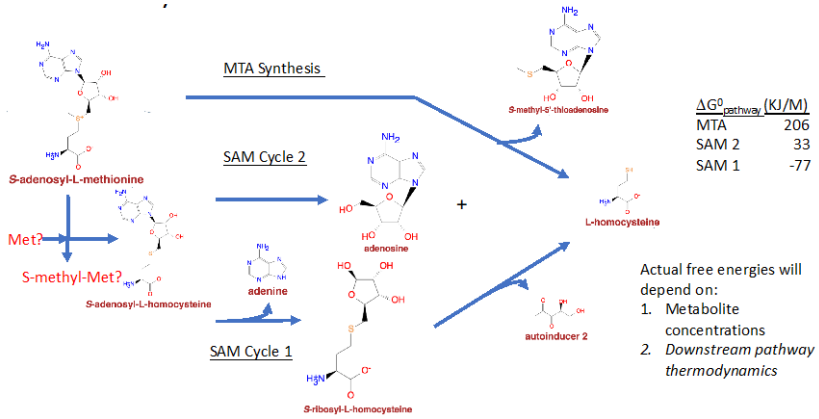


Fig.3. Development of kinetic and thermodynamic model for *R. palustris* metabolism.

(B) Computational Models: Computational models will be used to evaluate the thermodynamics of ethylene production and predict strategies for optimizing these conditions. Possible bottlenecks are first the redirection of SAM flux away from biomass production and towards MTA and subsequently ethylene production (see above). A second possible bottleneck may be the final steps in the ethylene pathway, which have not been characterized yet. We have already started developing a full kinetic and thermodynamic model of *R. palustris* metabolism including these pathways to evaluate the kinetics and thermodynamics (Fig. 3).

Aim 2: Discovery of more efficient enzymes of the anaerobic ethylene pathway: Previous studies had indicated that more effective aldolase ans isomerase genes could replace endogenous genes and substantially enhance ethylene levels (1, and unpublished studies). To scale up the search for ethylene-enhancing orthologs we apply targeted functional metagenomics to systematically query genes from the environment. Candidate genes will be synthesized for expression in our *R. palustris* genetic system and assayed in (i) a high throughput lysate activity assay and (ii) via physiological complementation to identify orthologs that enhance ethylene production. We have mined JGI IMG/M genome and metagenome sequence databases for candidate orthologs to the MTR-1P isomerase (*mtnA*) and the MTRu-1P aldolase (*ald2*) genes in collaboration with JGI scientists. The initial search identified 1,371,813 and 96,049 candidate genes for *mtnA* and *ald2*, respectively. Subsequent filtering based on synteny, phylogeny, and sequence homology with experimentally validated enzymes has yielded over 2500 candidate orthologs for each target gene, from multiple metagenomic samples covering a wide variety of environments including wetlands, forest soils, rhizosphere, and bioreactors. Candidates clustering with sequences from experimentally validated representatives, having both *mtnA* and *ald2* on the same contig, or having metatranscript data are ranked higher as synthesis targets than those identified by gene sequence alone. This approach should both ensure recovery of active enzymes and maximize sampling of undefined biochemical diversity.

References

1. North, J. A., Miller, A. R., Wildenthal, J. A., Young, S. J., and Tabita, F. R. Microbial pathway for anaerobic 5'-methylthioadenosine metabolism coupled to ethylene formation. Proc. Natl. Acad. Sci. U S A. 2017 Nov 28;114(48):E10455-E10464. doi: 10.1073/pnas.1711625114. Epub 2017 Nov 13.

Understanding the microbial controls on biogeochemical cycles in permafrost ecosystems

Neslihan Taş[1*], Megan Dillon[1], Yaxin Xue[1],Yaoming Li[1], Craig Ulrich[1], Yuxin Wu[1] and Hoi-Ying Holman[2]

Earth and Environmental Sciences Area[1], Biosciences Area[2], Lawrence Berkeley National Laboratory

Project Goals: Permafrost soils are one of the world's largest terrestrial carbon stores. The permafrost carbon reservoirs are currently protected from microbial decomposition by frozen conditions. Organic carbon becomes more available for mineralization by microorganisms as permafrost thaws. Microbial metabolism leads to decomposition of soil organic matter, substantially impacting the cycling of nutrients and significantly affecting the arctic landscape. This project use state of the art molecular techniques to resolve complex microbial processes governing the biogeochemical cycles in arctic soils and permafrost to better inform efforts to access uncertainties surrounding ecosystem responses.

Permafrost soils contain a broad diversity of cold-adapted microbes, whose metabolic activity depends on environmental factors such as temperature changes that cause cycles of freezing and thawing in the soil. Permafrost microbial communities are complex, diverse, and active at subzero temperatures. While carbon turnover at depth is proposed to be slower than surface, especially the fate of carbon in deep permafrost, which is currently protected from the warming climate, is uncertain. Permafrost microbiome is a seed bank of mostly novel organisms that have a diverse and broad metabolic potential. In-depth functional characterization of the permafrost microbes is needed to provide a foundation for understanding their responses to thaw. In order to address this gap in our knowledge we performed a pan-Arctic comparative analysis of permafrost metagenomes in which we study biogeography and metabolic functions of permafrost metagenomes assembled genomes (MAGs). This meta-analysis of permafrost MAGs across multiple locations in the Arctic (Alaska, Sweden, Norway, Canada and Russia) quantifies the impact of environmental drivers including but not limited to ice content, topography, continuity, active layer depth, and vegetation on permafrost microbiomes. The microbial communities inhabiting these soils have some common members and metabolic capacity, including the prominence of *Actinobacteria* and dominant functions such as fermentation and stress response. We also observe stark differences in which Eukaryotic and viral populations are represented and in the abundance and biochemistry of methanogens. These results improve our understanding of global variation in permafrost microbiomes. Recognizing geospatial patterns in soil properties and microbiome characteristics across Arctic permafrost landscapes will allow us to better inform on how permafrost microbes can respond to global climate change.

The microbial response to thaw in arctic environments is not uniform and the relationship between permafrost microbiomes and greenhouse gas (GHG) emissions is not well understood. Following

thaw, redistribution of water is a key event that conditions the permafrost for microbial decomposition. While inundation can give rise to low redox conditions and support anaerobic processes, drainage can promote aerobic respiration. Recent metagenomic studies conducted under different thaw scenarios suggest a capacity for utilization of many terminal electron acceptors (TEAs) across intact and thawed permafrost soils. These processes are likely to be active during thaw, although the impact of redox potential and the chemical form of TEAs on GHG emissions after permafrost thaw are not well understood. We initiated batch-scale permafrost incubation experiments dry, natural, and saturated moisture states and under microaerophilic or anaerobic headspaces. Our initial findings following microbial responses to permafrost thaw in short-term (i.e. seasonal) via metagenomics and metatranscriptomics show dominance of fermentation and competition between iron and sulfate reduction processes, highlighting the importance of interactions between iron, sulfur and carbon metabolism. We couple these small-scale manipulation experiments with 1-D controlled thaw and re-freeze soil column experiments. This experimental set-up preserve soil structure during thaw while enabling real time monitoring of GHG fluxes, soil pore water chemistry and microbial communities. After a simulation of thaw and re-freeze event over the whole soil column we analyzed the changes in microbial communities and in soil chemistry via synchrotron fourier transform infrared (sFTIR) spectral imaging at the Berkeley Infrared Structural Biology beamline of the Advanced Light Source (LBNL). Analysis showed that different organic compounds and metabolites were accumulated in post-thaw/refreeze seasonally active top soils (active layer) and permafrost soils. While carbohydrates were enriched in the active layer, permafrost layers showed an accumulation of aliphatic compounds and depletion in carbohydrates.

This project use field observations, laboratory manipulations, and multi-omics approaches to examine how microbial processes, biogeochemical transformations, and hydrology interact during permafrost thaw in different sites in Alaska in order to determine how these factors drive biogeochemical cycles in different arctic soils.

This research is supported by DOE Early Career Program by the Office of Biological and Environmental Research in the DOE Office of Science.

Harnessing robustness of *Yarrowia lipolytica* for effective utilization of undetoxified biomass hydrolysates

Seunghyun Ryu[1*] (sryu2@utk.edu), Josh C. Quaterman,[2] Caleb Walker[1], Richard Giannone,[3] Patricia J. Slininger,[2] Bruce S. Dien,[2] and **Cong T. Trinh**[1]

[1]University of Tennessee, Knoxville, TN 37996; [2]National Center for Agricultural Utilization Research, USDA-ARS, Peoria, IL 61604; [3]Oak Ridge National Laboratory, Oak Ridge, TN 37831

Project Goals: To elucidate and harness the exceptional robustness of novel, undomesticated *Y. lipolytica* isolates from genetic diversity screening as a bioenergy-relevant microbial platform for efficient conversion of undetoxified biomass hydrolysates into designer bioesters with continuous recovery using solvent extraction.

Abstract text. Robustness is an important phenotype for bioenergy microbes to acquire but is difficult to engineer. Even though the oleaginous yeast *Yarrowia lipolytica* has been extensively studied for high lipid accumulation, its robustness to enable effective utilization of undetoxified biomass hydrolysates is not fully understood. In our study, we first discovered that *Y. lipolytica* has native metabolism, capable of degrading complex sugars such as xylose, arabinose, and cellobiose that are dominant in plant biomass. Through bioinformatic, metabolic, and transcriptomic analyses of *Y. lipolytica*, we identified and activated the endogenous genes encoding for metabolic enzymes and sugar-specific transporters that assimilate these complex sugars. Interestingly, we found that xylose assimilation is synergistically activated by glucose with xylitol accumulation whereas arabinose assimilation is synergistically activated by xylose with arabitol accumulation. By tuning the expression of native xylitol dehydrogenase and arabitol dehydrogenase of *Aspergillus oryzae* in *Y. lipolytica*, the engineered strains could enhance xylose and arabinose utilization and eliminate xylitol and arabitol accumulation. Remarkably, by tuning expression of sugar-specific transporters, the assimilation efficiency of not only the pentose sugars but also cellobiose was improved. To further enhance robustness of *Y. lipolytica* for effectively utilizing undetoxified biomass hydrolysates, we next screened a comprehensive set of 57 undomesticated *Y. lipolytica* isolates belonging to the 13 unique species in *Yarrowia* clade. Unlike the model strain W29, we found some *Yarrowia* isolates, *Y. lipolytica* YB-392, YB-419, and YB-420, capable of exhibiting robust growth and lipid accumulation in up to 90% undetoxified dilute acid-pretreated switchgrass hydrolysates. Some isolates, e.g., *Candida phangngensis* PT1-17, were also identified as the top lipid producers that can accumulate neutral lipids up to 3-fold higher than W29. We sequenced the whole genome of these novel *Y. lipolytica* strains and in the process of activating and rewiring their robust metabolism for improved complex sugar assimilation, lipid accumulation, and designer bioester synthesis in undetoxified biomass hydrolysates.

References[1-5]

1 Ryu, S., Hipp, J. & Trinh, C. T. Activating and elucidating metabolism of complex sugars in *Yarrowia lipolytica*. *Applied and environmental microbiology* **82**, 1334-1345 (2016).

2 Ryu, S. & Trinh, C. T. Understanding Functional Roles of Native Pentose-Specific Transporters for Activating Dormant Pentose Metabolism in *Yarrowia lipolytica*. *Applied and Environmental Microbiology*, doi:10.1128/aem.02146-17 (2017).
3 Quarterman, J., Slininger, P. J., Kurtzman, C. P., Thompson, S. R. & Dien, B. S. A survey of yeast from the *Yarrowia* clade for lipid production in dilute acid pretreated lignocellulosic biomass hydrolysate. *Applied Microbiology and Biotechnology*, 1-16 (2016).
4 Walker, C. *et al*. Draft Genome Assemblies of Five Robust *Yarrowia lipolytica* Strains Exhibiting High Lipid Production, Pentose Sugar Utilization, and Sugar Alcohol Secretion from Undetoxified Lignocellulosic Biomass Hydrolysates. *Microbiol Res Announc* **7**, e01040-01018 (2018).
5 Quarterman, J. C., Slininger, P. J., Hector, R. E. & Dien, B. S. Engineering *Candida phangngensis*—an oleaginous yeast from the *Yarrowia* clade—for enhanced detoxification of lignocellulose-derived inhibitors and lipid overproduction. *FEMS yeast research* **18**, foy102 (2018).

Funding statement. This work is supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under Award Number DESC0019412.

Understanding Exceptional Ionic Liquid Tolerance in *Yarrowia lipolytica*

Caleb Walker[1*] (fyz692@vols.utk.edu), Seunghyun Ryu,[1] and **Cong T. Trinh**[1]

[1]University of Tennessee, Knoxville, TN 37996

Project Goals: To elucidate and harness the exceptional robustness of novel, undomesticated *Y. lipolytica* isolates from genetic diversity screening as a bioenergy-relevant microbial platform for efficient conversion of undetoxified biomass hydrolysates into designer bioesters with continuous recovery using solvent extraction.

Abstract text. Microbial biocatalysis in organic solvents such as ionic liquids (ILs) is attractive for making high-value chemicals. However, IL toxicity at a level of 0.5% ~1% (v/v) can drastically reduce microbial activity. In this study, we engineered a mutant *Yarrowia lipolytica* YlCW001 that can thrive in up to 18% (v/v) 1-ethyl-3-methylimidazolium acetate ([EMIM][OAc]), lethal to almost all microorganisms. Remarkably, YlCW001 also exhibits universal tolerance in most commonly used ILs beyond [EMIM][OAc]. Scanning electron microscopy revealed that ILs significantly damage cell wall and/or membrane of wildtype *Y. lipolytica* while YlCW001 maintains healthy cellular morphology even in high concentration of ILs up to 18% (v/v). By performing comprehensive metabolomics, lipidomics, and transcriptomics to elucidate this unique phenotype, we discovered that both wildtype *Y. lipolytica* and YlCW001 prominently exhibit upregulation of most glycerophospholipids (GPs), sphingolipids, and sterols under IL-stressful environment. However, the mutant reconfigured membrane composition and structure by increasing the content of GPs and sterols more than the wildtype. By modulating the sterols pathway, we validated that sterols is a key component of the cell membrane that enables *Y. lipolytica* to tolerate high IL concentrations. This study provides a fundamental understanding of exceptional robustness of *Y. lipolytica* and helps guide metabolic engineering of *Y. lipolytica* as a microbial manufacturing platform for production of high-value chemicals in organic solvents.

References[1,2]

1 Walker, C., Ryu, S. & Trinh, C. T. Exceptional Solvent Tolerance in Yarrowia lipolytica Is Enhanced by Sterols. *bioRxiv*, 324681 (2018).

2 Ryu, S., Labbé, N. & Trinh, C. T. Simultaneous saccharification and fermentation of cellulose in ionic liquid for efficient production of α-ketoglutaric acid by Yarrowia lipolytica. *Applied microbiology and biotechnology* **99**, 4237-4244 (2015).

Funding statement. This work is supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under Award Number DESC0019412.

Employing Bacterial Microcompartments To Create Privileged Redox Pools

for Biofuel Production

Sasha Shirman,[1] Taylor Dickman,[1] Svetlana Ikonomova,[1] Keith Tyo,[1] Niall Mangan,[1] and **Danielle Tullman-Ercek**[1*] (ercek@northwestern.edu)

[1]Northwestern University, Evanston, Illinois

Project Goals: To compartmentalize metabolic pathways along with enzyme cofactor recycling pathways to increase the yield and efficiency of bioproduction processes

https://dtelab.northwestern.edu/research/#nanobioreactors

Metabolic engineering holds great promise for creating efficient, competitive routes for the production of biofuels and biochemicals without the necessity for harsh chemicals and hazardous byproducts. Successes in biochemical production include the production of Dupont's Sorona fibers from 1,3-propanediol from glucose using bacteria and the manufacture of the anti-malarial drug artemisinin from yeast. However, roadblocks to biosynthesis prevent many biochemicals from being produced biologically given current technology. Nature uses compartmentalization (eg in organelles in eukaryotes and in bacterial microcompartments in prokaryotes) to solve issues such as intermediate leakage, toxicity, and byproduct formation. Here we propose to deploy compartmentalization as a strategy to overcome a critical roadblock: the requirement for redox cofactor recycling. In traditional systems, redox cofactors are lost to cellular growth and maintenance needs. By compartmentalizing redox cofactors with the biochemical synthesis enzymes, we anticipate increasing the thermodynamic efficiency and preventing the loss of valuable intermediates and cofactors. If successful, it would be the first direct demonstration of this feature of a bacterial microcompartment, and would provide a tool for improving metabolic pathway performance for all enzymes with redox or other cofactors.

With this poster, we will describe how we are coupling modeling with experiments to improve the performance of our first target metabolic pathway: 1,3-propanediol production. We will also show how modeling is leading to new insight into the native function of these structures.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under Award Number DE-SC0019337.

Lignin Removal Controls Kinetics of Biomass Deconstruction by Consolidated Bioprocessing

Christian Alcaraz[1*] (calca005@ucr.edu), Priyanka Singh[1], Ninad Kothari[1], Charles Cai[1], Charles E. Wyman[1], and **Gerald A. Tuskan[2]**

[1]University of California, Riverside; [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Pretreatment or cotreatment is necessary for biological conversion of lignocellulosic biomass to realize high yields of sugars that can be converted to liquid fuels. Most effective pretreatments apply mechanical energy and/or chemicals to disrupt the cellulose/hemicellulose/lignin matrix in plant cell walls so enzymes can access polysaccharides for deconstruction to sugars. Current application of dilute acid retains most of the cellulose and lignin in the solids while hemicellulose is released into the liquid with high yields. Then, cellulase and accessory enzymes can saccharify cellulose and any other polysaccharides remaining in the solids into monomeric sugars. However, substantial negative impacts on enzymatic hydrolysis yields have been attributed to the lignin left in the solids, particularly at economically viable enzyme loadings. For dilute acid pretreatment, application of high temperatures and long times can effectively disrupt lignin, but enzyme loadings are still too high to be economically attractive. In addition, monomeric sugars released during pretreatment degrade if the severe combinations of temperature and time needed to improve glucose release at lower enzyme loadings are employed. Therefore, conventional pretreatment has suffered from balancing between application of sufficiently severe conditions to disrupt lignin while minimizing sugar degradation.

Co-solvent Enhanced Lignocellulosic Fractionation (CELF) is a chemical pretreatment that applies miscible THF-water mixtures as a co-solvent to reduce the temperature needed for dilute acid to be effective while also removing over 80% of lignin. In addition, CELF can also recover nearly 100% of the sugars from the combined operations of pretreatment and enzymatic hydrolysis even for application of enzyme loadings as low as 2 mg of protein /g glucan, albeit for times approaching one month for such low loadings. Additionally, CELF releases distinctive

lignin fractions from the solids that can be targeted for further processed to fuels, chemicals, and/or materials, thereby achieving more complete utilization of these 3 major biopolymers.

Combining enzyme production, saccharification, and fermentation in a single organism by a process known as CBP would reduce costs for biological conversion of biomass to fuels and chemicals. The thermophilic anaerobe *Clostridium thermocellum* has proven to be a particularly effective CBP organism that can achieve nearly complete breakdown of the polysaccharides left in solids from CELF pretreatment in only 2 days. Integration of CELF with C*lostridium thermocellum* therefore offers the potential to realize high yields from the polysaccharides in lignocellulosic biomass in short times while also recovering a large portion of the lignin for potential conversion to valuable products.

The CELF-CBP combination has been shown to be highly effective for virtually 100% sugar solubilization from switchgrass, corn stover, and poplar while removing over 80% of lignin. However, it is not yet clear what changes in CELF pretreated solids are responsible for CBP realizing such high yields in short times compared to enzymes. For instance, *C. thermocellum* releases more sugars than fungal enzymes from solids produced by CELF pretreatment at mild conditions that remove less lignin. Thus, *C. thermocellum,* appears to be able to bypass lignin in the solids to some extent despite not possessing lignolytic enzymes. In this study, CBP was applied to CELF solids with varying degrees of lignin removal to better understand the impact of the lignin on CBP and key differences that enhance deconstruction of CELF solids by CBP compared to enzymes. Solids were CELF pretreated at 150°C for 5, 15, and 25 minutes at 7.5 wt% solids loading in a 1:1 THF:water (w/w) co-solvent ratio that also contained 0.5 wt% sulfuric acid. Afterwards, the pretreated solids were subjected to CBP for 7 days with samples taken at 12 hours and every 24 hours thereafter. The results showed that despite applying CELF at 150°C for only 5 minutes, *Clostridium thermocellum* was still able to release 95% of the sugars in 3 days, while for solids pretreated at 15 and 25 minutes, sugar release reached 98% in 4 days and 100% in 3 days, respectively. Kinetic models were applied to predict polysaccharide deconstruction during CBP and understand the impact of removing different fractions from the biomass matrix on rate of the rate of breakdown. Comparison of the impact of xylan removal on the rate and extent of cellulose deconstruction by CBP to that by lignin release revealed that lignin had a greater effect on the effectiveness of biological deconstruction by both *C. thermocellum* and fungal enzymes. Overall, this work provides new insights into factors governing biomass deconstruction by CELF-CBP that can lead to more economical biofuels production.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Mechanisms of Plant Cell Wall Polysaccharide Biosynthesis: A Two-Phase Model for the Non-Processive Biosynthesis of Homogalacturonan Polysaccharides by the GAUT1:GAUT7 Complex

Robert A. Amos[1*] (baa@uga.edu), Debra Mohnen[1], and **Gerald A. Tuskan[2]**

[1]University of Georgia, Athens; [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Hydrolysis and fermentation of plant cell wall polysaccharides produces fuel products from biomass feedstocks. Cell walls in mature plant tissues contain a heterogeneous mixture of cellulose, hemicellulose, and pectin as well as arabinogalactan protein-linked polysaccharides. Each of these cell wall components contributes to a highly-crosslinked matrix of glycan polymers that resists enzymatic and chemical depolymerization.

Pectic polysaccharides are deposited early during cellular development and are enriched in the adhesive middle lamella between adjacent cells and in the primary wall. Pectic crosslinking interactions control cell wall stiffness, porosity, cellular adhesion, and cell expansion. Recent efforts within CBI and the BioEnergy Science Center have shown that transgenic modification of the expression of pectin biosynthetic genes in woody and grass feedstocks (poplar and switchgrass) affect phenotypes related to growth and recalcitrance.

The pectic glycan, homogalacturonan (HG), is synthesized by the galacturonosyltransferase (GAUT) gene family. Two members of this family, GAUT1 and GAUT7, form a heteromeric enzyme complex in Arabidopsis. Here, we established a heterologous GAUT expression system in HEK293 cells and show that co-expression of recombinant GAUT1 with GAUT7 results in the production of a soluble GAUT1:GAUT7 complex that catalyzes elongation of HG products *in vitro*.

Study of GAUT1:GAUT7 activity *in vitro* revealed that the complex synthesizes high-molecular-weight polymeric HG (> 100 kDa). Unlike cellulose synthase, GAUT1:GAUT7 is a non-

processive enzyme, and tight control of the size of the HG products can be achieved by varying substrate concentrations during *in vitro* reactions. Unexpectedly, small changes to the degree of polymerization (DP) of acceptor oligosaccharides resulted in major differences in reaction rates and in the apparent mechanism of synthesis. GAUT1:GAUT7 displayed > 45-fold increased catalytic efficiency with DP11 acceptors relative to DP7 acceptors, and reactions primed with short-chain acceptors resulted in a bimodal product distribution of glycan products that has previously been reported as evidence for a processive model of GT elongation.

