U.S. DEPARTMENT OF ENERGY
Office of Science

USDA
United States Department of Agriculture

Genomic Science Contractors–Grantees Meeting XIII

USDA-DOE Plant Feedstock Genomics for Bioenergy Meeting

Abstract Book

February 22-25, 2015

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

This document was produced under contract number DE-AC05-06OR23100 between the U.S. Department of Energy (DOE) and Oak Ridge Associated Universities (ORAU).

The research grants and contracts described in this document are supported by the DOE Office of Science, Office of Biological and Environmental Research, Biological Systems Science Division.

Welcome

Department of Energy
Washington, DC 20585
February 23, 2015

Dear Colleague:

On behalf of the Genomic Science Team as well as the entire Biological Systems Science Division, welcome to the 2015 Genomic Science Annual Contractor-Grantee Meeting/USDA-DOE Plant Feedstock Genomics for Bioenergy Program Meeting! We are pleased to welcome you back this year, and are eagerly looking forward to meeting with you and discussing your latest scientific accomplishments. As the principal meeting for the Genomic Science program, this meeting brings together researchers supported by the program, representatives from the Department of Energy, and colleagues from other Federal Agencies.

I would like to extend a special welcome to our new PIs. This is an excellent opportunity to confer with colleagues, discuss new ideas, and survey the resources available to you within the program. Through discussion with your fellow PIs throughout the meeting and particularly at the poster sessions you will be able to view your research in the context of the larger program and among colleagues engaged in complementary research. We encourage you to take full advantage of this annual opportunity to enlarge your network and potentially develop collaborations with scientists from the broad range of disciplines represented here.

I would also like to extend a welcome to our colleagues from other DOE Offices and Federal Agencies. Please feel free to discuss your program(s) with researchers at this meeting working in similar areas.

The past year has been exciting