As an alternative to the processive glycosyltransfer model, we propose a two-phase, non-processive elongation model. A slow phase, which includes the *de novo* initiation of HG and elongation of short-chain acceptors, is distinguished from a phase of rapid elongation of intermediate and long-chain acceptors. Upon reaching a critical chain length of DP11, GAUT1:GAUT7 elongates HG to high molecular weight products.

Non-processive GTs also synthesize several other classes of extended plant cell wall polysaccharide chains, particularly xylans, arabinogalactans, and the RG-I backbone. Key enzymes in these biosynthetic pathways have been identified, but the mechanisms of high MW polysaccharide elongation remain to be studied. The preference for longer acceptor substrates has been observed in other GT families, but the mechanistic significance of these findings has not been previously appreciated. We propose that the two-phase, non-processive mechanism observed during the study of GAUT1:GAUT7 activity may apply to the biosynthesis of other polysaccharides. Establishment of HEK293 cells as a high-yield heterologous expression system for GTs should enable the study of enzymatic mechanisms of plant cell wall synthesis, functional redundancy among homologous enzymes, polysaccharide elongation length control, and the role of GT complexes in the synthesis of high MW polymers. Mechanistic understanding of the synthesis of HG, with its critical roles in plant growth, biomass yield, and cell wall architecture will support the development of high-yielding, robust bioenergy feedstocks.

References

1. Amos, R.A., Pattathil, S., Yang, J.-Y., Atmodjo, M.A., Urbanowicz, B.R., Moremen, K.W., Mohnen, D., A two-phase model for the non-processive biosynthesis of homogalacturonan polysaccharides by the GAUT1:GAUT7 complex. (2018) *J. Biol. Chem.* 293:19047-19063.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Lignin Valorization: Biosynthesis and Bioengineering of an Ideal Lignin

Fang Chen[1,2]* (fang.chen@unt.edu), Chan Man Ha[1,2], Xirong Xiao[1,2], Chunliu Zhuo[1], Aaron Harkelroad[1], Chung-Jui Tsai[3], Richard A. Dixon[1,2], and **Gerald A. Tuskan[2]**

[1]University of North Texas, Denton; [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; [3]University of Georgia, Athens

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Natural lignins are generally composed of *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units, that are biosynthesized by polymerization of three primary monolignols, p-coumaryl, coniferyl, and sinapyl alcohols, respectively. C-lignin is a polymer of caffeyl alcohol, found in the seed coats of a number of exotic plant species [1,2]. C-lignin exists as a linear homopolymer with only one type of benzodioxane linkage, with promising properties for generation of carbon fibers and high value chemicals [3-5].

In the seed coats of the ornamental plant *Cleome hassleriana*, lignin composition switches from conventional G-lignin to C-lignin during development. Bioinformatic analysis of RNA sequencing data identified a complete set of lignin biosynthesis genes for *Cleome*. Transcript analysis, coupled with kinetic analysis of recombinant enzymes in *E. coli*, revealed that the switch to C-lignin formation was accompanied by down-regulation of transcripts encoding functional caffeoyl CoA- and caffeic acid 3-O-methyltransferases (CCoAOMT and COMT) and a cinnamyl alcohol dehydrogenase (CAD) isoform with preference for coniferaldehyde as substrate, and up-regulation of another CAD isoform with preference for caffealdehyde. Blockage of lignin monomer methylation by down-regulation of both OMTs appears to be the major factor in diversion of flux to C-lignin in the *Cleome* seed coats, although the change in CAD specificity also contributes to the shift in lignin composition based on transgenic manipulation of CAD.

Based on this information, we started to engineer the C-lignin biosynthesis pathway in the model bioenergy crop poplar. Since the *ccoaomt comt* double mutant shows severely retarded growth in *Medicago truncatula* and *Arabidopsis*[6], we have designed to downregulate CCoAOMT or COMT, and at the same time, overexpress the CAD with preference for caffealdehyde in

Medicago hairy roots. Our preliminary results show that it is necessary to downregulate both methyltransferases to direct lignin pathway to C-lignin precursor since completely blocking COMT alone while overexpressing CAD does not lead to the C-lignin accumulation. We are also generating transgenic plants to express C-lignin pathway specific transcription factors and genes related to lignin polymerization. At the same time, we started to generate transgenic poplar plants in which the lignin biosynthesis pathway genes are altered through CRISPR gene-editing technology. These plants will be used for gene pyramiding later so that carbon flux will be redirected to the C-lignin precursor, the caffeyl alcohol and C-lignin biosynthesis.

References

1. Chen F, Tobimatsu Y, Havkin-Frenkel D, Dixon RA, Ralph J (2012) A polymer of caffeyl alcohol in plant seeds. Proceedings of the National Academy of Sciences 109: 1772-1777.
2. Tobimatsu Y, Chen F, Nakashima J, Escamilla-Treviño LL, Jackson L, et al. (2013) Coexistence but independent biosynthesis of catechyl and guaiacyl/syringyl lignin polymers in seed coats. The Plant Cell Online 25: 2587-2600.
3. Nar M, Rizvi HR, Dixon RA, Chen F, Kovalcik A, et al. (2016) Superior plant based carbon fibers from electrospun poly-(caffeyl alcohol) lignin. Carbon 103: 372-383.
4. Li Y, Shuai L, Kim H, Motagamwala AH, Mobley JK, et al. (2018) An “ideal lignin” facilitates full biomass utilization. Science Advances 4: eaau2968.
5. Stone ML, Anderson EM, Meek K, Reed ML, Katahira R, et al. (2018) Reductive Catalytic Fractionation of C-Lignin. ACS Sustainable Chemistry & Engineering.
6. Do C-T, Pollet B, Thévenin J, Sibout R, Denoue D, et al. (2007) Both caffeoyl Coenzyme A 3-O-methyltransferase 1 and caffeic acid O-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in Arabidopsis. Planta 226: 1117-1129.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. Part of this work was supported by grants from the National Science Foundation Integrated Organismal Systems program (Award number 1456286).

Expression Quantitative Trait Nucleotide (eQTN) Mapping Reveals Transcriptional Regulatory Networks in *Populus trichocarpa*

Jin Zhang[1], Jay Chen[1*] (chenj@ornl.gov), Meng Xie[1], Yongil Yang[1], Kaijie Zheng[1], Lee Gunter[1], Sara Jawdy[1], Timothy J. Tschaplinski[1], Nancy Engle[1], Erika Lindquist[2], Vasanth Singan[2], Kerrie Berry[2], Jeremy Schmutz [2,3], Wellington Muchero[1], and **Gerald A. Tuskan[1]**

[1]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; [2]U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA; [3]HudsonAlpha Institute for Biotechnology, Huntsville, AL

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Populus are being used for biomass production for a suite of industrial applications including biofuels conversion. Our understanding of biomass productivity and quality is limited by the fact that this complex trait requires the regulation and coordinated interactions of many genes. Identification of genetic networks regulating biomass productivity and quality remains largely unaccomplished and is urgently needed to inform genetic improvement of *Populus* feedstocks for biomass production and conversion.

To uncover the genetic regulatory landscape in the woody perennial bioenergy crop *Populus trichocarpa*, we performed an expression quantitative trait nucleotide (eQTN) mapping enabled by the whole-genome resequencing and RNA-seq analysis of *P. trichocarpa*. natural variants. A panel of >8.2 million single nucleotide polymorphisms (SNPs) and nucleotide insertions and deletions (InDels) were obtained from whole-genome resequencing of 917 unrelated individuals of *P. trichocarpa*. Transcriptome data from 390 leaf and 444 xylem samples were analyzed and revealed that 16,030 and 15,496 genes, respectively, exhibited significant expression variation across the population. Through genetic mapping, *cis*- and *trans*-eQTNs were identified. Enriched transcription factor binding sites (TFBS) including *cis*-eQTN showed tissue-specific divergence. *trans*-eQTN analysis identified multiple hotspots that were significantly associated with expression of more than 100 putative target genes. Combined with genome-wide association studies (GWAS) of trait phenotypes, the upstream regulators of these phenotype-associated genes and their regulatory network were identified. These analyses have provided a comprehensive understanding of the genetic regulatory mechanisms underlying complex traits.

References

1. Zhang, J. et al. (2018). GWAS and eQTL analyses reveal roles of *HCT2* in caffeoylquinic acid biosynthesis and its regulation by defense-responsive transcription factors in *Populus*. *New Phytologist* 220: 502-516.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Engineering CRISPR/Cas Systems for Genome Editing in *Clostridium thermocellum*

Julie Walker[1], Anthony Lanahan[2], Emily Freed[1], Jacob Fenster[1], Jeff Cameron[1], Dan Olson[2], Lee Lynd[2], Carrie Eckert[1,3]* (carrie.eckert@nrel.gov), and **Gerald A. Tuskan**[4]

[1]University of Colorado, Boulder; [2]Dartmouth College, Hanover; [3]National Renewable Energy Laboratory, Golden, CO; [4]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN

http://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI has identified key barriers for the current bioeconomy in (1) high-yielding, robust feedstocks, (2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and (3) methods to create valuable byproducts from the lignin residues. CBI will identify and utilize key plant genes for growth, yield, composition and sustainability traits as a means of achieving lower feedstock costs, focusing on the perennial feedstocks - poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and esters) using CBP at high rates, titers and yield. CBP will be combined with cotreatment or pretreatment. CBI will maximize process and product value by *in planta* modifications of lignin and biological funneling of lignin to value-added chemicals. Techniques for rapid domestication of non-model microbes and plants will be improved.

Efficient microbial conversion of lignocellulose to fuels and chemicals is key to an economically viable bioproduction platform, particularly if coupled with feedstocks designed for optimal microbial performance. At their core, current approaches for accelerated domestication of model organisms such as *E. coli* mirror the cycle of "Design, Build, Test," revealing underlying design principles that inspire creative solutions to modern engineering challenges. Although understanding of complex biological systems and "design rules" for metabolic engineering of model organisms have progressed significantly, we are still unable to effectively engineer many desired properties necessary for a sustainable and efficient bioprocessing platform in non-model microbes and plant feedstocks. Therefore, it is essential that we expand genetic engineering capabilities in novel non-model systems with desired traits for bioprocessing to accelerate advancement of the U.S. bioeconomy.

Advances in DNA synthesis and powerful genome editing tools such as CRISPR (**C**lustered **R**egularly-**I**nterspaced **S**hort **P**alindromic **R**epeats) have led to a rapid expansion of genome engineering in model and now increasingly in non-model organisms. Our goal is to develop efficient CRISPR-mediated genome editing systems to enable these and other cutting-edge genome editing technologies in CBI microbes and feedstocks. Although the *Streptomyces pyogenes* Cas9 system has been utilized across a number of microbial and eukaryotic platforms, unfortunately it is not active in the growth temperature range of *Clostridium thermocellum* (55-60^{o}C), a thermophilic bacterium capable of directly converting cellulose to sugars, bypassing the need for chemical processing of lignocellulosic feedstocks. Diverse CRISPR systems exist amongst Archaea and Bacteria, and the genome of *C. thermocellum* DSM1313 encodes two Type III and one Type I CRISPR/Cas systems. Unlike Type II systems (e.g. Cas9), which require only a single Cas protein for targeting (via the guide RNA, or gRNA), Type I and III systems require a multi-subunit Cas complex. Since Type III systems generally target RNA instead of DNA, we focused on characterizing the native Type I system as a tool for genome editing in *C. thermocellum.* We have been exploring the Type I-B CRISPR system in *Clostridium thermocellum* with the goal of using it for gene deletion and editing. In order to overcome sensitivity issues associated with modest expression of the endogenous CRISPR locus genes and spacers in *C. thermocellum,* we constructed an artificial CRISPR spacer expression cassette to probe CRISPR activity. We used this construct in a plasmid

transformation inhibition assay to test and validate predicted PAM sequences. This assay coupled with a PAM library depletion assay allowed us to identify the optimal PAM sequence of 5'-T (preferred, or C)/N/A, G or T-3' in *C. thermocellum*. In efforts to further enhance/optimize CRISPR activity in *C. thermocellum*, we inserted heterologous promoters upstream of the Cas operon. Quantitative PCR analysis showed upregulation to various degrees of the operon in all strains relative to the parental strain LL1299. This upregulation correlated with an increase in CRISPR activity as assayed in the plasmid transformation inhibition assay. Thus far efforts at gene editing have focused on the *pyrF* locus. We show a spacer dependent decrease in transformation, likely due to targeting of the wildtype *pyrF* locus. As expected, this decrease in transformation is higher in the promoter engineered strains where the Cas operon is expressed at elevated levels relative to those in LL1299. As seen in other *Clostridium* strains, we find that serial transfer of individual transformants increases the appearance of the pyrF deletion. allele. Current efforts are aimed at increasing the efficiency of CRISPR editing.

In addition, three thermophilic Cas9 systems have recently been characterized (1–3) and were additionally evaluated for activity/genome editing in *C. thermocellum*. A toxicity assay indicated that the *Geobacillus stearothermophilus* thermophilic Cas9 was the most active in *C. thermocellum.* However, multiple attempts for CRISPR/Cas9 homology directed genome repair were unsuccessful most likely due to unregulated expression of Cas9 as well as the low efficiency of homologous recombination in *C. thermocellum.* To overcome the first limitation, we placed Cas9 expression under a tightly regulated inducible promoter. To overcome the second limitation, recombineering machinery was isolated from the thermophilic organism *Acidithiobacillus caldus* and expressed in *C. thermocellum*. An increase in homologous directed repair was observed in *C. thermocellum* strains expressing recombineering machinery when compared to parental strains. We are now poised to enable rapid CRISPR/Cas9 genetic engineering in *C. thermocellum*.

We expect that CRISPR systems can be harnessed to not only accelerate precise genome editing in CBI organisms, but also to create targeted, *trans*-acting regulatory systems as has been demonstrated in other microbes including CRISPR interference (CRISPRi) for targeted knockdown of genes of interest to begin to determine gene-to-trait attributes towards desired phenotypes. Preliminary data on these evaluations will be presented.

Ultimately, we aim to utilize these CRISPR/Cas systems for a rapid, HTP method for phenotype-to-genotype discovery in *C. thermocellum* such as: 1) rational protein engineering, 2) complete residue substitution libraries, 3) pathway optimization and 4) discovery of new gene functions by genome-wide targeting strategies. These tools will expand and accelerate the canonical the "Design, Build, Test" cycle for gene-to-trait discovery in support of CBI research needs.

References

1. Mougiakos I, Mohanraju P, Bosma EF, Vrouwe V, Finger Bou M, Naduthodi MIS, et al. Characterizing a thermostable Cas9 for bacterial genome editing and silencing. BioRxiv. 2017 Aug 18;
2. Harrington L, Paez-Espino D, Chen JS, Ma E, Staahl BT, Kyrpides NC, et al. A thermostable Cas9 with increased lifetime in human plasma. BioRxiv. 2017 May 16;
3. Liu Q, Gao R, Li J, Lin L, Zhao J, Sun W, et al. Development of a genome-editing CRISPR/Cas9 system in thermophilic fungal Myceliophthora species and its application to hyper-cellulase production strain engineering. Biotechnol Biofuels. 2017 Jan 3;10:1.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

CRISPR-Tool Development in *Pseudomonas putida* KT2440 for High-Titer Strain Engineering and Multiplexed Approaches

Jacob A. Fenster[1]* (jacob.fenster@colorado.edu), Audrey Watson[1], Jeff Cameron[1,3], Carrie Eckert[1,3], and **Gerald A. Tuskan[2]**

[1]University of Colorado Boulder, Boulder; [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; [3]Renewable and Sustainable Energy Institute, Boulder, CO

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

The programmability of CRISPR-Cas machinery enables library based high-throughput and multiplexed experiments that have allowed for rapid genotype-phenotype mapping (Peters et al., 2016), protein engineering (Garst et al., 2017), strain engineering (Tarasava, Liu, Garst, & Gill, 2018), and gene discovery (Shalem et al., 2014). Both CRISPR-Cas editing and interference have been demonstrated in the bacteria *Pseudomonas putida* KT2440, a promising candidate for the industrial production of renewable chemicals from lignin. Although these studies establish the foundation for CRISPR-tool development in *P. putida*, the need remains for characterization and optimization of these tools before they can be leveraged for high-throughput and multiplexed experiments.

In this study, CRISPR-interference (CRISPRi) is optimized by screening inducible promoter systems that express catalytically-dead spCas9. The arabinose inducible promoter system performed best and was used to image the repression of the essential division protein ftsZ in real time. Future studies will quantify the dynamic range of repression by targeting a gnomically integrated fluorescent reporter gene. In addition, this toolset will be used to increase 2-pyrone 4,6-dicarboxylic acid (PDC), β-ketadipate (βKA), and muconate titers by temporally repressing downstream enzymes that funnel these desired products into central metabolism.

To optimize CRISPR-Cas gene editing, the transformation protocol from Sun et al. 2018 was improved to increase the colony forming units by 100-fold while keeping the editing efficiency

at 100%. The minimum homology arm (HA) length requirements for gene deletion was determined and the HA length requirements for gene integration and introducing single codon mutations are underway.

Together these tool optimization studies will enable high-throughput and multiplexed gene editing and gene silencing experiments to be conducted in *P. putida*. This will be leveraged to rapidly map genotype-phenotype relationships as well as engineer high-titer strains.

References

1. Garst, A. D., Bassalo, M. C., Pines, G., Lynch, S. A., Halweg-Edwards, A. L., Liu, R., … Gill, R. T. (2017). Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nature Biotechnology*, *35*(1), 48–55. doi:10.1038/nbt.3718
2. Peters, J. M., Colavin, A., Shi, H., Czarny, T. L., Larson, M. H., Wong, S., … Gross, C. A. (2016). A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria. *Cell*, *165*(6), 1493–1506. doi:10.1016/j.cell.2016.05.003
3. Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelson, T., … Zhang, F. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, *343*(6166), 84–87. doi:10.1126/science.1247005
4. Sun, Jun, et al. "Genome editing and transcriptional repression in Pseudomonas putida KT2440 via the type II CRISPR system." *Microbial cell factories* 17.1 (2018): 41.
5. Tarasava, K., Liu, R., Garst, A., & Gill, R. T. (2018). Combinatorial pathway engineering using type I-E CRISPR interference. *Biotechnology and Bioengineering*, *115*(7), 1878–1883. doi:10.1002/bit.26589

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Effectiveness of Cotreatment during Fermentation on Various Feedstocks, Impact on Different Microbes, and Changes in Biomass Physical Properties

Sanchari Ghosh[1*] (sanchari.ghosh.th@dartmouth.edu), Michael L. Balch[1], Galen D. Moynihan[1], Evert K. Holwerda[1], Brenden P. Epps[1], Lee R. Lynd[1], and **Gerald A. Tuskan[2]**

[1]Dartmouth College, Hanover; [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Milling during fermentation, also known as cotreatment[1], is being investigated as an alternative to thermochemical pretreatment for increasing the accessibility of lignocellulose to biological attack. Here we examine the effectiveness of cotreatment in combination with CBP fermentation by *Clostridium thermocellum* at low solids concentrations using switchgrass (with and without genetic modification aimed at decreasing recalcitrance), using more and less recalcitrant natural variants of *Populus*, and using corn stover and corn fiber. The ability to carry out fermentation in the presence of aggressive milling, previously assessed only for *Clostridium thermocellum,* is evaluated for *Thermoanaerobacterium saccharolyticum, Escherichia coli*, *Bacillus subtilis,* and *Zymomonas mobilis.* The substantial change in viscosity of corn stover slurries over the course of *Clostridium thermocellum* fermentation is evaluated with initial solids loading representative of an industrial process[2]. Results from screening various milling modalities for cotreatment effectiveness are also reported.

References

1. Balch *et al*. (2017), *Energy Environ. Sci.*, 10:1252-1261 doi:10.1039/C6EE03748H
2. Ghosh *et al*. (2018), *Biotechnol Biofuels*, 11:246. https://doi.org/10.1186/s13068-018-1248-z

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Isolation and Rapid Domestication of Non-Model Microbial Hosts for Biofuels Production

Adam M. Guss[1,2]* (gussam@ornl.gov), Lauren A. Riley[1,2], E. Anne Hatmaker[1], Kaela B. O'Dell[1], Neely M. Wood[1], Olivia N. Cannon[1], Robert J. Schmitz[3], Janet Westpheling[3], James G. Elkins[1], and **Gerald A. Tuskan**[1]

[1]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; [2]University of Tennessee, Knoxville; [3]University of Georgia, Athens

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Many organisms naturally possess complex physiological traits that are of interest for biotechnology research. Often, these traits are challenging to transfer into traditional host organisms such as *Escherichia coli* and *Saccharomyces cerevisiae*. Therefore, the ability to identify and rapidly domesticate non-model organisms to harness these traits could usher in a new era of biotechnology where synthetic biology is routinely applied to these organisms. One such complex phenotype is cellulose utilization at low pH. Even in the absence of organic acid fermentation pathways, the CO_2 produced during fermentation requires the use of pH control to maintain neutral conditions. However, so far, only one bacterium, *Acidothermus cellulolyticus*, has been identified that is capable of growing on cellulose at or below pH 5 or lower. Therefore, we are enriching and isolating additional organisms capable of growth on crystalline cellulose at pH 5 and lower. Once identified, these new organisms will be explored for use in consolidated bioprocessing (CBP), the one-pot conversion of cellulose to product without the addition of biomass-degrading enzymes.

No known organisms have all the necessary phenotypes for biomass conversion to fuels and chemicals, so genetic manipulation will be needed to engineer them to have the desired traits. However, non-model organisms are typically unable to be bioengineered due to a lack of available genetic tools and an insufficient foundation of knowledge about the organism. The development of genetic tools is limited largely by the inability to efficiently transform DNA into these organisms. A critical barrier to transformation is DNA restriction-modification systems, which act as a bacterial immune system to cut DNA that is methylated differently than the host. Typically, these systems are comprised of methylation and restriction subunits. To prevent host

death, the cognate DNA methyltransferases recognize the same target sequence as the restriction enzymes, and the methylated DNA is protected from restriction. Therefore, in order to prevent restriction of heterologous DNA, the DNA needs to be methylated in the same manner as the host organism prior to transformation. In order to determine the sites targeted for restriction, we performed methylome analysis on 22 metabolically and phylogenetically diverse organisms of interest in collaboration with the Department of Energy Joint Genome Institute. This information was used to choose methyltransferases for expression in *E. coli* to protect DNA for 16 of these organisms, which has led to improved or first demonstration of genetic transformation in 8 of these organisms thus far. This approach is also being applied to *A. cellulolyticus* and will applied to new isolates that utilize cellulose at low pH. Overall, this work is leading to the development of new bioengineering platform organisms for use in the production of renewable fuels and chemicals, and similar approaches are being taken with other organisms of interest to demonstrate that this is a broadly applicable approach to developing new host organisms for advanced bioengineering.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Exploiting Nature's Knockouts: How the Pangenome Reveals New Insights into Poplar Phenotypic Variation

Doug Hyatt[1], David Kainer[1]* (kainerd@ornl.gov), Piet Jones[2], Manesh Shah[1], Jared Streich[1], Wellington Muchero[1], Jay Chen[1], Tim Tschaplinski[1], Posy Busby[3], Daniel Jacobson[1] and **Gerald A. Tuskan[1]**

[1]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; [2]University of Tennessee, Knoxville; [3]Oregon State University, Corvallis

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Most GWAS are performed using SNPs obtained from reads aligned to a reference genome of the same species, or a closely related species. However, it is becoming clear that a reference genome assembled from one individual may not adequately represent the genomic feature space of the species. A large number of genes and other genomic elements may be present or absent in any given individual, including the original reference individual. The overall species-wide genetic space is known as the Pangenome, containing core genes found in most or all individuals, and distributed genes that are found in fewer individuals and may be rare. Distributed genes that are absent in the original reference individual (e.g. Nisqually-1) may explain some of the 'missing' heritability of important quantitative traits, provide new sources of introgression material for breeding, and new explanatory power for understanding population structure and the evolution of the species. We can view the presence and absence of these genes within a population as a naturally evolved assay of knockouts.

We have assembled the pangenome of *Populus trichocarpa* using DNA resequencing and RNAseq data from almost 1000 genotypes, resulting in a new marker dataset representing the presence or absence variation of putative distributed genes across the population. Here we present results of association between these markers and phenotypes assayed in the same population.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Aboveground Effects of Beneficial Microbe Co-Culture with *Populus* Cell Wall Chemistry Variants

Udaya C. Kalluri* (kalluriudayc@ornl.gov)[1], Anna Matthiadis[1], Mindy Clark[1], Sara Jawdy[1], Lee Gunter[1], Wellington Muchero[1], Dale Pelletier[2], Jessy Labbe[1], Timothy Tschaplinski[1], and **Gerald A. Tuskan[1]**

[1]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; [2]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Beneficial microbes associated with bioenergy feedstock crops are the subject of immense recent interest from research as well as industrial sectors for their potential in improving bioenergy crop productivity and robustness. The goal of present research is to understand the effect of *Populus* cell wall chemistry variation on interactions with beneficial microbes. Previous studies have shown that modifications of lignin and cellulose content can have quantifiable effects on associated microbiome in *Populus* (Beckers et al. 2016; Kalluri eat al, 2016; Veach et al. 2018). In the present study, we tested two hypotheses on beneficial plant-microbe interactions: first, plant cell wall phenotype (e.g. high/low lignin) alone is a predictor of functional interactions with microbes, and second, microbes with a known root proliferative effect have a positive impact on aboveground biomass production.

Towards this end, effects of single and mixed microbe cultures on plant growth, stem and root biomass and chemistry were studied in a 15-week long greenhouse experiment. The study included sixteen distinct genotypes of *Populus trichocarpa* (representing extreme % lignin and % C6 sugar phenotypic variants) and bacterial and fungal isolates including *Pantoea*, *Variovorax*, *Paraburkholderia*, *Pseudomonas*, *Mortierella*, and *Laccaria*. These microbes have previously been shown to have root profilerative effects on *Populus* in short-term co-culture experiments (Labbe et al. 2014; Timm et al. 2016; Bible et al. 2016). Leaf, stem and root samples were harvested from six biological replicates for dry biomass, cell wall composition, metabolome, microbiome and transcriptome analyses. Our analyses of aboveground properties, including changes in plant height, stem diameter and leaf chlorophyll changes in plants, with or

without co-culture, show that there was no uniform pattern of response across a given phenotype class; rather a significant genotype-specific effect was observed. While a significant gain in stem biomass was observed as a result of co-culture with *Variovorax*, *Paraburkholderia*, *Mortierella*, and *Laccaria*, a significant overall negative impact on above-ground growth with *Pantoea* was observed across most *Populus* genotypes. Contrary to a theoretical expectation that co-culture with root-growth promoting microbes can have a beneficial aboveground growth effect, several genotypes displayed contrasting stem growth responses to the same microbial treatments. Ongoing analyses are expected to shed light onto the basis of the strong genotype-specific response, and clarify the interconnections among plant genotype, chemistry, above- and belowground- biomass responses to microbes. These insights and approaches will be useful for identifying genetic underpinnings of favorable plant responses, identifying microbes with broadly favorable aboveground effects and for informed design of non-transgenic approaches to maximizing bioenergy crop productivity.

References

1. Beckers et al. Proc Natl Acad Sci U S A. 2016. doi: 10.1073/pnas.1523264113.
2. Kalluri et al. Frontiers in Plant Sciences. 2016. doi: 10.3389/fpls.2016.01455.
3. Veach AM et al. Plant and Soil. 2018. 429: 349–361.
4. Timm et al. 2016. Frontiers in Plant Sciences. doi.org/10.3389/fpls.2016.00497
5. Bible et al. 2016.Frontiers in Microbiology 2016: doi:10.3389/fmicb.2016.00491.
6. Labbe et al. 2014. Frontiers in Plant Sciences, Plant-Microbes interactions 5(579):1-10.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Glycosylation is Vital for Industrial Performance of Hyper-Active Cellulases

Brandon C. Knott[1*] (brandon.knott@nrel.gov), Daehwan Chung[1], Nicholas S. Sarai[1], Neal Hengge[1], Jordan F. Russell[3], Roman Brunecky[1], Nitin Supekar,[3] Todd VanderWall[1], Deanne W. Sammond[1], Michael F. Crowley[1], Parastoo Azadi[3], Janet Westpheling[3], Michael E. Himmel[1], Yannick J. Bomble[1], and **Gerald A. Tuskan**[2]

[1]National Renewable Energy Laboratory Golden, CO; [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; [3]University of Georgia, Athens

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

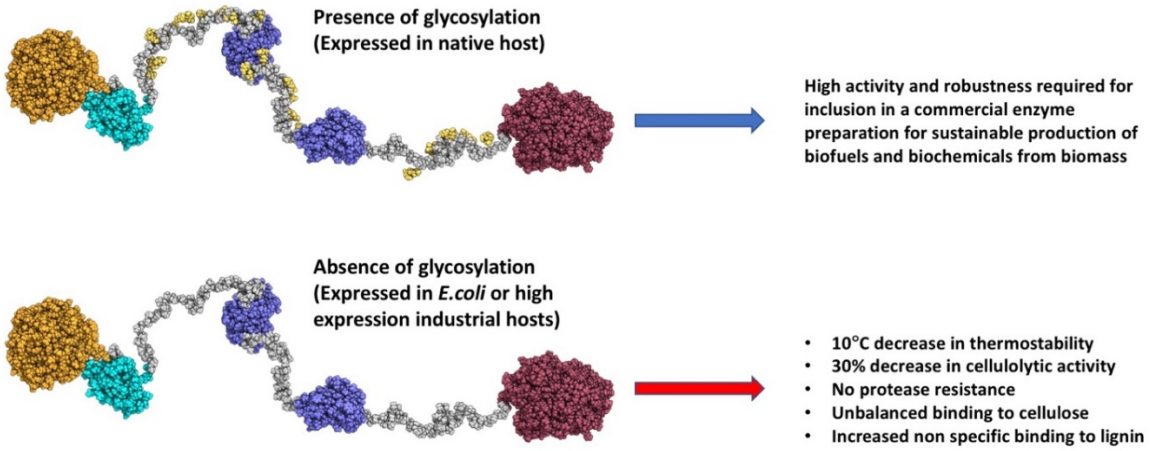
In the terrestrial biosphere, biomass deconstruction is conducted by microbes employing a variety of complementary strategies, many of which remain to be discovered. These strategies are also employed industrially, where the development of cost-effective biofuels requires increasingly more efficient (and less costly) cellulase formulations. The glycan decoration of fungal cellulases has been shown to protect these enzymes from protease action and to enhance binding to cellulose. Protein glycosylation is one of the most common protein post-translational modifications (PTM) and is thought to permit microorganisms to expand the combinatorial complexity of their gene products at a level beyond sequence space alone, opening new routes to structural, catalytic, and thermodynamic diversity. Glycosylation as a PTM in Nature has a variety of proposed roles that include enhancing protein solubility, biasing protein folding pathways, providing stability against proteolysis, and modulating signaling and molecular recognition pathways. Eukaryotic cellulolytic enzymes are often glycosylated, and this PTM has been shown to be important for both function and stability. Here we examine glycosylation's importance in cellulolytic procaryotes.

The hyperthermophilic anaerobic bacterium, *Caldicellulosiruptor bescii,* isolated from hot springs in the geothermally active Valley of Geysers in Siberia, can efficiently solubilize biomass without pretreatment. *C. bescii* relies primarily on a suite of complex multi-catalytic domain and multifunctional gene products to deconstruct biomass. The secretome of *C. bescii* displays high cellulolytic activity and the combination of the four most highly expressed enzymes in the secretome are enough to reproduce the activity of the entire secretome, making these enzymes appealing for the biofuels industry. One of these enzymes, CelA, has been shown to be the most efficient single gene product on several biomass substrates. CelA is a complex, thermally stable cellulase, containing an *N*-terminal GH9A-$CBM3_c$ processive endoglucanase, two family 3 carbohydrate-binding modules ($CBM3_b$), and a *C*-terminal GH48 exo-β-1,4-glucanase domain linked by Pro/Thr rich linkers. Recent

characterization of *C. bescii* carbohydrate-active enzymes (CAZymes), and especially native CelA has shown that the enzyme is glycosylated upon secretion from the cell. Despite the recognition of its important role in cellulolytic eukaryotes, detailed studies of bacterial cellulase glycosylation have not been reported.

We demonstrate that glycosylation of multimodular bacterial cellulase CelA is uniform across its three linker peptides and composed primarily of galactose disaccharides. This pattern of glycosylation is unique among previous studies of eukaryotic and bacterial cellulases. Furthermore, by combining experiment and computation, we find that glycosylation on the three intrinsically disordered regions of CelA plays key roles in modulating its proteolytic stability, thermodynamic stability, substrate binding, hydrolytic activity, and overall tertiary structure. The collective effects of glycosylation provide this multifunctional cellulase with the ability to function optimally in harsh environments, including geothermal hot springs and industrial biorefineries.

The understanding of bacterial cellulase glycosylation developed by this study opens the door to further studies of glycosylation in these systems but also to applications in the biofuels and biomaterials industries where better biomass-degrading enzymes are needed. For example, nearly all industrial applications of secreted enzymes require access to hosts capable of large-scale production yielding high titers of proteins. Challenges with the expression of bacterial glycosyl hydrolases, many of which are highly active or display diverse and important specificities, continue to limit the introduction of these enzymes into commercial markets. Enhancing our understanding of the roles played by glycans decorating bacterial enzymes will greatly enable use of these diverse catalysts at large scale to help enable the sustainable production of biofuels and biochemicals.



The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Quantitative Network Analysis in *Clostridium thermocellum* using ^{13}C and ^{2}H-tracers

Travis Korosh[1]* (korosh@wisc.edu), Saratram Gopalakrishnan[3], Tyler Jacobson[2], Satyakam Dash[3], Charles Foster[3], Daniel G. Olson[4], Lee R. Lynd[4], Costas Maranas[3], Daniel Amador-Noguez[2], and **Gerald A. Tuskan[1]**

[1]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; [2]University of Wisconsin-Madison, Madison; [3]Pennsylvania State University, State College; [4]Dartmouth College, Hanover

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Recent advances in genetic engineering have enabled production of a small suite of specialty biofuels at a commercial scale using industrially relevant organisms such as *Saccharomyces cerevisiae* and *Escherichia coli*. However, industrial biotechnology has been hampered by low titers, yields, and productivities (*1*). Product toxicity has long been cited to be one of the main limitations to achieving high product titers and has often been linked to physiochemical effects on the cell membrane (*2–4*). However, thermodynamic constraints can also limit the metabolic potential of a given host to produce a particular compound (*5*) by affecting the rate of a biochemical reaction or pathway through the flux-force relationship (*6*).

Thermophilic, cellulolytic anaerobic bacteria such as *Clostridium thermocellum* have a pyrophosphate dependent metabolism in close proximity to thermodynamic equilibrium (*7*), which is thought to help with ATP yield over traditional glycolysis (*8*). This energetic efficiency may be causing the current ethanol titer limit of ≈ 30 g/L produced by metabolically engineered strains of *C. thermocellum*, which is in stark contrast to the 70 g/L titer achieved by metabolically engineered strains of the hemicellulose-fermenting *Thermoanaerobacterium saccharolyticum* (*9*).

In order to test the relationship of the thermodynamic driving force to product titers, we used metabolic flux analysis in conjunction with flux ratio analysis to compare the Gibbs free energy values and fluxes of wild-type *C. thermocellum* and *T. saccharolyticum* to metabolically

engineered strains in the presence of isotopically labelled ethanol. In the present study we identify GAPDH, TPI, and ADH as close to thermodynamic equilibrium in wild-type *C. thermocellum*, compare the effects of alternative pyruvate oxidation pathways towards increasing the thermodynamic driving force and altering fluxes of fermentation pathways, and suggest further engineering strategies to overcome metabolic bottlenecks.

References

1. G.-Q. Chen, *Microb. Cell Fact.* 11, 111 (2012).
2. B. Wilbanks, C. T. Trinh, *Biotechnol. Biofuels*. 10, 1–11 (2017).
3. L. A. Royce, P. Liu, M. J. Stebbins, B. C. Hanson, L. R. Jarboe, *Appl. Microbiol. Biotechnol.* 97, 8317–8327 (2013).
4. R. M. Lennen *et al.*, *Appl. Environ. Microbiol.* 77, 8114–8128 (2011).
5. J. Asplund-Samuelsson, M. Janasch, E. P. Hudson, *Metab. Eng.* 45, 223–236 (2018).
6. D. A. Beard, H. Qian, *PLoS One*. 2, 1–4 (2007).
7. J. Zhou *et al.*, *Appl. Environ. Microbiol.* 79, 3000–3008 (2013).
8. E. Mertens, *Parasitol. Today*. 9, 122–126 (1993).
9. S. Hon *et al.*, *Biotechnol. Biofuels*. 11, 242 (2018).

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Investigating the Role of Beneficial Microbes Reducing the Rust Infection Severity in *Populus*

Jessy Labbé[1]* (labbejj@ornl.gov), Mang Chang[2], Timothy Tschaplinski[1], and **Gerald A. Tuskan[1]**

[1]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; and [2] University of Tennessee, Knoxville

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Populus is an important commercial crop and one of the leading candidates for bioenergy production; however, the commercial use of *Populus* is easily impacted by abiotic and biotic stresses. These stresses include major pathogens such as rust fungi, *Melampsora*. In order to maximize sustainable *Populus* biomass production, we proposed to engineer microbial community to alleviate stresses and nutrient deficiency limiting growth and productivity using a native microbial collection. We assembled a fungal-bacterial community that benefit each other (helper effect, Labbé et al., 2014) and are symbionts of *Populus*. At field scale, we demonstrated these beneficial microbes are stable in soil and *Populus* roots (after two years) while they increase plant yield of 30%. We also observed that this microbial treatment significantly reduces the severity of infection of rust pathogens (*Melampsora spp*.). As a consequence of this result, we further investigate the effect of the beneficial microbes during *Melampsora* infection in *Populus*. Yet there is little known on the role of beneficial root microbes in the modulation of the poplar immune system to decrease susceptibility to leaf pathogens. Then, we employ closed microcosm system with beneficial microbial inoculant and sterile *Populus* plant treated or not with the rust pathogen. Thus, coupling transcriptomic and metabolomic approaches, we aim to characterize the defense-related metabolites and genes primed in poplar by beneficial microbes and to identify specific-defense microbial metabolite produced or transferred to the plant partner. This pioneer work contributes to the understanding of the molecular mechanisms conferring protection against pathogens in woody plants during beneficial associations to further exploit them and develop a sustainable bioenergy feedstock.

References

1. Labbé, J. L., Weston, D. J., Dunkirk, N., Pelletier, D. A., & Tuskan, G. A. (2014). Newly identified helper bacteria stimulate ectomycorrhizal formation in *Populus*. Frontiers in plant science, 5, 579.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Recombination and Segregation Genomic Patterns in *Populus* to Facilitate Genomic Selection

Chanaka Roshan Abeyratne[1], David Macaya-Sanz[1]* (david.macayasanz@mail.wvu.edu), Gancho Slavov[2], Lee Gunter[3], Kathleen Haiby[4], Richard Shuren[4], Jay Chen[3], Daniel Jacobson[3], Brian Stanton[4], Stephen DiFazio[1], and **Gerald A. Tuskan[3]**

[1]West Virginia University, Morgantown; [2]Rothamsted Research, Harpenden, UK; [3]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; and [4]GreenWood Resources Inc., Portland, OR

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Black cottonwood (*Populus trichocarpa*) is a pioneer tree species identified as a promising renewable feedstock for bioenergy and bioproducts. *Populus* can be used to study many aspects of perennial development related to phenology, wood formation, vegetative propagation, and dioecy that cannot be studied using conventional plant model systems such as *Arabidopsis*. Breeding techniques to identify elite parent trees face considerable challenges: lengthy sexual reproduction cycles (of at least 4 years), difficulty to design Recombinant Inbred Lines (RIL), highly heterozygous genomes, and poor juvenile-mature correlation for many traits, necessitating expensive long-term field trials. Genomic selection incorporates genomic information into breeding strategies to alleviate these issues.

Prediction models can be improved by detailed understanding of how recombination and segregation vary across the genome, among individuals and between sexes. Linkage information also allows the creation of genetic maps that can be used to identify Quantitative Trait Loci (QTL) for phenotypes of breeding relevance (height, diameter, bud set, and disease resistance). Those QTL can also be used to improve prediction model training.

Here, we have re-sequenced the genomes of 49 families ($N = 821$ offspring), corresponding to a full factorial cross of seven females and seven males. Using benchmark software GATK4, we have called biallelic SNPs. Using the pedigree information, we have revealed patterns of segregation in the genome and selected high-quality markers under Mendelian segregation. We have phased and imputed the progeny genomes and recovered the gametic haplotypes of the

fourteen parents, allowing us to estimate fine-scale genomic patterns of recombination. By means of the R package Onemap, fourteen genetic maps have been created, and specific patterns of variation have been revealed.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Development and Optimization of Unmanned Aerial Vehicle High-Throughput Phenotyping of Field-Grown Switchgrass

Reginald J. Millwood[1]* (rmillwood@utk.edu), Mitra Mazarei[1], Francisco J. Palacios[1], Chandler D. Douglas[1], C. Neal Stewart Jr.[1], and **Gerald A. Tuskan[2]**

[1]University of Tennessee, Knoxville; [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Unmanned aerial vehicle (UAV) sensor-based analysis is an emerging and powerful platform for high-throughput plant phenotyping under field conditions. In this study, we explored the potential of UAV-based sensing to measure plant height and area as well as chlorophyll content of field-grown switchgrass. A pilot field experiment was performed in Knoxville, Tennessee from late August to December 2018. A total of 120 lowland (tetraploid) switchgrass GWAS accessions were transplanted in four plots with plant spacing of 1.5 m, 1.75 m, 2.0 m, and 2.25 m. Included within each plot were height reference points of 0 m, 0.5 m, 1.0 m, 1.5 m, and 2.0 m. Plant height and area were determined from UAV flights equipped with a high resolution camera mounted to a 3-axis gimbal. Flights were performed at the end of the growing season while manual measurements for plant height and area were collected for comparison. We found UAV-based plant height measurement correlated strongly with manual measurements (r = 0.92), and the same was observed for plant area measurements (r = 0.94). In addition, UAV flights equipped with a multispectral camera were performed during the growing season to estimate chlorophyll content. Multispectral images were analyzed through the use of established vegetation indices, and our findings indicated UAV-derived chlorophyll content estimates can be made in switchgrass. However, these data have yet to be “ground-proofed”. Owing to a strong positive correlative relationship with chlorophyll, lab-based leaf nitrogen analysis will be performed for confirmation. The UAV-based methods developed here will be incorporated in a large-scale switchgrass GWAS analysis experiment to identify loci associated with high biomass yield and nitrogen use efficiency with subsequent integration into genomic selection models.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Manipulation of Seasonal Dormancy as a Strategy to Improve Switchgrass Biomass Yield

Rasyidah Razar[1], Ali M. Missaoui[1*] (cssamm@uga.edu), Katrien Devos[1], and **Gerald A. Tuskan[2]**

[1]University of Georgia, Athens; [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic approaches to increase the sustainability of biomass production to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

To establish a sustainable biomass energy system, improving biomass yield and the seasonal distribution of production is crucial. The short growing cycle of switchgrass (*Panicum virgatum* L) cultivars limits not only the amount of biomass accumulated but also the seasonal distribution, which is restricted to one harvest per year at the end of summer. Like many warm season species, switchgrass evolved winter dormancy as an adaptation strategy to survive cold winters, especially in the most northern latitudes. Because of winter dormancy, switchgrass growth and biomass accumulation are repressed substantially during the short-photoperiod months even in southern environments where winter temperatures are mild. Reducing winter dormancy might be a viable strategy to increase yield in southern environments with mild winters. The overall objective of this research is to investigate the effects of manipulating winter dormancy in switchgrass, and to identify the genetic components (QTL and candidate genes) underlying winter dormancy and related traits such as flowering time and spring regrowth in a F1 pseudo-testcross from a non-dormant parent (B6) with a dormant parent (AP13). Genotyping-by-sequencing generated 2.5 million raw variants of which 3761 were single-dose (SDA) maternal alleles and 4133 were paternal SDAs. Linkage maps constructed with Mendelian segregating and distorted SDAs with a p-value $\geq E^{-15}$ resulted in 18 linkage groups representing the nine K and nine N sub-genome chromosomes. A total of 3555 markers were mapped in the maternal map (2238 cM) and 3711 markers in the paternal map (1826 cM). QTL identification for one environment for fall regrowth height, NDVI, and spring emergence date showed several promising chromosome regions associated with the traits. Field data will continue to be collected and QTL analysis across years will be conducted. The substantial genetic variation in the onset and duration of dormancy shows potential for breeding of non-dormant or semi-dormant cultivars for the southern regions where winters are mild.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Development of Consortia of Beneficial Microbes for Switchgrass and Consequences of Their Deployment on Native Soil Microbiome

Prasun Ray[1]* (pray@noble.org), Myoung-Hwan Chi[1], Yingqing Guo[1], Kelly D. Craven[1] and **Gerald A. Tuskan[2]**

[1]Noble Research Institute, LLC, Ardmore, OK; [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

The family *Serendipitaceae* represents a diverse group of fungi in the Basidiomycota that encompasses endophytes and lineages that repeatedly evolved ericoid, orchid and ectomycorrhizal abilities. *Serendipitaceae* species provide excellent models for root endophytism, given the availability of sequenced genomes, genetic tractability, and a broad host plant range. Previous research performed in our lab with an Australian strain of *Serendipitaceae* named *Serendipita vermifera* indicates its plant growth promoting abilities in switchgrass in normal or stressed conditions. Considering their proven beneficial impact on plant growth and their ubiquity, we describe *Serendipitaceae* as an effective microbial tool for enhancing switchgrass productivity and stress tolerance. The agronomic utility of these fungi has been hampered by the paucity of strains available, the large majority isolated from Australian orchids. We have addressed this constraint by isolating the first North American strain of *Serendipita*, named *Serendipita vermifera* subsp. *bescii* NFPB0129, from switchgrass root in Ardmore, Oklahoma.

In parallel, more than 500 bacterial strains that were previously isolated from switchgrass roots, from the Tall Grass Prairie Preserve, northern Oklahoma have been screened for plant growth promoting traits such as phosphate solubilizing activity, ACC deaminase activity and nitrogen fixation ability. We found 31 strains with phosphate-solubilizing activity, 39 strains with ACC deaminase activity and 42 strains that harbor the *NifH* gene. These strains were subsequently evaluated for their switchgrass growth promoting ability. The top three performers across these three different categories are selected to make a bacterial consortium. The consortia itself as well as individual members of this consortia were evaluated for switchgrass growth promotion under

greenhouse conditions. Ultimately, we would pair this bacterial consortium with *Serendipita bescii*, the unique endomycorrhiza isolated from switchgrass and evaluate their growth promoting potential in two different field locations in Ardmore, Oklahoma.

Deployment of microbes with known plant growth promoting traits such as this has played a major role in improving the productivity of agronomically important crops for many decades. Around the world, experiments have been designed and conducted in laboratory set up, green house and field conditions to optimize and increase the efficiency of microbes or consortia of microbes to improve productivity of agronomically important crops. However little or no effort has been taken to measure impact of these introduced microbes on native soil or associated rhizosphere microbiome. We are now beginning to acknowledge the critical role played by the soil microbiome in shaping the plant's overall physiology and development. Hence, it is critical to put forth similar emphasis to understand the interactions between introduced microbes with the rhizosphere microbiome. Recent studies have suggested that such microbial cross talk are regulated, available soil nutrients, and defense hormones, among others. Therefore, care must be taken to ensure that synthetic communities are not so aggressive that they invade local ecosystems and negatively affect soil health, and eventually compromise plant productivity.

From this perspective, in our field trial we are going to determine not only the impact of bacterial endophyte and *Serendipita bescii* inoculation on switchgrass growth performance, but also its impact on native soil microbiome by 16rRNA and ITS2 amplicon metagenomics analysis to estimate diversity and relative proportion of bacterial and fungal communities. This will enable us to evaluate how these microbial inoculants interact with the native microbiome, and subsequently establish, compete, and function in agricultural soils for improving plant productivity.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Unraveling New Mechanisms for Lignin Catabolism in Nature

Davinia Salvachua[1*] (davinia.salvachua@nrel.gov), Allison Werner[1], Paul E. Abraham[2], Richard J. Giannone[2], Martyna Michalska[3], Sandra Notonier[1], Bryon S. Donohoe[1], Philip D. Laible[3], Robert L. Hettich[2], Gregg T. Beckham[1], and **Gerald A. Tuskan[2]**

[1]National Renewable Energy Laboratory Golden, CO; [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN; [3]Argonne National Laboratory, Chicago, IL

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

A primary CBI goal is to understand how lignin is converted by microbes to facilitate cost-effective, efficient lignin valorization. *Pseudomonas putida* KT2440 is a Gram-negative soil bacterium reported to efficiently catabolize a variety of aromatic compounds. We have also demonstrated that *P. putida* can utilize high molecular weight lignin [1]. The intracellular mechanisms implicated in the catabolism of aromatic compounds have been extensively studied, but the enzymes involved in the breakdown of oligomeric lignin and/or their spatial location remains unknown. The work presented here provides new insights into the location of aromatic-catabolic and ligninolytic enzymes during microbial conversion of lignin.

To identify the enzymes involved in lignin breakdown, we performed a differential proteomics study in the intracellular and extracellular fraction of *P. putida* when grown in lignin and minimal media. The number of proteins found exclusively in lignin was considerably higher in the extracellular fraction than in the intracellular one and the former fraction contained oxidoreductases and aromatic-catabolic enzymes that had been previously described to be intracellular. To discern between cell lysis and secretion of enzymes to the extracellular milieu, we conducted cytometry and microscopy analysis. Cytometry revealed that 5 and 12% of the population were dead cells in minimal and lignin media, respectively. However, scanning and transmission electron microscopy images also uncovered the presence of outer membrane vesicles (OMVs), both in lignin and minimal media.

With the aim of understanding the function of these vesicles, we first analyzed their cargo through proteomics analysis. For that purpose, we isolated the vesicles from *P. putida* cultures grown in minimal media and lignin. The results indicate that some proteins were enriched in the

OMV fraction compared to the supernatant. In fact, some of the enzymes enriched in OMVs from lignin cultures have been previously reported to be involved in the catabolism of aromatic compounds or correlated with lignin breakdown in other organisms. This discovery opens new directions for investigation, from fundamental research to biotechnological applications, to understand how bacteria interact with lignin or aromatic compounds in the extracellular locus and to engineer improved microbes for the conversion of lignin to renewable chemicals.

References

1. Salvachúa, D., Karp, EM, Vardon, D.R., Nimlos, C., Beckham, G.T.*, 2015. Towards lignin consolidated bioprocessing: simultaneous lignin depolymerization and product generation by bacteria. Green Chemistry, 17, 4951-4967.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

The GWAS Times: A Time-Series Aware GWAS to Detect Natural Climate Adaptations in *Arabidopsis* and *Populus*

Jared Streich[1*] (ju0@ornl.gov), David Kainer[1], Ashley Cliff[1], Jonathan Romero[1], Daniel Jacobson[1], and **Gerald A. Tuskan**[1]

[1]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

Elite crop lines are often bred for specific traits and trialed in farming environments which are synthetic compared to natural conditions. Maintenance of crops due to environmental effects can be costly and drive down profits. The Center for Biofuel Innovation aims to increase cellulosic biomass for biofuels and chemical feedstocks made from lignin residuals by use and improvement of two plant species, Switchgrass (*Panicum virgatum*) and Poplar trees (*Populus trichocarpa*). To improve genomic selection of potential breeding lines we created a time-series aware association approach called GWATS (Genome Wide Association Time-series Studies) to detect climate adaptive alleles and demonstrated them on *Populus trichocarpa*. GWATS involves running a GWAS analysis using time-series phenotype data, in this case for each day of the year. BioClim has been an ideal data series for species distribution modeling and finding adaptive alleles but lacks a true seasonal component and leaves much of each seasonal period unexamined. Instead, from raw monthly climate data, we interpolated 365 daily values of 78 climate/environment layers using 498 locations of 970 *Populus trichocarpa* individuals from where our GWAS population were sourced using ~10M genome-wide SNPs. Geographically isolated alleles tend to coincide with adaptive alleles making them difficult to distinguish from false positives. However, true positive p-values become distinct from false positives by using a Fourier Transform to analyze co-variate variation and output similarity matrix calculations across the time-series. Further, using a machine learning algorithm called iRF (iterative Random Forest) we can filter complex climate phenotypes into epistatic interactions across multiple suites of alleles ranging as high as five or more orders of epistasis. Using GWATS we detected hundreds of candidate climate adaptive loci for Solar Radiation Stress, Temperature Stress, Aridity Stress, Light Quality and many more. From this analysis we expect to create a list of climate-ideal lines to breed both climate wide and climate specific ideotypes in addition to being optimized for CBI-target phenotypes to improve agricultural production of *Populus trichocarpa*.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

CRISPR Genome Editing in *Populus*

Chung-Jui Tsai[1*] (cjtsai@uga.edu), Patrick Bewg[1], Liang-Jiao Xue[1], Jacob Reeves[1], Margot Chen[1], and **Gerald A. Tuskan[2]**

[1]University of Georgia, Athens; and [2]Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN

https://cbi.ornl.gov

Project Goals: The Center for Bioenergy Innovation (CBI) vision is *to accelerate domestication of bioenergy-relevant, non-model plants and microbes to enable high-impact innovations at multiple points in the bioenergy supply chain.* CBI will address strategic barriers to the current bioeconomy in the areas of: 1) high-yielding, robust feedstocks, 2) lower capital and processing costs via consolidated bioprocessing (CBP) to specialty biofuels, and 3) methods to create valuable byproducts from the lignin. CBI will identify and utilize key plant genes for growth, composition and sustainability phenotypes as a means of achieving lower feedstock costs, focusing on poplar and switchgrass. We will convert these feedstocks to specialty biofuels (C4 alcohols and C6 esters) using CBP at high rates, titers and yield in combination with cotreatment or pretreatment. CBI will maximize product value by *in planta* modifications and biological funneling of lignin to value-added chemicals.

The overarching goal of the Rapid Genetics team is to develop cutting-edge genetic toolkits to enable efficient and multiplexed genome editing for development of optimal feedstocks specifically Poplar. This requires the establishment of an efficient transformation system, the development and optimization of a robust CRISPR genome editing platform for knock-out (KO) and knock-in (KI) mutagenesis, and development of computational resources for genome editing in heterozygous plants. An improved Variant database (http://aspendb.uga.edu/s717) for *P. tremula* x *alba* INRA 717-1B4 has been deployed to facilitate variant-free guideRNA (gRNA) design. The database serves a global research community. By using a single gRNA, we are able to target individual, as well as duplicated genes, with 100% KO efficiencies for traits ranging from lignin biosynthesis to trichome formation. The use of one specific gRNA per target gene is preferred over multiple gRNAs per gene, especially for multiplex editing when several unrelated genes are targeted for KO simultaneously, as it simplifies vector design and reduces potential off-target effects. The lignin and trichome-less KO mutants are being exploited to develop a KI system based on MMEJ (microhomology-mediated end-joining) repair pathway, rather than the standard homology-directed repair, which requires extensive homology with poor efficiencies in eukaryotes. Successful repair of the previously generated KO alleles will restore gene functions, and in the case of trichome-less plants, will lead to visually identifiable phenotypes. This will alleviate lengthy molecular screening during various optimizations. The effects of homology length, template excision, and target gene choice on KI efficiency are currently being assessed. In collaboration with other CBI researchers, the KO system is being applied to target genes selected for improving growth, sustainability and bioprocessing characteristics of *Populus*.

The Center for Bioenergy Innovation is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Prospecting thiamine diphosphate-dependent carboligases and characterizing their promiscuity to create novel metabolic pathways from primary metabolites

Bradley W. Biggs[1]* (bradleybiggs2014@u.northwestern.edu), Jonathan Strutz,[1] Sara Wixon,[1] Matthew T. Robey,[2] Valerie J. Winton,[3,4] Neil L. Kelleher,[2,3,4] Paul M. Thomas,[4] **Keith E.J. Tyo**[1]

[1]Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL; [2]Department of Molecular Biosciences, Northwestern University, Evanston, IL; [3]Department of Chemistry, Northwestern University, Evanston, IL; and [4]Proteomics Center of Excellence, Northwestern University, Evanston, IL

https://pamspublic.science.energy.gov/CCBond

Project Goals: The goal of this project is to characterize a library (>100) of thiamine-diphosphate dependent carboligase enzymes against a diversity of α-ketoacid substrates to determine the reaction landscape of this family of enzymes using machine learning, to identify ideals candidate enzymes from this family for biosynthesis applications, and then to use this information to assembly favorable enzymatic pathways to target bioproducts.

Abstract. Recent work has shown that enzyme promiscuity, the ability of an enzyme to accept non-native substrates and perform non-native chemistries, is widespread in nature. This provides an opportunity for biological engineers to leverage this capacity for valuable unnatural transformations and to hone desired activities. One particularly interesting family of enzymes to this end is thiamine-diphosphate dependent carboligases, which condense two α-ketoacids (or aldehydes) to form new carbon-carbon bonds. Because of an abundance of α-ketoacids in the central metabolism of common metabolic engineering hosts like *Escherichia coli*, this provides opportunity to assemble new favorable biochemical pathways to targets of interest. One-step condensations could allow for more efficient routes to desired targets and access to novel molecules, including chiral compounds. Our goal is to characterize a library (>100) of carboligases, map their reactivity on a diversity of α-ketoacid substrates using machine learning, and then utilize promising enzyme candidates for biosynthesis applications. Here, we demonstrate our characterization workflow with proof-of-concept enzymes and reactions.

This work is supported by DOE grant DE-SC0019339.

Tools for Easy, Fast and Accurate Quantitative Characterization of the Methanotroph-Photoautotroph Coculture

Kiumars Badr[1*] (kzb0054@auburn.edu), Kyle A. Stone[1], Q. Peter He[1], Alexander S Beliaev[2], Marina Kalyuzhnaya[3] and **Jin Wang**[1]

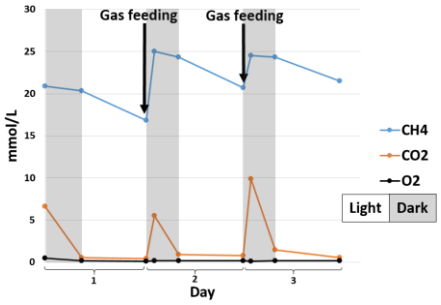
[1]Auburn University, Auburn, AL; [2]Pacific Northwest National Laboratory, Richland, WA; [3]San Diego State University, San Diego, CA

Project Goals: In nature, microbial communities have developed a highly efficient way to recover the energy and capture carbon from both CH_4 and CO_2 through interspecies coupling of methane oxidation to oxygenic photosynthesis. However, in order to successfully utilize mixed culture for biotechnology applications, both fundamental knowledge and technological gaps have to be addressed. The knowledge gap refers to the lack of systematic study for identifying and quantifying the interactions between community members and how the interaction feedbacks affect system dynamics. The technological gap refers to the lack of effective methodology, and fast and low-cost analytical tools to characterize mixed culture systems frequently or in real-time. The overall research objective of this project is to help address those gaps through developing experimental/computational tools to characterize a synthetic photoautotroph-methanotroph binary consortium, and to identify and validate interspecies interactions at both systems and cellular levels for a model methanotroph-photoautotroph coculture pair.

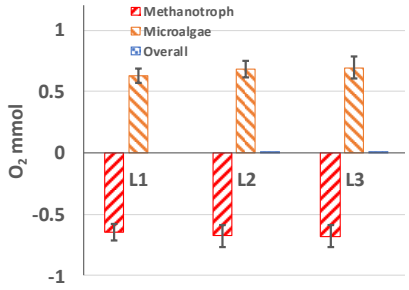
Abstract: Multispecies associations are ubiquitous in nature as they provide key ecosystem services such as carbon, nutrient, and metal cycling. It has been recognized that a mixed culture could offer a number of advantages over a conventional single-culture, such as complete utilization of substrate, better stability and robustness, higher product yield, higher growth rate, as well as the capability to carry out multistep transformation that would be impossible for a single organism. Despite these potential significant advantages, utilization of mixed cultures for biotechnological applications in bioenergy and related areas have been limited, partially due to the lack of effective tools to characterize the mixed culture accurately and frequently. In this project, we have developed experimental and computational protocols to quantitatively characterize the photoautotroph-methanotroph coculture. Specifically, we have developed experimental protocols to obtain accurate measurements of overall consumption and production rates for gas components CH_4, O_2 and CO_2[1]. We have also developed computational procedures to estimate individual gas consumption and production rates by each organism [2]. Such quantitative characterizations laid the foundation for the proposed modelling work.

Accurate measurement of overall consumption and production rate for CH_4, O_2 and CO_2

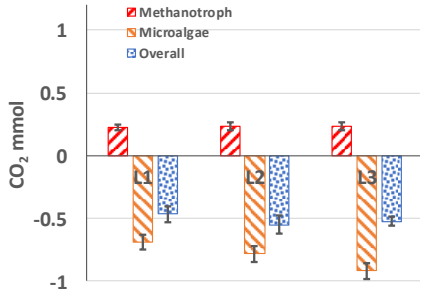
Due to gas phase volume/pressure (for batch experiments) or flow rate (for continuous experiments) change, and the pH dependent solubility of CO_2, using the direct GC measurements of the headspace or off-gas composition to calculate the gas consumption/production rates can cause large errors. To address this challenge, we have developed two easy-to-implement experimental protocols and associated calculation procedures to obtain accurate measurements of gas component consumption and production rates, one for batch operation and one for continuous operation. For batch operations, we use nitrogen (or other inert gases) to re-pressurize the system to atmosphere pressure before taking samples; while for continuous operations, we use helium (or other inert gases) as an internal tracer to accurately measure off-gas flow rate. In addition, we use total inorganic carbon (TIC) to track dissolved CO_2. The effectiveness and accuracy of the two



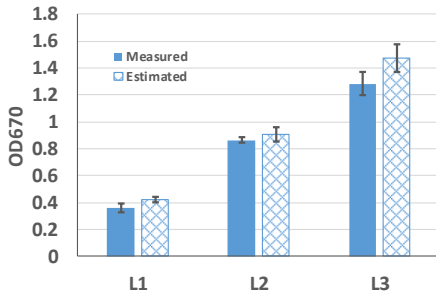
(a) Gas phase measurements



(b) O_2 consumption/production



(c) CO_2 consumption/production



(d) Total biomass

Fig. 1. Gas phase measurement over time (a), individual consumption/production (estimated) and overall change (measured) of O_2 (b) and CO_2 (c) for photoautotroph-methanotroph coculture during three light cycles. The results are validated by the good agreement between the estimated and measured total biomass during the same cycles (d).

protocols and associated calculation procedures are demonstrated using several case studies with both abiotic and biotic systems (both methanotroph and the coculture) [1, 2].

Estimate individual gas consumption and production rates by each organism Estimating individual consumption/production rate of CH_4, O_2, CO_2 and other metabolites in co-culture is challenging, but is absolutely necessary for understanding the dynamics of the system. The main challenges are: 1. individual consumption/ production rates of CO_2 and O_2 cannot be measured directly because of the coupling between photoautotroph and methanotroph; (2) the amount of dissolved CO_2 in the liquid medium must be estimated to complete the carbon balance. To address these challenges, we have developed a computational procedure based on mass balances and growth stoichiometric information (such as biomass yield) to compute the amount of O_2 and CO_2 consumed or produced by each organism [2]. One example is shown in Fig. 1. The accuracy of the developed protocol was validated by the good agreement between the estimated and measured total biomass.

Publications:

1. Stone K., He Q.P., & Wang J. (2018), Two Experimental Protocols for Accurate Measurement of Gas Component Uptake and Production Rates in Bioconversion Processes, Nature Scientific Report, under review.

2. Bahr K., Hilliard M., Roberts N., He Q.P. and Wang J. (2019), Photoautotroph-Methanotroph Coculture – A Flexible Platform for Efficient Biological CO_2-CH_4 Co-utilization, Proceedings of 12th Dynamics and Control of Process Systems (DYCOPS 2019), Apr. 23-26, 2019, Florianópolis, BRAZIL, accepted;

Funding statement: This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under Award Number DE-SC0019181; and by the U.S. Department of Education, Graduate Assistance in Areas of National Need Program under Award Number P200A150074.

Methanotroph-Photoautotroph Coculture – A Flexible Platform for Efficient Biological CO_2-CH_4 Co-utilization

Kiumars Badr[1], Matthew Hilliard[1*] (mvh0006@auburn.edu), Q. Peter He[1], Alexander S Beliaev[2], Marina Kalyuzhnaya[3] and **Jin Wang**[1]

[1]Auburn University, Auburn, AL; [2]Pacific Northwest National Laboratory, Richland, WA; [3]San Diego State University, San Diego, CA

Project Goals: In nature, microbial communities have developed a highly efficient way to recover the energy and capture carbon from both CH_4 and CO_2 through interspecies coupling of methane oxidation to oxygenic photosynthesis. However, in order to successfully utilize mixed culture for biotechnology applications, both fundamental knowledge and technological gaps have to be addressed. The knowledge gap refers to the lack of systematic study for identifying and quantifying the interactions between community members and how the interaction feedbacks affect system dynamics. The technological gap refers to the lack of effective methodology, and fast and low-cost analytical tools to characterize mixed culture systems frequently or in real-time. The overall research objective of this project is to help address those gaps through developing experimental/computational tools to characterize a synthetic photoautotroph-methanotroph binary consortium, and to identify and validate interspecies interactions at both systems and cellular levels for a model methanotroph-photoautotroph coculture pair.

Abstract: Industrial, municipal, and agricultural waste streams contain stranded organic carbon, which can be converted into biogas through anaerobic digestion. It has been demonstrated that biogas has immense potential as a renewable feedstock for producing high-density fuels and commodity chemicals. However, the utilization of biogas presents a significant challenge due to its low pressure, high proportion of CO_2 and presence of contaminants such as H_2S, ammonia, and volatile organic carbon compounds. To tap into this immense potential, effective biotechnologies that co-utilize both CO_2 and CH_4 are needed. Using the basic metabolic coupling principles utilized by many natural consortia, we have demonstrated that photoautotroph- methanotroph co-cultures offers a flexible and highly promising platform for biological CO_2/CH_4 co-utilization. In this work, we focused on filling part of the knowledge gap, i.e., identifying the potential interactions at the systems level through a hypothesis-driven approach. Also, we present the very first effort to quantitatively model the growth dynamics of a photoautotroph-methanotroph coculture. These advancements will enable us to further examine the interspecies interactions within the coculture at molecular level.

Advantages of the proposed coculture platform Coupling methanotrophic metabolism to photosynthesis offers several advantages for engineering biogas conversion. That includes the exchange of *in situ* produced O_2 and CO_2 which could dramatically reduce mass transfer resistance of the two gas substrates. However, one very important question should be answered first: are there clear benefits of using the coculture than using single cultures sequentially for biogas conversion? This is a critical question applicable to any consortia-based biotechnologies, as the operation of the mixed culture can be more challenging than maintaining two single cultures sequentially. To answer the question, we have designed and conducted the comparison experiments for three cases, using *Methylotuvimicrobium buryatense* (methanotroph) - *Arthrosipira platensis* (cyanobacterium) as the model coculture system. Case A is the coculture; Case B is the sequential culture of cyanobacterium followed by methanotroph; Case C simulates the effect of the exchange of *in situ* produced O_2 between the coculture, where the amount of O_2 produced by the

cyanobacterium in the coculture was injected into the methanotroph single culture. Our initial experimental results[1] clearly showed that the methanotroph-photoautotroph coculture offers significantly more benefit than sequentially operated single cultures. In addition, the model parameters of unstructured kinetic model, which were estimated using experimental data, further confirm that the synergy within the coculture is beyond the exchange of in situ produced O_2 and CO_2.

Kinetic Modeling of the Coculture The development of multi-organism platforms for commercial biogas conversion presents significant challenges which center around our ability to control function and composition of species in the coculture. An essential tool for the optimization, design and analysis of the coculture based biogas conversion is the development and validation of kinetic models that can accurately describe and predict the co-culture growth under different conditions. In this work, we have developed an unstructured dynamic model to capture the dynamics of the coculture growth under a variety of culture conditions. By considering the substrate exchange and self-shading effect, the unstructured kinetic model is able to capture the coculture growth dynamics accurately throughout the entire batch duration and can accurately predict the growth behaviour under a different light intensity [1], as shown in Figure 1A. Finally, both the model prediction and experimental results validated our hypothesis that the inoculum ratio of the two strains will not affect the final steady state ratio, as shown in Figure 1B.

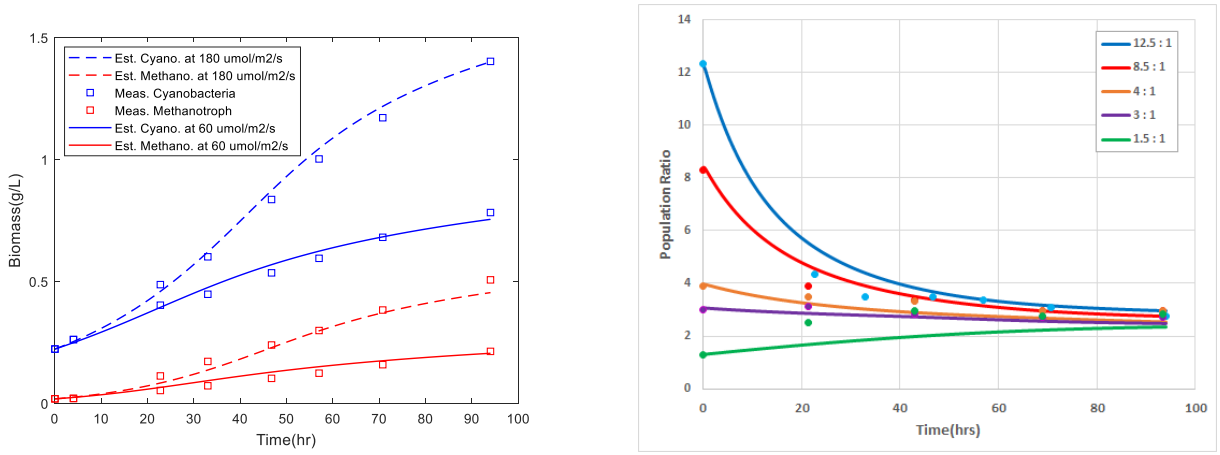


Fig. 1 A (Left). Comparison of the biomass concentration in the coculture predicted by the model (lines and dashed lines) with measurements (squares). B (Right). Both model simulation (lines) and experimental measurements (dots) confirm that despite different inoculum ratios (cyanobacterium : methanotroph), after reaching steady state, the population ratio of the two strains converge to the same value.

Publications:

1. Badr K., Hilliard M., Roberts N., He Q.P. and Wang J. (2019), Photoautotroph-Methanotroph Coculture – A Flexible Platform for Efficient Biological CO_2-CH_4 Co-utilization, Proceedings of 12th Dynamics and Control of Process Systems (DYCOPS 2019), Apr. 23-26, 2019, Florianópolis, BRAZIL, accepted;

Funding statement: This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under Award Number DE-SC0019181; and by the U.S. Department of Education, Graduate Assistance in Areas of National Need Program under Award Number P200A150074.

Determining the genetic and environmental factors underlying mutualism within a *Sphagnum* peatmoss microbiome system

David J. Weston[1*] (westondj@ornl.gov), Travis J. Lawrence[1], Alyssa Carrell[1], Megan Gable[1], Sara Jawdy[1], Adam Healey[2], John T. Lovell[2], Jeremy Schmutz[2]

[1]Oak Ridge National Laboratory, Oak Ridge, TN; [2]Joint Genome Institute, Walnut Creek, CA

Project Goals: To characterize the *Sphagnum*-diazotroph symbiosis by incorporating plant host *Sphagnum* and microbial genetic variation, variable climatic drivers, and complex communities that scale across biological organizations to regulate peatland carbon and nitrogen cycling.

The importance of plant-microbiome systems on terrestrial carbon and nitrogen processes is perhaps most pronounced in *Sphagnum* dominated ecosystems, which occupy 3% of the Earth's land surface yet store approximately 25% of terrestrial carbon as recalcitrant organic matter (i.e., peat). The foundation plant genus *Sphagnum* is responsible for much of the primary production in peatland ecosystems and produces recalcitrant dead organic matter. Together with associated N_2-fixing microorganisms, *Sphagnum* contributes to substantial peatland nitrogen inputs. *Sphagnum* growth and production (carbon gain) depends, in part, on a symbiotic association with N_2-fixing, diazotrophic microbes. Under changing environmental conditions, a central question about these ecosystems is whether the *Sphagnum*-diazotroph symbiosis will maintain its beneficial interaction, or will it shift to neutral or even antagonistic interactions that ultimately influence peatland carbon gain and storage. To begin to address this question, we are initiating a 5-year project using synthetic communities, genotype-to-phenotype associations, and metabolic characterization to address two overarching hypotheses, 1) *Sphagnum* host and diazotroph genetic variations play a key role in determining the environmental tipping point of beneficial symbiosis (i.e., environmental disruption), and 2) the surrounding microbiome can further adjust the tipping point through facilitation, competition, and antagonism.

To facilitate the testing of hypothesis one, we collaborated with the DOE JGI to develop a *Sphagnum fallax* genetic map for quantitative trait analysis (QTL). A 220 member pedigree was established in tissue culture conditions with 167 individual genotypes resequenced to 7X depth. The initial map is 5395.77 cM in length with an average spacing of 1.8 cM between markers. The largest gap is 9.3 cM between markers. Overall there are 19 well-defined linkage groups with an additional linkage group possibly representing a sex chromosome or micro-chromosome. Furthermore, 15 additional *Sphagnum* species have been sequenced to enable comparative genomic analysis. The end goal of this approach is to link gene and gene family evolution to QTLs playing a key role in symbiosis and tolerance to environmental perturbations. This combined approach has the potential to revel the history of selective pressure underlying symbiosis QTLs.

To address the second hypothesis, we developed a community transfer approach where native *Sphagnum* microbiomes conditioned to two-years of elevated temperature (ambient + 9 ˚C) or ambient temperature were isolated and applied to germ-free tissue culture *Sphagnum and* exposed to temperature manipulations. Our initial results indicate that the warming conditioned microbiome reduced *Sphagnum* mortality by 50% at warming temperatures relative to *Sphagnum* cultured with an ambient derived microbiome or no microbiome at all. Ongoing metagenome

analyses have assembled 38 microbial genome bins with greater than 90% completeness and less than 10% contamination. Relative abundance of binned genomes shows putative patterns of temperature specific taxa suggesting changes in metabolic potential. These data are now being augmented with metatranscriptomics.

The second year of this project will continue to link phylogenetic and quantitative genetic analyses to investigate the evolution of adaptive symbiosis traits. Furthermore, we are developing an additional pedigree and genetic map for *Sphagnum fuscum*, and will screen the temperature conditioned microbiomes across all *Sphagnum* pedigree members.

Funding: This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Early Career Research Program; and the DOE JGI Community Science Program 504399; and FICUS 504306.

Development of emerging model microorganisms: *Megasphaera elsdenii* for biomass and organic acid upgrading to fuels and chemicals

Lauren A. Riley[1,2], E. Anne Hatmaker[1], Adam M. Guss[1,2], **Janet Westpheling (janwest@uga.edu)[3]**

[1]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [2]University of Tennessee, Knoxville; [3]University of Georgia, Athens

The native ability to condense acetyl-CoA groups to efficiently generate C4 to C8 compounds makes *Megasphaera elsdenii* a compelling platform for the production of fuels and chemicals from lactate and plant carbohydrates. Our overall objective is to develop *Megasphaera elsdenii* as a platform for the conversion of lignocellulosic biomass sugars and organic acids into hexanol and other valuable chemicals. *M. elsdenii* produces organic acids as fermentation products when growing on lactate and glucose, including formation of butyric (four carbon), hexanoic (six carbon), and in some cases octanoic (eight carbon) acids as major fermentation products, likely via a chain elongation pathway using acetyl-CoA. As the carbon chain length increases, fuel properties improve, with the energy density increasing and hygroscopicity decreasing. Virtually nothing is known about the metabolic pathways in *M. elsdenii* that result in organic acid formation beyond predictions based on genome annotation. We have developed the first methods for DNA transformation of two strains of *Megasphaera elsdenii*, opening this organism to advanced physiological studies and bioengineering. We have begun the use of newly generated genome sequences to create a metabolic reconstruction of glucose and lactate conversion to hexanoic acid using the DOE KBase platform as the foundation of a metabolic model. We continue to develop genetic tools to enable more rapid and complex strain construction to develop strains capable of producing long chain carbon molecules at high yield and high titer. We will present a progress report on these efforts.

This work was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences Energy Frontier Research Centers program under Award Number DE-SC0019401.

Title: Developing the thermotolerant yeast *Kluyveromyces marxianus* as a host for next generation bioprocessing

Ian Wheeldon,[1]* (iwheeldon@engr.ucr.edu), Tami L. McTaggart[2], Danielle Bever[2], Shane Bassett[2], Xiyan Sun[1], Stefano Lonardi[1], and Nancy A. Da Silva[2]

[1]University of California, Riverside; [2]University of California, Irvine

Project Goals: This systems and synthetic biology project seeks to understand and engineer the native stress tolerance phenotypes of the yeast *Kluyveromyces marxianus* with the goal of developing a new synthetic biology chassis for fuel and chemical production.

Kluyveromyces marxianus is a promising nonconventional yeast for biobased chemical production due to its rapid growth rate, high TCA cycle flux, and tolerance to low pH and high temperature. Unlike *S. cerevisiae*, *K. marxianus* grows on low-cost substrates to cell densities that equal or surpass densities in glucose, which can be beneficial for utilization of lignocellulosic biomass (xylose), biofuel production waste (glycerol), and whey (lactose). This project seeks to understand and exploit these native traits to create a new thermotolerant eukaryotic host chassis for biofuel and biochemical products. Our efforts include developing genome-wide mutagenesis and regulation strategies to identify essential genes and the genotype-phenotype relationships underpinning the novel traits. The new systems-level data created in these experiments will inform metabolic engineering to increase the yield of acetyl-CoA derived products such as triacetic acid lactone (TAL) and ethyl acetate. To date, we have created and used a multiplexed CRISPR interference (CRISPRi) system to understand and engineer ethyl acetate synthesis on glucose, increasing production by upward of 3.8-fold.[1] We have also engineered TAL biosynthesis, with the highest titers achieved in xylose at 37 °C. The ~1.0 g/L titer achieved in mL-scale cultures shows promise for high titer, high yield production once scaled to bioreactor conditions. Ongoing work utilizes our newly developed CRISPR tools to modify native and heterologous pathways to increase the levels of the acetyl-CoA and malonyl-CoA, precursors to ester and TAL biosynthesis. Functional genetic screens to understand the origins of high thermal and acid tolerance are also ongoing.

References

1. Löbs AK, Schwartz C, Thorwall S, and Wheeldon I. Highly multiplexed CRISPRi repression of respiratory functions enhances mitochondrial localized ethyl acetate biosynthesis in *Kluyveromyces marxianus*. *ACS Synthetic Biology*, 2018;7:2647−2655.References are not required but should be placed here if needed.

This work was supported by DOE DE-SC0019093.

From genomes to methane emission: targeting critical knowledge gaps in wetland soils

Adrienne Narrowe[1*] (adrienne.narrowe@colostate.edu), Rebecca Daly[1], Jorge Villa Betancur[2], Garrett Smith[2], Mikayla Borton[1], David Hoyt[3], Lindsey Solden[2], Jordan Angle[2], Josue Rodriguez Ramos[1], Tim Morin[4], William Riley[5], Christopher Miller[6], Trent Northen[5], Gil Bohrer[2], **Kelly Wrighton**[1]

[1] Colorado State University, Fort Collins, CO

[2] The Ohio State University, Columbus, OH

[3] Pacific Northwest National Laboratory, Richland, WA

[4] State University of New York College of Environmental Science and Forestry, Syracuse, NY

[5] Lawrence Berkeley National Laboratory, Berkeley, CA

[6] University of Colorado Denver, Denver, CO

Project Goals: The overarching objective of this early career research project is to identify the biogeochemical and genomic determinants of methane production, as well as the scale and physical distribution over which they operate, along freshwater wetland gradients. Here we use field and laboratory data to provide highly resolved information on small scale spatial heterogeneity in forcing and identify the soil conditions impacting methane emission. Results from this project will be integrated into biogeochemical models to resolve methane flux predictions at the terrestrial-aquatic interface.

Abstract

Despite their relatively small land coverage, wetlands represent the largest natural source of atmospheric methane. Wetland emission budgets for this potent greenhouse gas are highly variable, with over 25% uncertainty. Accurately predicting net methane fluxes from wetlands depends on multiple interrelated geochemical, ecological, and metabolic constraints that are poorly understood, oversimplified, or missing in global biogeochemical models. This project studies the distribution, diversity, and gene expression of methane cycling microbial communities along wetland spatiotemporal gradients. The current paradigm is that oxygen inhibits methanogenesis and that most methane is produced from acetate or hydrogen; assumptions incorporated into global biogeochemical models today. In contrast, our wetland observations challenge these widely held views on microbial methane metabolism in soils.

Over two seasons, we have shown clear geochemical and biological evidence for 10 times greater methane production in well-oxygenated bulk soils of a freshwater wetland **(1)**. Recovery of the first near-complete genomes for a novel methanogen species, *Candidatus* Methanothrix paradoxum, using metagenomic and metatranscriptomic sequencing showed that acetoclastic methane production is dominant in oxygenated soils. Moreover, this microorganism was a dominant member of the soil community across this wetland over a three-year period, but was also active in other methane emitting ecosystems, suggesting a global significance. Importantly, in this wetland, we estimated that up to 80% of methane fluxes could be attributed to methanogenesis in oxygenated soils. Additionally, we showed that this surface soil methanogen activity is relatively stable by season, but methanotroph activity decreases in summer, thereby a reduction in consumption, rather than increased production, may be an important contributor to increased methane fluxes observed in summer **(2)**.

In addition to the dominant acetoclastic methanogen from surface oxygenated soils, our previous descriptions of archaeal diversity **(3)** revealed methylotrophic *Methanomassiliicoccales* were prevalent across the wetland over multiple years, especially in deeper anoxic soils. Given that little is known about the contribution from methylotrophic methanogenesis in soils, the microbial taxa, substrates, and possible contributions to methane production are here characterized. In contrast to other reports from wetland soils, we demonstrated that trimethylamine amended soils produce 60-fold greater methane than soils with endogenous methanogen substrates alone. NMR of porewater fluids indicated methanogen substrate availability varied by depth, with acetate and formate in higher concentration in surface soils, and methylotrophic substrates like TMA and methanol enriched in deeper soils. Consistent with the substrate profile, metatranscripts from *Methanomassiliicoccus* are three times higher in the deep soils. Metatranscript data also demonstrated that methanol, rather than methylamine, is the preferred substrate for these methanogens. Together our field and laboratory findings highlight that methylotrophic methanogenesis may be an underrepresented contributor to overall methane emissions in this system.

As a part of this DOE career award, we have developed an extensive database of over 1,900 genomes. We use this database, along with highly resolved spatial (cm scale) and temporal (monthly) sampling, to map the distribution of chemical and biological determinants of methane across wetland gradients. Toward this goal, in 2018 we completed an extensive field campaign, collecting over 700 soil samples from 9 sites every month for 6 months. Soil samples were paired to *in situ* methane measurements collected at the same cm-scale resolution, with data linked to overall methane fluxes. 16S rRNA gene analyses, along with untargeted soil metabolomics, will track methane cycling taxa distribution and substrate profiles across these samples. A JGI community sequencing project is currently generating 80 metagenomes and over 230 metatranscriptomes. Collectively, these data provide an unprecedented, high resolution view of methane production and consumption activities, resolved spatially and temporally across a wetland site. This research can facilitate the transition of climate models from treating methane flux in wetlands as a "black box", primarily relying on overall flux rates, to more process or spatially oriented models.

References

1. Angle JC, Morin TH, Solden LM, Narrowe AB, Smith GJ, Borton MA, Rey-Sanchez C, Daly RA, Mirfenderesgi G, Hoyt DW, Riley WJ, Miller CS, Bohrer G, Wrighton KC. 2017. Methanogenesis in oxygenated soils is a substantial fraction of wetland methane emissions. *Nature Communications* 8:1567.

2. Smith GJ, Angle JC, Solden LM, Borton MA, Morin TH, Daly RA, Johnston MD, Stefanik KC, Wolfe R, Gil B, Wrighton KC. 2018. Members of the Genus *Methylobacter* Are Inferred To Account for the Majority of Aerobic Methane Oxidation in Oxic Soils from a Freshwater Wetland. *mBio* 9.

3. Narrowe AB, Angle JC, Daly RA, Stefanik KC, Wrighton KC, Miller CS. 2017. High resolution sequencing reveals unexplored archaeal diversity in freshwater wetland soils. *Environmental Microbiology* 19:2192-2209.

This material is based upon an Early Career Award to KCW from the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018022

Application of a Designed Protein Cage to Increase Pathway Flux of Cellulolysis

Kevin Cannon (kacannon@g.ucla.edu)[1,2]*, Scott McConnell[1,2], Christian Morgan[3], Robert T. Clubb,[1,2,4] and **Todd O. Yeates**[1,2,4]

[1]UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, CA; [2]UCLA Department of Chemistry and Biochemistry, Los Angeles, CA; [3]UCLA Department of Ecology and Evolutionary Biology, Los Angeles, CA; and [4]UCLA Molecular Biology Institute, Los Angeles, CA

https://www.doe-mbi.ucla.edu/

Project Goals: The goal of this work is to improve methods for breaking down cellulose to glucose in the laboratory by using designed protein cages as a scaffold for celullase enzymes. This technology would aid in the development of efficient routes to production of ethanol and other biocommodities from microbial sources.

Motivated by the rich diversity of protein molecules that have evolved in nature to form complex and highly symmetric supramolecular structures, recent engineering efforts in the field of protein design have exploited symmetry to create novel self-assembling protein structures of types unseen in biology. Advances in designing proteins to self-assemble into specific architectures are now opening up numerous exciting technology applications. One such application of protein cages is in enzyme display for improving pathway flux of sequentially acting enzymes. We have created protein cages with cellulolytic activity by using the sortase enzyme SrtA to covalently link multiple cellulase enzymes to the exterior of a designed protein cage. By decorating these cages with both the endoglucanase Cel8A and the exoglucanase Cel48S, we have demonstrated a more than 5-fold increase in pathway flux of cellulolysis compared to the free enzymes in solution. These cellulolytic cages could find application in engineering more efficient routes to the production of ethanol and other biocommodities.

This work was funded by the DOE BER Office of Science award # DE-FC02-02ER63421.

Cadmium Toxicity Impacts the Transcriptome and Global Mineral Homeostasis in *Chlamydomonas reinhardtii*

Madeli Castruita[1]* (madeli@chem.ucla.edu), Rey Martin,[1] Sean Gallaher,[1] Christian Hofmann[2], Stephan Clemens[2], and **Sabeeha Merchant[3]**

[1]University of California, Los Angeles; [2]University of Bayreuth, Bayreuth Germany; and [3]University of California, Berkeley

Project Goals: One mechanism for removing cadmium from polluted sites is phytoremediation, which works by extraction of the contaminant from the soil or water, using plants or algae[1]. The use of plants (or algae for contaminated water) offers the advantage of low cost and carbon neutrality. The ultimate objective of this project is understanding cadmium metabolism in relation to essential metal homeostasis in algal and plant cells with the long term goal of promoting phytoremediation strategies for Cd.

The 2015 World's Worst Pollutant Problems report identified Cd as one of the top 6 toxic threats to human's health[2]. Human exposure to Cd primarily occurs from food consumption as Cd enters the food web from contaminated soil and waters. Cd exposure in humans results in kidney and renal damage, osteoporosis, cancer, and cardiovascular disease. In plants, Cd toxicity diminishes growth, photosynthesis, and crop yield and causes an alteration in mineral nutrition[3]. Phytoremediation of Cd contaminated sites via plants and algae is of great interest as it is a low cost, solar energy driven clean up technique. *Chlamydomonas reinhardtii* is an excellent reference organism for understanding the metabolic responses to Cd exposure because it has been extensively studied at the cellular lever and it grows in a simple well-defined salts medium.

We have identified two Cd conditions that elicit a phenotypic vs a non-symptomatic response in Chlamydomonas (sub-toxic vs. toxic). Our goal is to identify the mechanisms utilized for tolerance/resistance to Cd and the overall impact of Cd toxicity in the cell. Stress markers such as growth, photosynthesis, GSH/GSSG, and phytochelatin production were used to determine the impact of Cd in the cell. Cd at toxic levels alters the elemental composition profile of Chlamydomonas resulting in the hyper-accumulation of Fe and Cu and in a reduction of P and Ca content. The genome-wide analysis via RNA-Seq of sub-toxic vs. toxic Cd exposed cells identified both overlapping and unique responses to Cd.

References

1. Ali, H., Khan, E. and Sajad, M. A. Phytoremediation of heavy metals--concepts and applications. *Chemosphere* **91**, 869-881, (2013).

2. Pure Earth and Green Cross Switzerland. World's Worst Pollution Problems 2015. (2015) http://worstpolluted.org

3. Nagajyoti, P. C., Lee, K. D., and Sreekanth, T. V. M. Heavy metals, occurrence and toxicity for plants: a review. *Environmental Chemistry Letters* **8**, 199–216, (2010).

This work was supported by the Department of Energy (DE-FC03-02ER63421) and by the National Institutes of Health (GM42143).

Orchestrated Response of Dunaliella for Coping with Iron Limitation

Lital Davidi[1*] (lital@chem.ucla.edu), Sean D. Gallaher[1], Mary S. Lipton[2], **Sabeeha S. Merchant**[3]

[1]University of California, Los Angeles; [2]Environmental Molecular Sciences Laboratory, PNNL, Richland, WA; [3]University of California, Berkeley

Project goals: The specific goal of this project is to provide a broader view of iron assimilation, sub-cellular compartmentation and metabolism, by discovering pathways and mechanisms of intracellular iron trafficking and storage for iron re-use and re-cycling in the green microalgae *Dunaliella bardawil.*

Algae are responsible for half of global primary productivity and have an important role in biogeochemical cycling. Algae can thrive in nutrient-rich and nutrient-depleted environments, indicating their wide plasticity. Iron, which is one of the key elements for life on Earth, limits all forms of life due to its limited bioavailability. Photosynthetic organisms which have greater iron demand due to its use in photosynthesis in addition to respiration, can be in a chronic state of iron deficiency; Consequentially, have adapted iron economy mechanisms involving sparing and recycling. The goal of this project is to provide a cellular view of iron metabolism in *Dunaliella bardawil* by discovering and understanding the pathways and mechanisms for visualizing the dynamics of iron distribution during metabolic transitions, from iron luxury to iron economy conditions.

The genus Dunaliella represents globally abundant, broadly distributed unicellular green microalgae that are well known for their high tolerance to extreme abiotic stress conditions of salinity, light, temperature and pH, but are otherwise not well studied at the molecular level. One species, *D. bardawil*, is famous for its ability to accumulate β-carotene. Most importantly, a draft genome sequence of the closely related *D. salina*, has been published recently, enabling genome-level studies. Biochemical and physiological experiments indicate that *D. bardawil* can reduce its iron quota with only moderate chlorosis. Combination of transcriptome and proteome experiments on iron limited vs. replete cells, identified sentinel genes and proteins for monitoring iron status, with the most dramatic increases in the transcriptome being recapitulated in the proteome, such as transferrin and FOX1 in iron assimilation and Tidi and flavodoxin in the photosynthetic apparatus. The iron assimilation components include activities and proteins related to those known in Chlamydomonas, like FOX1, proteins known in diatoms, like p130b/ ISIP1 and CREG1, and novel proteins like transferrin, supporting the model that algal Fe assimilation pathways are a heterogenous mix and match of individual Fe uptake modules. Choosing *D. bardawil* as an organism of interest with its importance in industrial applications and its evolutionary relationship to *C. reinhardtii* allows us to both build on the knowledge that has been acquired already for *C. reinhardtii* and also to discover diverse strategies for acclimation to iron limitation. A systematic understanding of diverse strategies for optimizing iron utilization is a pre-requisite for exploiting iron-poor environments for food and fuel production.

The research was performed using DOE grant (DE-FG02-04ER15529) and EMSL (grid.49840), a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research.

Advances to Determine the Factors that Govern Stable Protein-display in the Model Gram-positive Bacterium *Bacillus subtilis*

Grace L. Huang,[1] Jason Gosschalk[1]* (jaygosschalk@ucla.edu), Ye Seong Kim,[1] Xiyan Sun,[2] and **Robert Clubb**[1]

[1]UCLA-DOE Institute of Genomics and Proteomics, University of California, Los Angeles, 611 Charles Young Drive East, Los Angeles, CA 90095, USA. [2]Chinese Academy of Sciences, Kunming Institute of Botany, Yunnan Kunming, China.

Project Goals: We aim to define genetic and biochemical elements that affect the abundance and activity of proteins displayed on the surfaces of Gram-positive bacteria. Our goal is to leverage this knowledge to engineer microbes that progress the DOE's objectives in bioenergy and bioremediation. Toward this goal, we are studying how to stabilize proteins and polymers on the surface of eubacteria that could be useful in degrading lignocellulose. To gain broad insight, our efforts are concentrated on *Bacillus subtilis*, a model Gram-positive microorganism that is used industrially. For practical applications, a diverse collection of surface-displayed cellulases need to be bound to the peptidoglycan at a high density while remaining highly stable. In this study, we systematically explored several well-established protease-deficient *B. subtilis* strains' ability to display cellulases, process carbohydrate, and assessed the impact of surface-display on bacterial growth, cellular stress and morphology. The results of this work can be used to guide the construction of cellulase-displaying microbes that convert plant biomass into second generation biofuels, chemicals, and materials.

Lignocellulosic biomass is a promising feedstock from which to sustainably produce useful biocommodities, but its recalcitrance to hydrolysis limits its commercial utility. One attractive strategy to overcome this problem is to use consolidated bioprocessing (CBP) microbes that directly convert biomass into chemicals and biofuels. Several industrially useful microbes possess desirable consolidated bioprocessing characteristics, yet they lack the ability to degrade biomass. Engineering these microbes' surfaces to display cellulases and cellulosome-like structures could endow them with potent cellulolytic activity, enabling them to be used in CBP. To gain general insight into factors controlling protein display, we engineered the surface of the model Gram-positive bacterium *B. subtilis*. We constructed a *B. subtilis* protein display reporter system in which the *Clostridium thermocellum* Cel8A endoglucanase is fused to the LysM cell wall binding module. The effects of LysM positioning, extracellular proteases, and solution conditions on the copy-number and stability of the reporter protein were determined. We demonstrated that heterologous surface enzyme activity is rapidly lost, even when *B. subtilis* is genetically modified to eliminate all of its extracellular proteases (AprE, Epr, Bpr, Vpr, NprE, NprB, Mpr, WprA, HtrA and HtrB). This problem presumably occurs because the membrane's proton motive force (pmf) dissipates when nutrients are scarce, leading to autolysis and the concomitant release of cytoplasmic proteases that degrade the heterologous surface proteins. We overcame cellular autolysis using a two-step procedure in which the pmf is maintained by glucose or glycerol additives. This procedure enables the production of cellulase-coated *B. subtilis* cells that are stable for more than two days, as substantiated by whole-cell enzyme activity measurements and cell fractionation experiments in tandem with immunoblotting. We observed that protein display affects cell morphology in certain strains and that the secretory stress response is activated as a result of both Cel8A-LysM display and deficiency of the HtrA and HtrB proteases; we have made strides toward determining the ideal *B. subtilis* strain and growth conditions for CBP. The ability to produce stably enzyme-coated vegetative *B. subtilis* is a step toward their practical use in biotechnological applications and lays the foundation for their further optimization.

Publications

Huang GL, Gosschalk JE, Kim YS, Ogorzalek Loo RR, Clubb RT. Stabilizing displayed proteins on vegetative Bacillus subtilis cells. *Appl Microbiol Biotechnol*. 2018;102(15):6547-6565.

This work was supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research program under Award Number DE-FC02-02ER63421. GL Huang was supported by a Ruth L. Kirschstein National Research Service Award GM007185.

LARKSdb: a tool for proteome-wide identification of functional low-complexity domains of proteins

Michael Hughes (mphughes@ucla.edu)[1]* and **David Eisenberg**[1]

[1]Department of Biological Chemistry and Department of Chemistry and Biochemistry, University of California Los Angeles (UCLA), Howard Hughes Medical Institute (HHMI), UCLA-Department of Energy (DOE) Institute for Genomics and Proteomics, Los Angeles, CA 90095, USA

Project Goals: We have previously identified structures, LARKS, that correlate with and explain labile amyloid associated with protein hydrogels and membraneless organelles. LARKS promote formation of reversible amyloid-like fibrils. Here we present LARKSdb, an online tool to predict LARKS forming segments and domains in proteins, and present preliminary results of LARKS predictions in non-human proteomes.

Membraneless organelles are assemblies of nucleic acids and proteins. Examples in eukaryotic cells include P-bodies, the nucleolus, nuclear gems, and stress granules. All of these are dynamic, rapidly rearranging bodies that perform functions for the cell. These assemblies are self-organizing in the absence of an enveloping membrane. How are such membraneless organelles organized in the absence of a physical barrier?

Low-complexity domains (LC) of proteins are important in organizing membraneless organelles (Kato *et al*., 2012). LC domains are regions of protein sequences with a biased composition of amino acids. In humans, LC domains often over represent glycine, serine, and alanine. *In vitro*, these domains phase separate and form labile hydrogels. These phase transitions mirror some of the behavior of membraneless organelles, so we sought structure of proteins that form labile hydrogels to better understand organization of membraneless organelles.

Closer inspection of hydrogels formed by the LC domains of FUS, hnRNPA1, and TDP43 revealed that they form protein fibrils that give diffraction indicative of amyloid fibrils (Kato *et al*., 2012). This was unexpected because pathogenic amyloid is typically irreversible, unlike the behavior of the hydrogels and membraneless organelles. We found that adhesive segments of LC domains form β-rich fibrils as in typical amyloid, but that the protein segments have sharp kinks in their peptide backbones. These kinks prevent extensive interfaces forming between mated β-sheets (Hughes *et al*., 2018). A consequence of a smaller interface is weakened interactions allowing for labile amyloid-like fibrils (Guenther *et al*., 2018). Reflecting their novel structure, we named them Low-complexity, Amyloid-like Reversible Kinked Segments (LARKS).

In previous work we computationally predicted the location of LARKS throughout the human proteome and found them to be enriched in proteins that form membraneless organelles (Hughes *et al*., 2018). We are launching an online server and database called LARKSdb enabling

researchers to predict LARKS within their proteins of interest. In humans we find that LARKS are enriched in LC domains. We present the following findings in our preliminary search for LARKS across kingdoms of life:

1) Organisms have variable LC coding regions in their proteomes; E. coli has 1.7% of proteins with LC domains, the corresponding values are 8.8% for yeast, 10.0% for tuberculosis, 42.0% for malaria, 26.0% for tetrahymena, and 15.9% for humans.

2) The number of LARKS rich proteins does not correlate with abundance of LCRs:

	E. coli	Yeast	Tuberculosis	Human
# proteins	4309	6049	3983	20396
# LC domain proteins	65	484	398	3243
# proteins > 10% LARKS	0	3	56	73

3) Functions of LARKS rich proteins varies by organism. In humans most LARKS-rich proteins are intracellular nucleic acid binding proteins and keratins (Hughes *et al.*, 2018). In tuberculosis the proteins are extracellular, possibly to interact with intracellular proteins from host eukaryotes. Nature seems to have co-opted LCR proteins and LARKS rich proteins to different functions.

LARKSdb enables scientists to identify LARKS in proteins of interest as an aid in understanding protein function. Therefore we believe that LARKSdb will be an important tool moving forward for scientists to interrogate the role of proteins in membranless organelles throughout the kingdoms of life.

References

E.L. Guenther, Q. Cao, H. Trinh, J. Lu, M.R. Sawaya, D. Cascio, D.R. Boyer, J.A. Rodriguez, M.P. Hughes, D.S. Eisenberg. Atomic structures of TDP-43 LCD segments and insights into reversible or pathogenic aggregation. Nat. Struct. Mol. Biol. (6) (2018), pp. 463-47.

M. Kato, T. W. Han, S. Xie, K. Shi, X. Du, L. C. Wu, H. Mirzaei, E. J. Goldsmith, J. Longgood, J. Pei, N. V. Grishin, D. E. Frantz, J. W. Schneider, S. Chen, L. Li, M. R. Sawaya, D. Eisenberg, R. Tycko, S. L. McKnight, Cell-free formation of RNA granules: Low complexity sequence domains form dynamic fibers within hydrogels. Cell 149, 753–767 (2012).

M.P. Hughes, M.R. Sawaya, D.R. Boyer, L. Goldschmidt, J.A. Rodriguez, D. Cascio, L. Chong, T. Gonen, D.S. Eisenberg. Atomic structures of low-complexity protein segments reveal kinked β sheets that assemble networks. Science, 359 (2018), pp. 698-701.

This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. We thank NSF MCB-1616265, NIH AG-054022, DOE, and HHMI for support

Characterization of the *Methanosaeta concilii* Sheath, a Remarkably Resilient Protein Assembly

John Muroski* (jmuroski@chem.ucla.edu), Farzaneh Sedighian, Rob Gunsalus, Rachel R. Ogorzalek Loo, and **Joseph A. Loo**

University of California, Los Angeles

https://www.doe-mbi.ucla.edu/

Project Goals: Understanding the proteinaceous sheath of *Methanosaeta concili* is critical for understanding how the methanogen interacts with its environment. Given the difficulty of solubilizing the assembly, these efforts will also provide insight into the analysis of such intractable protein subunits.

The sheath of certain methanogens makes up the outermost component of cells. Sheaths assemble into long structures that encapsulate multiple cells in a filamentous segment. The components that compose this structure are critical for two reasons. First, this filamentous layer must regulate entry of critical molecules, acetate specifically, into the cells, as well as the exit of carbon dioxide and methane. There are also mechanisms expected to be present that facilitate interspecies electron transport, a unique capability evidenced in many methanogens, including *Methanosaeta concilii*. The sheath also protects the cells of the methanogens, which can often exist in particularly harsh conditions. Elucidating the components of these structures would provide great insight into their physiology and regulation.

The chemical properties of some methanogen sheaths have been studied, and more recently, the protein identity of the *Methanospirillum hungatei* sheath was discovered. Surprisingly however, no homolog to the *M. hungatei* sheath protein was found in the genome of *M. concilii*, keeping its identity unknown. Further, *M. concilii* sheaths have demonstrated significantly different chemical properties from *M. hungatei* sheaths, and the former appears to be a much more stable assembly. Previous reports have also indicated that the sheaths of these archaea bear significantly more glycan than other identified sheaths. This investigation provides unique challenges and opportunities for analysis of this proteinaceous layer.

We have recently looked to characterize this sheath by exploring protocols to solubilize it and to identify proteins that associate with it. Cells were grown and sheaths isolated in the Gunsalus laboratory. Primarily SDS-PAGE and liquid chromatography coupled with tandem mass spectrometry have been employed in analyses. We have identified a number of protein factors that associate with the sheath and have found conditions to solubilize it, although the sheath protein's identity remains elusive.

By characterizing the sheath of *M. concilii*, we should be able to elucidate the environmental/cellular interface of this organism. This study lays groundwork for understanding the mechanisms by which these archaea comprise extremely durable structures and how they survive and thrive in atypical biological environments.

This work was funded by the Department of Energy.

Multivariant RNAseq reveal novel players in algae nitrogen metabolism

Stefan Schmollinger[1*](schmolli@chem.ucla.edu) and **Sabeeha S. Merchant[1,2]**

[1]University of California, Los Angeles, CA; [2]University of California, Berkeley, CA

URL: http://www.chem.ucla.edu/dept/Faculty/merchant/#research

Project goals: The single-celled, eukaryotic green alga *Chlamydomonas reinhardtii* is an excellent model to study plant metabolism while offering fast generation times and all the advantages of microbial systems. All the genomes are sequenced and well annotated, allowing for large scale transcriptomic or system biology analyses. In order to discover novel genes involved in essential metabolic processes and further the understanding of the alga we generated a multivariant RNASeq dataset within the JGI Gene Atlas project, with the goal to compare different carbon and nitrogen sources, growth regimes, light intensities and cell densities in a single dataset.

Abstract: *Chlamydomonas reinhardtii* is a unicellular green alga that has been widely used as a plant reference system for six decades, it has a quick generation time (~ 6h), can be synchronized and grown to high densities and its three genomes are sequenced and well-annotated. We have utilized Chlamydomonas as a reference organism to understand the principles underlying trace metal utilization and economy in a photosynthetic cell, and have identified a repertoire of assimilatory and distributive transporters, discovered mechanisms for reducing the metal quota and recycling metal cofactors from non-essential to essential proteins in situations of sustained elemental deficiency.

We used a multi-variant RNAseq approach to identify novel genes in nitrogen metabolism. A total of 22 different conditions were analyzed in triplicate, allowing for 300 individual comparisons, including 69 comparisons where only a single variable is changing. Among the utilized perturbations where variations of the carbon source (acetate vs CO_2), nitrogen source (NH_4, NO_3, urea), light intensity (100 vs 500 PAR), growth regime (16/8h day/night cycle vs continuously grown cultures) and culture density (10^5 vs 10^6 cells/ml). Variation of the growth regime was the most influential parameter altering gene expression, followed by the carbon source and cell density, while the variation of the nitrogen source and light intensity only affected a select group of transcripts. Consistently, growth regime and carbon source were responsible for the first two principal components in the dataset. To maximize discovery and improve confidence we used a consensus strategy that utilized multiple algorithms for differential gene expression analysis. Known genes involved in the individual comparisons allowed to verify the validity of the approach. In total we identified ~650 novel genes in the dataset, with various degrees of annotation, that we were able to categorize for their regulation to one of the studied perturbations.

This material is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research program under Award Number DE-FC02-02ER63421. The work conducted at the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Synthetic Biochemistry: Making Biofuels and Commodity Chemicals the Cell-Free Way

Saken Sherkhanov[1]* (saken@chem.ucla.edu), Salem Faham[2], Paul H Opgenorth[3], Sum Chan[1], Tyler P Korman[1], and **James U Bowie[1]**

[1] UCLA-DOE Institute for Genomics and Proteomics, University of California, Los Angeles, [2] Orthogon Therapeutics, Canton, MA, and [3] Joint BioEnergy Institute, Emeryville, CA

http://bowielab.mbi.ucla.edu/research-2/

The goal of the project is to develop cell-free production of isobutanol and terpene biofuels via commercially viable enzyme systems that can produce continuously for at least 1 week at volumetric productivities >1.5 g/L/hr, in > 90% yield and reach titers that lead to product phase separation. Significant enzyme engineering has been done on a large scale (*e.g.* on at least 10-20 individual enzymes) and the current isobutanol system is being scaled up for high-volume production.

Considerable effort is currently directed to engineer micro-organisms to produce useful chemicals. The greatest potential environmental benefit of metabolic engineering will be the production of high volume commodity chemicals, such as biofuels. Yet the high yields and concentrations required for the economic viability of low-value chemicals are particularly hard to achieve in microbes due to the myriad competing biochemical pathways and product toxicity. We are developing an alternative approach, which we call synthetic biochemistry. Synthetic biochemistry throws away the cells and builds biochemical pathways in reaction vessels using complex mixtures of isolated enzymes and extracted cofactors. As the only pathway in the vessel is the desired transformation, yields can approach 100%. The challenge for synthetic biochemistry is to replace the complex regulatory systems that exist in cells in a simplified form. We have designed and built highly robust systems that can operate continuously for long periods of time. We are working to improve longevity further and to scale synthetic biochemistry reactors. So far we have generated a highly stable enzyme system that provides continuous production of isobutanol, with titers of ~65 g/L and productivity of 1.6 g/L/hr.

This project is supported by DOE grants DE-FC02-02ER63421 and DE-AR0000556

Correlation of periodic changes in the transcriptome and proteome in the Chlamydomonas cell cycle

Daniela Strenkert[1*] (danielas@chem.ucla.edu), Samuel O. Purvine[2], Mary Lipton[2] and **Sabeeha Merchant**[1,3]

[1]Institute for Genomics and Proteomics, University of California, Los Angeles, CA
[2]EMSL, Richland, WA,
[3]University of California, Berkeley, CA

Project goals: The green alga *Chlamydomonas reinhardtii* has been a reference organism for addressing many questions in biology, and it also serves as a reference organism for commercial applications in photosynthetic eukaryotes, including biofuel production. We have reported previously on a multilayer data set of gene expression, metabolomics, proteomics and physiology during the Chlamydomonas cell cycle, utilizing a flat panel bioreactor system for reproducible synchronization of *Chlamydomonas reinhardtii*. We further expanded our understanding of the cell cycle by generating an additional full-scale proteomic data set augmented by a concept called the "proteomic ruler", which allows for absolute protein quantification and copy number estimates. We integrate this additional dataset to review the relationship between the proteome and the transcriptome during the course of one day/night cycle.

Chlamydomonas reinhardtii is a unicellular green alga that has been widely used as a plant reference system, it has a short generation time and its three genomes are sequenced and well-annotated. Previously, we analyzed expression patterns of all three genomes of *Chlamydomonas reinhardtii* over the course of a day in cultures synchronized by a 12 h dark 12 h light period. Nearly 85% of transcribed genes show differential expression, with different sets of transcripts being up-regulated during each phase of the cell cycle. Parallel measurements of select metabolites and pigments, physiological parameters and a subset of proteins offered the opportunity for inferring metabolic events and for evaluating the impact of the transcriptome on the proteome. Assessment of starch, total organic carbon and respiratory activity suggested that fermentative metabolism may dominate during the night. This multi-omics approach offers an unprecedented high-resolution, systems-level view of cellular processes as cells grow in the light period and divide in the dark from one to two cells (1).

In this study, we generated a high resolution proteomic data set of synchronized Chlamydomonas cells. Over the course of the day we could identify 10696 different proteins, ~ 2120 proteins were identified with at least two peptides in each of the 16 time points. We further expanded our proteomic data set by a concept called the "proteomic ruler" (PR), which allows for absolute protein quantification and estimates of protein copy number per cell (2). This method is based on the assumption that the MS signal of histones can be used as an internal standard, the so-called "proteomic ruler". Since the amount of histones is proportional to the DNA content of a sample and correlates with cell number, this internal normalization allows absolute protein quantification without external spike-in standards. In order to validate protein quantification of our cell cycle data set using the Proteome Ruler (PR) method, we first compared the quantitative readout for total protein content per cell obtained using PR with results generated independently experimentally from BCA positive material. We obtained excellent correlation between these data, with a ratio of total protein content per cell PR:BCA of 1:1.11 (±0.15). We also validated correct absolute scaling of our proteomic data by comparison with earlier studies using either a spike-in

of Quantification Concatamers (QconCATs) (3), spike-in of purified proteins combined with heavy-labeled Chlamydomonas cell extracts (4) or spiked-in synthetic, isotope-labeled peptides and LC-MS analysis by SRM (5). Since other proteomic data was obtained using asynchronous cultures, we used the averaged protein quantifications over the diurnal cycle for comparison. In the proteomic data set using QconCATs it was estimated that unsynchronized Chlamydomonas cells contained an average of 2.9 amol PS1 per cell (3). Using the proteome ruler approach, we observed an average of 2.7 amol PS1 / cell, ranging from 2.3 (night) to 4.4 (day) amol PS1 / cell over the course of the day. We used this quantitative dataset to confirm ratios of proteins in protein complexes involved in photosynthesis. Our observed ratio of PS1 to plastocyanin of 1:1.17 compares well with the 1:1.12 ratio observed by the method based on quantification concatamers (3) and the 1:1.4 ratio that was determined using a spike-in of purified proteins (4).

Rubisco is composed of eight chloroplast encoded large subunits (rbcL) and eight nucleus-encoded small subunits (RBCS) (6). Based on known mechanisms for the regulation of RbcL and RBCS gene expression (7,8), we assumed both subunits to be present in equimolar amounts to allow a 1:1 subunit stoichiometry of the holoenzyme. Thus, ratios between RbcL and RBCS of 11–44:1 (4), 5:1 (9), determined in earlier studies were quite surprising, with a ratio of 1.56:1 obtained using QconCATs being closest to expectations (3). Using the proteome ruler approach, we estimated that the large subunit of Rubisco is present in average 8.9 amol / cell during the cell cycle, with a ratio of rbcL to RBCS of 1:0.98.

We further use these data to review the relationship between the proteome and the transcriptome during the course of a full diurnal cycle. Changes in transcript and protein abundances correlate in the majority of cases, if a delayed response at the protein level is taken into account.

1. Strenkert, D., Schmollinger, S., Gallaher, S. D., Salome, P. A., Purvine, S. O., Nicora, C. D., Mettler-Altmann, T., Soubeyrand, E., Weber, A. P. M., Lipton, M. S., Basset, G. J., and Merchant, S. S. (2019) *Proc Natl Acad Sci U S A*
2. Wisniewski, J. R., Hein, M. Y., Cox, J., and Mann, M. (2014) *Mol Cell Proteomics* **13**, 3497-3506
3. Hammel, A., Zimmer, D., Sommer, F., Muhlhaus, T., and Schroda, M. (2018) *Frontiers in Plant Science* **9**
4. Nikolova, D., Heilmann, C., Hawat, S., Gabelein, P., and Hippler, M. (2018) *Photosynthesis research* **137**, 281-293
5. Wienkoop, S., Weiss, J., May, P., Kempa, S., Irgang, S., Recuenco-Munoz, L., Pietzke, M., Schwemmer, T., Rupprecht, J., Egelhofer, V., and Weckwerth, W. (2010) *Mol Biosyst* **6**, 1018-1031
6. Taylor, T. C., Backlund, A., Bjorhall, K., Spreitzer, R. J., and Andersson, I. (2001) *Journal of Biological Chemistry* **276**, 48159-48164
7. Wostrikoff, K., and Stern, D. (2007) *P Natl Acad Sci USA* **104**, 6466-6471
8. Khrebtukova, I., and Spreitzer, R. J. (1996) *P Natl Acad Sci USA* **93**, 13689-13693
9. Recuenco-Munoz, L., Offre, P., Valledor, L., Lyon, D., Weckwerth, W., and Wienkoop, S. (2015) *J Proteomics* **113**, 143-153

This material is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research program under Award Number DE-FC02-02ER63421. Part of the research was performed using Environmental Molecular Sciences Laboratory (Ringgold ID 130367).

Title: Cell-free enzymatic biosynthesis of cannabinoids

Meaghan A. Valliere[1,2], (meaghan.valliere@gmail.com), Tyler P. Korman[1,2], Nicholas B. Woodall[3], David Baker[3], James U. Bowie[1,2]

[1]Molecular Biology Institute, University of California Los Angeles; [2]UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, CA; [3]Department of Biochemistry, Institute for Protein Design, University of Washington, Seattle

http://bowielab.mbi.ucla.edu/research-2/

Project Goals: To design cell-free enzymatic systems to produce complex natural products, like cannabinoids.

Abstract

Prenylation of natural compounds adds structural diversity, alters biological activity, and enhances therapeutic potential. Because prenylated compounds often have a low natural abundance, alternative production methods are needed. Metabolic engineering enables natural product biosynthesis from inexpensive biomass, but is limited by complexity of secondary metabolite pathways, intermediate and product toxicities, and substrate accessibility. Alternatively, enzyme catalyzed prenyl transfer provides excellent regio- and stereo-specificity, but requires expensive isoprenyl pyrophosphate substrates. Here we develop a flexible cell-free enzymatic prenylating system that generates isoprenyl pyrophosphate substrates from glucose to prenylate an array of natural products. The system provides an efficient route to cannabinoid precursors cannabigerolic acid (CBGA) and cannabigerovarinic acid (CBGVA) at > 1 g/L, and a single enzymatic step converts the precursors into cannabidiolic acid (CBDA) and cannabidivarinic acid (CBDVA) respectively. Cell-free methods may provide a powerful alternative to metabolic engineering for chemicals that are hard to produce in living organisms.

This work was supported by DOE grants DE-FC02-02ER63421 and DE-AR0000556 to J.B., and an NIH National Research Service Award 5T32GM008496 to M.V.

A Designed Scaffold for Near-Atomic Resolution Cryo-EM Imaging of Small Proteins

Yuxi Liu,[1] Matthew Agdanowski,[1,2] Duc Huynh,[2] Shane Gonen,[3] Tamir Gonen,[3] **Todd O. Yeates**[1,2]* (yeates@mbi.ucla.edu)

[1]UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, CA; [2]UCLA Department of Chemistry and Biochemistry, Los Angeles, CA; and [3]Howard Hughes Medical Institute, Janelia Research Campus, Ashburn, VA

https://www.doe-mbi.ucla.edu/

Project Goals: Research in the UCLA-DOE Institute for Genomics and Proteomics includes major efforts in imaging technology development. The specific goal of this work is to make it possible to determine high resolution structures of proteins smaller than 50 kDa by cryo-EM, so that this powerful technique can be applied to cellular proteins and enzymes within the research scope of the DOE.

Abstract:

Recent technical advances in cryo-electron microscopy (cryo-EM) have made it possible to determine the three-dimensional structures of large protein assemblies and macromolecular complexes with atomic level detail. However, proteins smaller than about 50 kDa are currently too small to be imaged at high resolution by cryo-EM, leaving most protein molecules in the cell beyond the reach of this powerful structural technique. In recently published work (1), we designed a modular, symmetric scaffolding system to rigidly bind and display smaller proteins to make them amenable to visualization by cryo-EM. Our novel protein scaffold solves two key problems – rigidity and modularity – that have limited the utility of previous scaffolding methods. A designed protein cage with cubic symmetry serves as the core. A small 17-kDa protein (DARPin) serves as a modular adaptor, which can be edited (e.g. based on phage display experiments) to bind a wide range of target proteins. The DARPin adaptor (which is itself alpha-helical) is genetically fused to the self-assembling cage subunit in a semi-rigid fashion by a continuous alpha helical connection. Imaging the scaffold by itself (without a bound cargo protein) showed that the DARPin was held rigidly enough to visualize it at near-atomic resolution (1). In new work (2), this protein scaffold is used for the first time to bind and display 12 copies of a small 26 kDa protein, sfGFP. We show that the bound cargo protein is held rigidly enough on the exterior of the scaffold to visualize it at a resolution of 3.8 Å. Structural details of the cargo protein are visible, making it the first demonstration of near-atomic resolution for a protein smaller than 50 kDa by cryo-EM. The designed scaffold is modular and can be modified through modest changes in its amino acid sequence to bind and display diverse proteins for imaging, thus providing a general method to break through the lower size limitation in cryo-EM. Future aims for optimization and broader application are noted.

Publications

1. Liu Y, Gonen S, Gonen T, Yeates TO. (2018). Near-atomic cryo-EM imaging of a small protein displayed on a designed scaffolding system. *Proc Natl Acad Sci USA* **115**, 3362-3367.
2. Liu, Y., Huyn, D., and Yeates, T.O. (2018) A 3.8 Å Resolution Cryo-EM Structure of a Small Protein Bound to a Modular Imaging Scaffold, https://www.biorxiv.org/content/10.1101/505792v1

This work was funded by the DOE BER Office of Science award # DE-FC02-02ER63421.

Systematic debottlenecking of isobutyraldehyde production in cyanobacteria

Yi Ern Cheah[1]* (yi.ern.cheah@vanderbilt.edu), Yao Xu[2], Sarah Sacco[1], Carl H. Johnson[2], **Jamey D. Young**[1,3]

[1] Chemical & Biomolecular Engineering, Vanderbilt University, Nashville, TN; [2] Biological Sciences, Vanderbilt University, Nashville, TN; [3] Molecular Physiology & Biophysics, Vanderbilt University, Nashville, TN

http://www.vanderbilt.edu/younglab

Project Goals: This project aims to develop a systematic approach to identify and alleviate bottleneck reactions that limit carbon flux towards desired products in engineered cyanobacteria.

Our lab has previously developed isotopically nonstationary ^{13}C-MFA (INST-MFA) (*1-4*) as a technology to assess the photoautotrophic metabolism of cyanobacteria (*4*) and plant leaves (*5*). Recently, we combined INST-MFA with rational metabolic engineering to improve the productivity of an isobutyraldehyde (IBA) producing mutant of the cyanobacterium *Synechococcus elongatus* PCC7942 (*6*).

This presentation describes our current effort at refactoring previously engineered cyanobacteria strains (in (*6*)) to identify the metabolic phenotypes that contribute to enhanced strain performance. The results led us to hypothesize that the bottleneck to IBA lies around the pyruvate node. By perturbing the fluxes around pyruvate, we were successful at further increasing IBA productivity.

Our efforts demonstrate that INST-MFA can play an important role in the strain optimization workflow of industrially relevant autotrophic systems.

References

1. L. J. Jazmin, J. D. Young, in *Systems Metabolic Engineering: Methods and Protocols,* H. S. Alper, Ed. (Humana Press, Totowa, NJ, 2013), pp. 367-390.
2. A. Adebiyi, L. Jazmin, J. Young, 13C flux analysis of cyanobacterial metabolism. *Photosynthesis Research*, 1 (2014).
3. J. D. Young, INCA: a computational platform for isotopically non-stationary metabolic flux analysis. *Bioinformatics* **30**, 1333 (May, 2014).
4. J. D. Young, A. A. Shastri, G. Stephanopoulos, J. A. Morgan, Mapping photoautotrophic metabolism with isotopically nonstationary C-13 flux analysis. *Metab. Eng.* **13**, 656 (Nov, 2011).
5. F. F. Ma, L. J. Jazmin, J. D. Young, D. K. Allen, Isotopically nonstationary C-13 flux analysis of changes in Arabidopsis thaliana leaf metabolism due to high light acclimation. *P Natl Acad Sci USA* **111**, 16967 (Nov 25, 2014).
6. L. J. Jazmin *et al.*, Isotopically nonstationary 13C flux analysis of cyanobacterial isobutyraldehyde production. *Metab. Eng.* **42**, 9 (2017/07/01/, 2017).

This work is funded by the U.S. Department of Energy (DOE) Awards DE-SC0019404 and DE-SC008118.

Modeling carbon metabolism of the diatom *Phaeodactylum tricornutum*

Amy Zheng[1*](amy.o.zheng@vanderbilt.edu), Yi Ern Cheah[1] and **Jamey D. Young**[1,2]

[1]Chemical and Biomolecular Engineering, [2] Molecular Physiology & Biophysics, Vanderbilt University, Nashville, TN

http://www.vanderbilt.edu/younglab

Phaeodactylum tricornutum (Pt), a unicellular diatom species, has the ability to store up to 45% of dry cell weight as triacylglycerol (TAG), a precursor to biodiesel [1]. To take advantage of this innate ability, we need to understand the metabolic pathways that supply the substrates needed to synthesize TAG.

Our lab focuses on using ^{13}C metabolic flux analysis (MFA) to map the central metabolism of numerous autotrophic systems, such as plants and cyanobacteria, using $^{13}CO_2$ labeling experiments [2,3]. In this project, we are developing a detailed metabolic model that describes the primary metabolism of Pt and encompasses precursors necessary for TAG biosynthesis. In addition, we are adapting our analytical methods (GC-MS and LC-MS/MS) to obtain the measurements needed to perform MFA in Pt under photoautotrophic conditions.

This presentation will discuss the construction, testing, and validation of an initial network model as well as optimal experiment design to maximize pathway resolution and flux precision. Our long term goal is to develop a ^{13}C MFA pipeline that would be integral for guiding rational metabolic engineering to optimize TAG production in Pt.

(Supported by grant *DE-SC0018344: Design, Synthesis, and Validation: Genome Scale Optimization of Energy Flux through Compartmentalized Metabolic Networks in a Model Photosynthetic Eukaryotic Microbe* from the Department of Energy.)

1 Hu, Q. et al. Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *Plant Journal* **54**, 621-639, doi:10.1111/j.1365-313X.2008.03492.x (2008).
2 Jazmin, L.J. et al. Isotopically nonstationary ^{13}C flux analysis of cyanobacterial isoburyaldehyde production. *Metabolic Engineering* **42**, 9-18, https://doi.org/10.1016/j.ymben.2017.05.001 (2017)
3 Ma, F. et al. Isotopically nonstationary ^{13}C flux analysis of changes in Arabidopsis thaliana leaf metabolism due to high light acclimation. *Proc Natl Acad Sci U S A.* **111**(47), 16967-16972, doi: 10.1073/pnas.131948511 (2014)

Investigating the Switchgrass Diterpenoid Metabolic Network Toward Improved Stress Tolerance and Biofuel Production

Kyle A. Pelot[1], Ruibing Chen[1], D. Max Hagelthorn[1], Cari A. Young[1], J. Bennett Addison[2], Andrew Muchlinski[3], Dorothea Tholl[3], **Philipp Zerbe**[1]* (pzerbe@ucdavis.edu)

[1]University of California, Davis; [2]San Diego State University, San Diego; [3]Virginia Polytechnic Institute and State University, Blacksburg

Project Goals:

Of the myriad specialized metabolites that plants form to adapt to environmental challenges, terpenes form the largest group. In many major crops, unique terpene blends serve as key stress defenses that directly impact plant fitness and yield. In addition, select terpenes are used for biofuel manufacture. Thus, engineering of terpene metabolism provides a versatile resource for advancing biofuel feedstock production, but requires a system-wide knowledge of the diverse biosynthetic machinery and defensive potential of often species-specific terpene blends. This project would merge genome-wide enzyme discovery with comparative –omics, protein structural and plant microbiome studies to define the biosynthesis and stress-defensive functions of switchgrass (*Panicum virgatum*) terpene metabolism. These insights would be combined with the development of genome editing tools to design plants with desirable terpene blends for improved biofuel production on marginal lands.

Abstract:

Diterpenoids constitute a diverse class of metabolites with critical functions in plant development, defense, and ecological adaptation. Major monocot crops, such as maize (*Zea mays*) and rice (*Oryza sativa*), deploy diverse blends of specialized diterpenoids as core components of biotic and abiotic stress resilience. Here, we describe the genome-wide identification and functional characterization of stress-related diterpene synthases (diTPSs) in the bioenergy crop switchgrass (*Panicum virgatum*). Mining of the allotetraploid switchgrass genome identified the largest thus far known diTPS family in plants, comprising 31 members. Biochemical analysis of 11 diTPSs revealed a modular metabolic network producing a diverse array of diterpenoid metabolites. In addition to *ent*-copalyl diphosphate (CPP) and *ent*-kaurene synthases likely involved in gibberellin biosynthesis, we identified *syn*-CPP and *ent*-labda-13-en-8-ol diphosphate (LPP) synthases as well as two diTPSs forming (+)-labda-8,13E-dienyl diphosphate (8,13-CPP) and *ent-neo-cis-trans*-clerodienyl diphosphate (CT-CLPP) scaffolds not observed previously in plants. Structure-guided protein mutagenesis of the (+)-8,13-CPP and *ent-neo*-CT-CLPP synthases revealed active site determinants that may resemble neo-functionalization events that occurred during diversification of the switchgrass diTPS family. Formation of several specialized diterpenoids, such as 9β-hydroxy-syn-pimar-15-ene, along with the expression of the corresponding diTPS genes was induced in roots and leaves in response to oxidative stress and UV irradiation, indicating their possible roles in abiotic stress adaptation. Together, these findings expand the known chemical

space of diterpenoid metabolism in monocot crops toward systematically investigating and ultimately improving stress resilience traits in bioenergy crop species.

References

1. Pelot KA, Chen R, Hagelthorn DM, Young CA, Addison JB, Muchlinski A, Tholl D, Zerbe P (2018) Functional Diversity of Diterpene Synthases in the Biofuel Crop Switchgrass. *Plant Physiol*. 178(1):54-71. doi:
2. Pelot KA, Hagelthorn DM, Hong YJ, Tantillo DJ, Zerbe P (2019) Diterpene Synthase-Catalyzed Biosynthesis of Distinct Clerodane Stereoisomers. *Chembiochem* 20(1):111-117.

Funding statement.

This work was supported by the Department of Energy (DOE) Early Career Research Program (DE-SC0019178, to P.Z.), a DOE Joint Genome Institute Community Science Program grant (CSP2568, to P.Z.), and a University of California–Davis Academic Senate Faculty Research Grant (to P.Z.). In addition, a NSF Graduate Research Fellowship Program (to K.A.P.) is gratefully acknowledged.

Transcriptome and metabolome analysis of the oleaginous yeast *Rhodosporidium toruloides* uncovered new gene targets for metabolic engineering

Anshu Deewan* (deewan2@illinois.edu), Sujit Sadashiv Jagtap, Jing-Jing Liu, Hanna Walukiewicz, Matthew Plutz, and **Christopher V Rao**

Institute for Genomic Biology, University of Illinois at Urbana-Champaign

http://www.scs.illinois.edu/rao/index.php

Project Goal: The goal of this project is to engineer the oleaginous yeast *Rhodosporidium toruloides* for the production of biofuels and bioproducts from plant based sugars and lipids. We are also interested in understanding the mechanism of substrate utilization, metabolite identification, and identification of the key genes governing the lipogenesis process.

Oleaginous yeasts are promising hosts for producing biofuels and bioproducts such as biodiesel, organic acids, polyols, jet fuels, and alcohols from renewable lignocellulosic biomass[1-2]. *Rhodosporidium toruloides* IFO0880 naturally accumulate lipids from multiple simple sugars when some other essential nutrient such as nitrogen is limiting[3]. Recently, we have engineered *R. toruloides* for increased lipid production during growth on glucose[4-5]. In addition to lipid-based chemicals, *R. toruloides* also produces a number of sugar alcohols at high titers. For example, during growth on xylose in nitrogen-rich medium, *R. toruloides* produced D-arabitol[6]. D-arabitol is listed by the Department of Energy as one of its top value-added chemicals from biomass.

We have performed transcriptome and metabolome analysis of *R. toruloides* during growth on glucose, xylose, acetic acid, and lipids. The principle component analysis (PCA) of the transcriptome and metabolome data showed a clear separation between different carbon sources in *R. toruloides*. We mapped the gene expression and metabolite concentrations on the metabolic pathways of *R. toruloides*. These results revealed that different metabolite pathways are activated under different carbon sources. The results provide a better understanding of the mechanism of

substrate utilization and the identification of the key genes governing the lipogenesis process. We also identified and functionally characterized a few putative sugar transporters from *R. toruloides* in *Saccharomyces cerevisiae*. The integration of the metabolite data of central carbon metabolism with gene regulation offers us a better understanding of the metabolic response of *R. toruloides* on different substrates. These results could provide the future directions for metabolic engineering of oleaginous yeasts for the production of biofuels and bioproducts.

References

1. Jagtap, S. S.; Rao, C. V., Microbial conversion of xylose into useful bioproducts. *Applied Microbiology and Biotechnology* **2018,** *102* (21), 9015-9036.
2. Isikgor, F. H.; Becer, C. R., Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polymer Chemistry* **2015,** *6* (25), 4497-4559.
3. Ageitos, J. M.; Vallejo, J. A.; Veiga-Crespo, P.; Villa, T. G., Oily yeasts as oleaginous cell factories. *Applied Microbiology and Biotechnology* **2011,** *90* (4), 1219-1227.
4. Zhang, S.; Ito, M.; Skerker, J. M.; Arkin, A. P.; Rao, C. V., Metabolic engineering of the oleaginous yeast *Rhodosporidium toruloides* IFO0880 for lipid overproduction during high-density fermentation. *Applied Microbiology and Biotechnology* **2016,** *100* (21), 9393-9405.
5. Zhang, S.; Skerker, J. M.; Rutter, C. D.; Maurer, M. J.; Arkin, A. P.; Rao, C. V., Engineering *Rhodosporidium toruloides* for increased lipid production. *Biotechnology and bioengineering* **2015**.
6. Jagtap, S. S.; Rao, C. V., Production of D-arabitol from D-xylose by the oleaginous yeast *Rhodosporidium toruloides* IFO0880. *Applied microbiology and biotechnology* **2018,** *102* (1), 143-151.

This work was supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award number DE-SC0018260.

A Comprehensive Genome-Scale Model for *Rhodosporidium toruloides* IFO0880 Accounted for Functional Genomics and Phenotypic Data

Hoang V. Dinh[1]* (hvd5034@psu.edu), Patrick F. Suthers[1], Siu Hung Joshua Chan[1], Yihui Shen[2], Anshu Deewan[3], Sujit Sadashiv Jagtap[3], Joshua D. Rabinowitz[2], Christopher V. Rao[3], Costas D. Maranas[1], and **Huimin Zhao**[3]

[1]The Pennsylvania State University, University Park, PA; [2]Princeton University, Princeton, NJ; [3]University of Illinois at Urbana-Champagne, Urbana, IL

https://www.igb.illinois.edu/DOEcenter

Project Goals: Non-model yeasts are yet under the limelight for metabolic engineering efforts despite possessing advantageous physiological traits as production hosts. To assist the effort in understanding and engineering non-model yeast, a genome-scale metabolic model can be reconstructed capturing the comprehensive metabolism of the organism. As a part of the Genome Scale Engineering project and in conjunction with genome-scale engineering (i.e., CRISPR-based tools) and analysis (i.e., -omics data) techniques, engineered strains can be built, understood, and tested in an automated manner at the Illinois Biological Foundry for Advanced Biomanufacturing. *Rhodosporidium toruloides* can accumulate lipid up to >65% cellular weight and is therefore an attractive host for the bioproduction fatty acids-derived products. This work presents a comprehensive genome-scale model for the oleaginous yeast *Rhodosporidium toruloides* IFO0880 accounted for functional genomics and phenotypic data.

Rhodosporidium toruloides is a basidiomycetes yeast that can accumulate large amount of lipids and natively produce carotenoids. A genome-scale model of *R. toruloides* IFO0880's metabolic network is reconstructed to better assess this non-model yeast's metabolic capabilities by integrating the latest available knowledge with in-house generated biomass composition, growth yield and viability. The model captures the most recent annotations from the latest version of the *R. toruloides* genome with organism-derived macromolecular composition in the biomass description and ATP maintenance requirements. It contains two separate biomass compositions depending on the growth condition (i.e., carbon or nitrogen limitation, respectively). The gene-protein-reaction rules and transporter system assignments are revised leading to gene essentiality prediction accuracy at a level similar to the latest *S. cerevisiae* model (yeast 7.6). The metabolic model is used to predict growth and lipid production phenotypes, contrast predicted metabolic flux redistribution and expression change, and facilitate *in silico* prediction of metabolic engineering strategies for the overproduction of triacylglycerol using the OptForce algorithm. The model predictions were in good agreement with functional genomics and phenotypic data and meaningful regulatory perturbation strategies were obtained from using the strain design algorithm on the model.

This work is supported by U. S. Department of Energy Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018260.

Development of a genetic toolbox for metabolic engineering of *Issatchenkia orientalis*

Vinh Tran[1], **Zia Fatma**[1,*] (ziafatma@illinois.edu), Mingfeng Cao[1], Xiaofei[1], Patrick F. Suthers[2], Costas D. Maranas[2], and Huimin Zhao[1,#]

[1]University of Illinois at Urbana Champaign, IL, USA
[2]The Pennsylvania State University, PA, USA, *Presenting author, [#]Correspondence

https://www.igb.illinois.edu/DOEcenter

Project Goals: To develop precise genetic tools for efficient metabolic engineering of non-model yeast *Issatchenkia orientalis* for organic acid production.

Abstract

There is a growing interest in engineering microbial cell factories for the production of chemicals and fuels from renewable feedstock. Among the twelve platform chemicals identified by the US Department of Energy (DOE), most of them are organic acids. Biosynthesis of these molecules leads to a multistress condition such as low pH, high temperature and accumulation of inhibitory products. *Issatchenkia. orientalis* is acid tolerant, ethanol tolerant, and thermotolerant, and is thus a multistress tolerant yeast and an ideal heterologous host for the organic acid production. However, lack of effective genome editing tools as well as nonavailability of its transcriptomics, proteomics, metabolomics, and genome-scale model have hindered its widespread application in biomanufacturing of industrially relevant molecule. In this study, an efficient genome editing method was developed for *I. orientalis* using a CRISPR-Cas9 system. We have constructed a single plasmid system, carrying 20 bp spacer, scaffold sequences, the gene for Cas9 protein expression, and 100 bp homology arm. The *ADE2* gene was deleted with the efficiency of more than 95%. Further, we have shown the application of these tools for the experimental validation of a genome-scale metabolic model of *I. orientalis*.

Funding statement: *This work is supported by U. S. Department of Energy Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018260.*

Model-guided analysis of the role of OPI1 and RPD3 transcription factors in hexadecanol overproduction strategies

Shekhar Mishra[1]*(smishr10@illinois.edu), Costas Maranas[2], **Huimin Zhao**[1,3]

[1]Department of Chemical and Biomolecular Engineering, Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States

[2]Department of Chemical Engineering, Penn State University, University Park, PA 16802, United States

[3]Departments of Chemistry, Biochemistry, and Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States

Project goals:

To explore the effect of transcription factor knockouts on lipid metabolism in the context of fatty alcohol overproduction to design beneficial strain engineering strategies.

https://www.igb.illinois.edu/DOEcenter

Abstract:

Obtaining a complete understanding of the lipid metabolic network has wide reaching implications in biofuel and biochemical production, one of them being fatty alcohols, where engineering the lipid network can produce higher yields of target biochemical. However, the strong regulation of lipid metabolism makes it difficult to engineer strategies for overproduction. Various studies regarding hexadecanol overproduction in yeast that explored the knockouts of transcription factors OPI1 and RPD3 report seemingly contrasting results about the effects of these knockouts on hexadecanol yield[1,2]. We employed a mathematical model-guided approach to study the flux control coefficients within lipid metabolism under different the transcription factor (TF) knockouts of OPI1 and RPD3. Our model-based analysis shows that knockout of these transcription factors results in the loss of a dynamic regulatory feature originally in place to conserve carbon flux in membrane lipid biosynthesis. As a result, upstream overproduction strategies coupled with the TF knockouts, as employed in the studies, result in an uncontrolled amount of flux being channeled towards membrane lipid biosynthesis, reducing the yield of fatty acid-derived chemicals. Our analysis demonstrates that mitigation strategies, such as downregulation of membrane biosynthesis, when applied in tandem with TF knockouts can be a useful method to harness the full benefit of these knockouts for overproduction of fatty acid-derived biochemicals such as fatty alcohols.

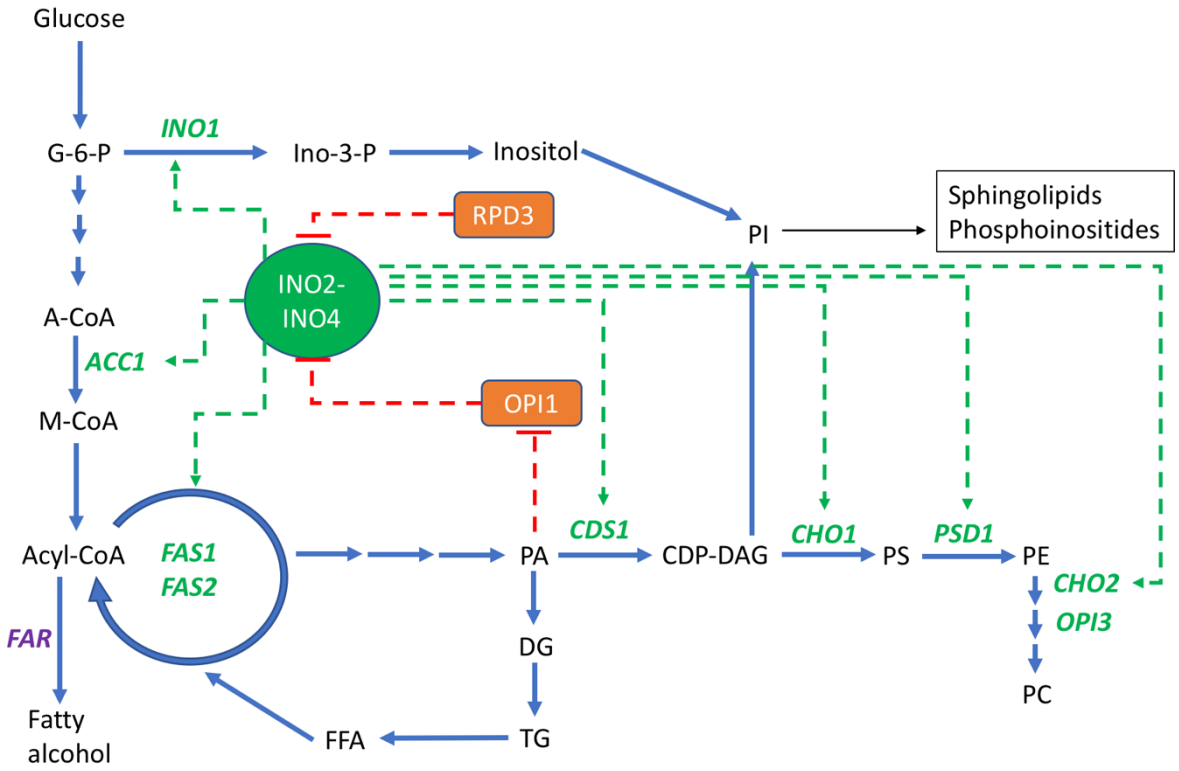


Figure 1: Map of the lipid metabolic network displaying the role of transcription factors OPI1 and RPD3 on the various biosynthetic genes in fatty acid chain elongation and membrane lipid biosynthesis reactions.

References:

1. d'Espaux L, Ghosh A, Runguphan W, et al. Engineering high-level production of fatty alcohols by Saccharomyces cerevisiae from lignocellulosic feedstocks. *Metab Eng*. 2017;42(April):115-125. doi:10.1016/j.ymben.2017.06.004.
2. Feng X, Lian J, Zhao H. Metabolic engineering of Saccharomyces cerevisiae to improve 1-hexadecanol production. *Metab Eng*. 2015;27:10-19. doi:10.1016/j.ymben.2014.10.001.

Funding statement:
This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DESC0018260.

Metabolic flux measurement in *I. orientalis* and *R. toruloides*

Yihui Shen[1,2]* (yihuis@princeton.edu), Tianxia Xiao[1,2], Xi Xing[1,2], Sarat Gopalakrishnan[3], Costas D. Moranas[3], and **Joshua D. Rabinowitz**[1,2]

[1]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ;
[2]Department of Chemistry, Princeton University, Princeton, NJ;
[3]Department of Chemical Engineering, The Pennsylvania State University, University Park, PA

https://www.igb.illinois.edu/DOEcenter

Project Goals: Measure metabolic flux in *I. orientalis* and *R. toruloides* using isotope tracers and mass spectrometry, in combination with quantitative modeling

Understanding of native metabolic pathway activity and regulation is valuable for accelerating metabolic engineering efforts. One reason for the great success of metabolic engineering in *S. cerevisiae* is decades of accumulated knowledge from its use as a model laboratory organism. Other organisms are likely intrinsically superior for many metabolic engineering objectives; however, their utilization is hindered by lack of knowledge. We are engaged in an effort to rapidly advance understanding of metabolism in two non-model yeasts, *I. orientalis* (which is desirably acid tolerant) and *R. toruloides* (which desirably synthesizes lipids). To this end, we have initially focused on mapping their metabolic pathway activity in batch culture and various nutrient limitations. Here, we will present data using different ^{13}C-glucose tracers to explore central carbon metabolic activity in these two organisms (relative to *S. cerevisiae*) in batch culture. In addition, we will present data on how nutrient uptake, waste excretion, and biomass composition varies across carbon, nitrogen, and phosphorus limitation. Finally, early efforts to computationally integrate these data to achieve experimentally informed quantitative flux maps will be described. The overarching goal is to first quantitatively determine metabolic pathway fluxes across a diversity of conditions, and then to use integrative 'omic approach to determine how these fluxes are controlled[1]. In this manner, we are optimistic that, within a few years, we can bring knowledge of these organisms' metabolic regulation up to a level approaching *S. cerevisiae*, dramatically increasing their value as metabolic engineering platforms.

References

1. Hackett, S. R. et al. Systems-level analysis of mechanisms regulating yeast metabolic flux. Science, 354, (2016).

This work is supported by U. S. Department of Energy Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018260.

Quantifying Metabolic Enzyme Variation in Non-Model Yeasts

Arjuna Subramanian[1,3*] (arjunas@princeton.edu), Meera Gupta[2,3], W. Lance Martin[3], and **Martin Wühr**[1,2,3]

[1]Department of Chemistry, [2]Department of Chemical and Biological Engineering, [3]Department of Molecular Biology & the Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ

https://www.igb.illinois.edu/DOEcenter

Project Goals: Pinpointing the role of variable enzyme expression across different yeast species is an essential aspect of understanding and exploiting yeast metabolism. To this end, we are developing high-throughput proteomics methods for deep quantification of metabolic enzyme abundances in three non-model yeast species (*Issatchenkia orientalis*, *Rhodosporidium toruloides*, and *Yarrowia lipolytica*) of interest for bioproducts synthesis, using *Saccharomyces cerevisiae* as a reference comparison. Analysis of all four yeasts will include experimental measurements of both absolute and relative protein abundances, suitable for integration with metabolomics and transcriptomics data towards complete genome-scale metabolic models for each organism.

The non-model yeasts *Issatchenkia orientalis*, *Rhodosporidium toruloides*, and *Yarrowia lipolytica* have shown substantial promise as microbial factories for synthesis of alcohols, organic acids, fatty acids, and other desirable bioproducts. The expression levels of the enzymes that catalyze reactions relevant to producing such small molecules are essential inputs for understanding, predicting, and engineering metabolic behavior in these organisms. We have adapted proteomics protocols for sample preparation, data collection,[1,2] and data analysis[3,4] to be compatible with the physiology and protein characteristics of these non-standard yeasts. Of particular importance was the development of robust and reproducible lysis methods for yeast cell walls, which are notoriously difficult to disrupt completely. Applying these tools, we confidently quantify over 4700 proteins in *S. cerevisiae*, 4800 proteins in *I. orientalis*, and 5600 proteins in *R. toruloides*, while mapping ~3000 orthologs between *S. cerevisiae* and each non-model species via reciprocal similarity. Using spike-in protein standards and precursor-based MS1 quantification we estimate that intracellular concentrations of individual detectable proteins vary over the ~10 pM-10 μM range, roughly seven orders of magnitude, in all three yeast species under study. Expression levels of orthologous proteins are correlated between each non-model species and *S. cerevisiae* with R^2-values between 0.487 and 0.587; ortholog correlation data is further resolvable at the level of individual metabolic pathways, including glycolysis, the citric acid cycle, and fatty acid metabolism. Additionally, we report shallow tandem mass tag (TMT)-based relative protein quantification of *I. orientalis*, *R. toruloides*, and *S. cerevisiae* in unconstrained, carbon-limited, nitrogen-limited, and phosphorus-limited growth conditions, reaching ~1700 proteins per species, ideal for integration with metabolomics measurements in order to gain a refined understanding of metabolic trends and tradeoffs in these non-standard organisms. The absolute and relative quantification data collected will be used to accelerate

collaborative genome-wide integrative ‘omics metabolic modeling of *I. orientalis*, *R. toruloides*, and *Y. lipolytica*; furthermore, the rigorous proteomics pipeline established may be generalized to more yeast species of interest as they arise.

References

1. McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wühr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P. MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal Chem* **2014**, *86*, 7150-7158. DOI: 10.1021/ac502040v

2. Sonnett, M.; Yeung, E.; Wühr, M. Accurate, sensitive, and precise multiplexed proteomics using the complement reporter ion cluster. *Anal Chem* **2018**, *90* (8), 5032-5039. DOI: 10.1021/acs.analchem.7b04713

3. Wühr, M.; Freeman, R. M.; Presler, M.; Horb, M. E.; Peshkin, L.; Gygi, S. P.; Kirschner, M. W. Deep proteomics of the Xenopus laevis egg using an mRNA-derived reference database. *Curr. Biol.* **2014**, *24*, 1467-1475. DOI: 10.1016/j.cub.2014.05.044

4. Peshkin, L.; Gupta, M.; Ryazanova, L.; Wühr, M. Bayesian confidence intervals for multiplexed proteomics integrate ionstatistics with peptide quantification concordance. *bioRxiV* **2019**. DOI: 10.1101/210476

This work was supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award number DE-SC0018260.

Genome-scale Metabolic Reconstruction of the Non-model Yeast *Issatchenkia orientalis* SD108

Patrick F. Suthers[1]* (pfs13@psu.edu), Siu Hung Joshua Chan[1], Hoang Dinh[1], Yihui Shen[2], Zia Fatma[3], Joshua D. Rabinowitz[2], Costas D. Maranas[1], **Huimin Zhao**[3]

[1]The Pennsylvania State University, University Park, Pennsylvania; [2]Princeton University, Princeton, New Jersey; [3]University of Illinois at Urbana-Champaign, Urbana-Champaign, Illinois

https://www.igb.illinois.edu/DOEcenter

Project Goals: One goal of Genome Scale Engineering is to develop new genome-scale design and engineering tools and implement them on a biofoundry for non-model yeast organisms, including *Issatchenkia orientalis*, to produce high-levels of organic acids. The goal of the current work is to develop a genome-scale metabolic model for *I. orientalis* for use in designing these engineered pathways.

Microorganism can convert renewable biomass into organic acids, many of which are important platform chemicals. However, intolerance to typical biomass hydrolysate's low pH conditions remains a challenge for their industrial production. *Issatchenkia orientalis* is a promising host for industrial production owing to its tolerance of acidic conditions down to pH 2.5. In this work, we describe the development of a genome-scale metabolic model for *I. orientalis* SD108 covering 1,022 genes. The model accounts for reported growth viability on a number of carbon substrates and in-house experimentally measured macromolecular composition specific for this strain. We draw comparisons of the *I. orientalis* SD108 model with recent *Saccharomyces cerevisiae* genome-scale models and examine the network topology for growth on glucose, including proposed essential genes. We demonstrate use of the model by proposing specific genetic interventions for production pathways for the industrially-relevant succinic acid, employing OptKnock to discern modifications which link production of a targeted chemical to biomass production.

This work is supported by U. S. Department of Energy Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018260.

Oxidative Pentose Phosphate Pathway is the Major Cytosolic NADPH Source in *Rhodosporidium toruloides*

Tianxia Xiao[1,2*] (tianxiax@princeton.edu), Christopher Rao,[2,3] and **Joshua Rabinowitz**[1,2]

[1]Princeton University, Princeton, NJ; [2]Carl R. Woese Institute For Genomic Biology

, Urbana, IL; [3]University of Illinois, Urbana-Champaign, Urbana, IL

https://www.igb.illinois.edu/DOEcenter

Project Goals: Understand the metabolic pathways in *R.toruloides*.

Understanding the native metabolism of microbes is important for facilitating metabolic engineering efforts. *Rhodosporidium toruloides* is a promising yeast for fatty-acids production, but relatively understudied. Nitrogen limitation elevates *R. toruloides*' already high native production of fatty acids. Fatty acid production causes a high demand for cytosolic NADPH. Here we investigated the source of cytosolic NADPH in *R. toruloides*. Specifically, we employed a ^{2}H-glucose strategy to trace directly the source of the redox-active hydride of NADPH, an approach that had not previously been applied to any yeast. This strategy requires complementary deuterated water tracing experiments to determine the extent of ^{1}H-^{2}H exchange in NADPH, which otherwise leads to underestimation of pathway contributions. By this approach, we find that the oxidative pentose phosphate pathway (oxPPP) contributes most of *R. toruloides*' cytosolic NADPH. The oxidative pentose phosphate pathway contribution is equivalent in both batch growth and N-limitation. 1,2-^{13}C-glucose tracing data indicate that the ratio of oxidative pentose phosphate pathway flux to glycolysis is also similar in these two conditions. These data suggest a shift during nitrogen limitation in NADPH utilization from reductive nutrient assimilation to fat synthesis. These observations lay foundation for future efforts to enhance fatty acid production in *Rhodosporidium* species via metabolic engineering.

References

1. Zhang, Z., Chen, L. *et al.* Chemical basis for deuterium labeling of fat and NADPH. *JACS,* **139**, 14368-14371 (2017).

Lead PI: Huimin Zhao
PI Organization: University of Illinois at Urbana-Champaign
Program manager: Pablo Rabinowicz
Presenter Organization: Lawrence Berkeley National Laboratory
Poster Abstract Funding Area: Biosystem Design
Abstract file naming convention: Zhao_**Yoshikuni**.pdf

Evaluation of *Issatchenkia orientalis* as a next-generation chassis for bioproduct production from lignocellulosic biomass

Jing Ke[1], Yusuke Sasaki[1], Zhiying Zhao[1], **Yasuo Yoshikuni (yyoshikuni@lbl.gov)**[1,2,3]

[1]US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory
[2]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory
[3]Biological Systems Engineering Division, Lawrence Berkeley National Laboratory

Project Goals: We aim to characterize *Issatchenkia orientalis* as a chassis to produce bioproducts directly from lignocellulosic hydrolysates and identify benchmark strains for metabolic engineering.

Abstract

During pretreatment and hydrolysis of lignocellulosic biomass, many inhibitors are generated as by-products, including organic acids, furan derivatives, and phenolic compounds. These inhibitors negatively affect microbial growth and bioproduct production from lignocellulosic sugars. Conventional industrially relevant (model) microbes such as *E. coli* and *S. cerevisiae* cannot tolerate these inhibitors. *I. orientalis* is a non-model ascomycetes yeast species distinguished by its ability to tolerate multiple stresses. This species may be a promising alternative for producing bioproducts from lignocellulosic substrates. To characterize this species and identify benchmark strains, we collected 160 *I. orientalis* strains from culture collections globally. We are currently evaluating their abilities to tolerate pH and temperature fluctuations and various lignocellulosic inhibitors (e.g., HMF, furfural, acetate, NaCl, and phenolics), as well as their glycolytic capacity, substrate range (e.g., growth on xylose, arabinose, and cellobiose), growth rate, and nutritional requirements (e.g., vitamins and amino acids). The *I. orientalis* strains evaluated thus far generally show promising ability to tolerate low (3.0) pH and various lignocellulosic inhibitors, as well as to produce ethanol from glucose at near quantitative (100%) yield. These strains were more robust when fermentation conditions were semi-aerobic than when they were aerobic, suggesting that that strains use oxygen to reduce inhibitor activity . Most of the strains could catabolize arabinose and cellobiose as carbon sources when glucose was depleted, but they could rarely use xylose.

Funding statement: *This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SCxxxxx). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.*

Feedback Responses of Grassland Microbial Communities to Experimental Warming, Precipitation Alternation, and Clipping

Xue Guo[1], Qun Gao[1], Mengting Yuan[1], Zhou Shi[1], Xishu Zhou[1], **Liyou Wu[1]**, Gangsheng Wang[1], **Zhili He**[1], Joy D. Van Nostrand[1], Lauren Hale[1], Daliang Ning[1], Aifen Zhou[1], Jiajie Feng[1], Ying Fu[1], **Jim Cole**[2], **James M. Tiedje**[2], **Edward A.G. Schuur**[3], **Konstantinos Konstantinidis**[4], **Yiqi Luo**[3], and **Jizhong Zhou**[1*](jzhou@ou.edu)

[1]Institute for Environmental Genomics and Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK; [2]Center for Microbial Ecology, Michigan State University, East Lansing, MI; [3]Center for Ecosystem Sciences and Society, Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ; [4]Center for Bioinformatics and Computational Genomics and School of Biology, Georgia Institute of Technology, Atlanta, GA

http://www.ou.edu/ieg.html

Project goal: The overall goal of this project is to advance systems-level predictive understanding of the feedbacks of belowground microbial communities to multiple climate change factors and their impacts on soil carbon (C) cycling processes. The specific objectives are to: (i) reveal the responses of soil microbial communities to climate warming and soil moisture alteration in both tundra and temperate grassland ecosystems; (ii) determine temperature sensitivity of recalcitrant C decomposition and characteristics of the microbial degraders; and (iii) develop integrated bioinformatics and modeling approaches to scale information across different organizational levels.

As a part of the integrated project, here we present results from; i) field experiments established in 2009 in a temperate grassland of central Oklahoma to reveal the influence of elevated temperature, altered precipitation and plant biomass clipping on long-term succession of plant and microbial communities, and ii) soil respirations and ecosystem C fluxes under long-term warming.

Warming reduces soil microbial biodiversity. To determine successional dynamics of microbial communities in response to warming, clipping, altered precipitation and their combinations, 264 annual soil samples from 2009 to 2016 were analyzed by sequencing of 16S rRNA genes for bacteria and archaea, ITS regions for fungi, and by functional gene arrays (GeoChip 5.0). Our analyses indicated that, experimental warming decreased bacterial and fungal species diversity as well as their phylogenetic diversity, irrespective of the types of clipping and precipitation treatments. Decreases were primarily linked to the decrease of plant functional diversity. By examining the functional traits of plants and microbes, warming also enhanced the dominance of fast-growing resource-acquisitive species in both plant and microbial communities, which could together contribute to the reduced temporal stability of community structure and ecosystem functioning.

Long-term succession of microbial communities. Our analyses indicated that global change factors including warming, clipping, half precipitation, double precipitation and most of their combinations differently shifted the temporal successional patterns of the taxonomic composition and phylogenetic structure of bacteria and fungi. More importantly, among these global change factors, climate warming played a dominant role in accelerating divergent succession of all soil microbial communities as evidenced that experimental warming enhanced microbial temporal divergences under the context of various global changes, which is published in *Nature Climate Change*. Secondly, our results also showed these global change factors and most of their interactions significantly ($P < 0.05$) changed species-time relationships (STRs) of different soil microbial populations including bacteria, fungi and micro-eukaryotes. And climate warming significantly ($P < 0.05$) promotes temporal scaling rates (STR exponent) of all microbial populations even under the context of various global changes, which is presented in our manuscript recently accepted by *Nature Ecology & Evolution*. All of these results indicated that warming plays a dominant role in accelerating temporal succession rates of soil microbial communities.

Network analysis of microbial temporal successions. Soil microorganisms coexist in complex arrays in which interactions among members are essential for community assembly and ecosystem functions. However, most of the studies in the last decades examined the responses of ecological communities to climate changes by just focusing on diversity, but whether and how climate changes affect ecological community organization and the interactions among members of ecological communities, particularly microbial communities, remains elusive. Our network analysis revealed that warming predominantly led to larger and more complex bacterial and fungal networks along time under the context of various global change factors, as indicated that the warmed soil networks significantly increased in size (r^2 = 0.836, P = 0.011) and connectivity (r^2 = 0.916, P = 0.003) over time. Secondly, more and larger modules with more positive and negative links were found in the warmed soil networks, suggesting that more mutualistic and competitive interactions may occur under climate warming. Thirdly, we identified more putative keystone taxa including module hubs, connectors and network hubs in the warmed soil networks. Almost all of these keystone taxa had low relative abundances (0.002% ~ 2.59%), suggesting low-abundance taxa may significantly contribute to soil microbial function. Intriguingly, no network hubs were identified in any of non-warmed soil networks, but one network hub was detected only in the last year of warmed soil network, which were assigned to the typical oligotrophic phylum Acidobacteria and exhibited 91% identity to an isolate of Acidobacteria Gp16. These results indicated that oligotrophic taxa may play more important roles than those copiotrophic taxa in the warmed soil communities.

Incorporating functional genes data into ecosystem modeling. To examine the temperature sensitivity of microbial respiration (Q_{10}), the measured field soil respiration data from individual plots were fitted with the Q_{10}-based Arrhenius equation, which fitted relatively well with the Rh data (r^2 = 0.19-0.99, $p < 0.05$). The Q_{10} estimates were significantly or marginally significantly higher under control than warming in various years. By fitting the respiration data from all 7 years, the overall Q_{10} of heterotrophic soil respiation was also significantly ($p < 0.001$) lower under warming (1.692 ± 0.041) than control (1.947 ± 0.055), suggesting that temperature sensitivity of heterotrophic soil respiation was possibly reduced under warming.

Due to the importance of microbes and their activities in controlling heterotrophic respiration, we further improved the Microbial-ENzyme Decomposition (MEND) model by incorporating GeoChip-detected functional genes information into the ecosystem models. We pooled the functional genes into two categories (i.e., oxidative and hydrolytic enzymes) corresponding to the MEND model. We constrained the model by achieving the highest correlation between MEND-modeled enzyme concentrations and GeoChip-detected enzyme densities in addition to a best fit between modeled and simulated Rh. The MEND mnodel simulated heterotrophic soil respiation agreed well with the observed heterotrophic soil respiation under warming and control. In addition, Also, our model simulation analysis revealed that some key model parameters, such as Q_{10}, intrinsic carbon use efficiency (CUE) and its temperature sensitivity, were better constrained with microbial information than without them. In addition, the MEND simulated oxidative and hydrolytic enzyme concentrations were significantly correlated with the observed gene abundances. The MEND modeling also produced satisfactory results on the enzyme concentrations under warming, as well as the responses of enzymes to warming.

Funding statement: *Our work is supported by the U.S. DOE Office of Science, Biological and Environmental Research Division (BER), Genomic Science Program, Award No. DE - SC0010715.*

BioImaging Abstract Book Information

PRESENTER NAME: Brian O'Callahan
PI NAME: Scott Lea
ORGANIZATION(S): Pacific Northwest National Laboratory
POSTER TITLE: Ultrasensitive vibrational nano-spectroscopy of protein complexes and biological crystals

PRESENTER NAME: Arthur Montazeri
PI NAME: Tina Jeoh & Hoi-Ying Holman
ORGANIZATION(S): University of California at Davis and Lawrence Berkeley National Laboratory
POSTER TITLE: Plasmonic graded gratings for super resolution mid-IR imaging and spectromicroscopy

PRESENTER NAME: Andreas Vasdekis
PI NAME: Andreas Vasdekis
ORGANIZATION(S): University of Idaho, Pacific Northwest National Laboratory
POSTER TITLE: Single-Cell Imaging of Enzyme, Production and Growth Dynamics in Biofuel Producing Microorganisms

PRESENTER NAME: Sang-Hyuk Lee and Shishir Chundawat
PI NAME: Sang-Hyuk Lee
ORGANIZATION(S): Rutgers University, Vanderbilt University, Oak Ridge National Laboratory
POSTER TITLE: In planta multimodal single-molecule imaging to study real-time turnover dynamics of polysaccharides and associated carbohydrate metabolites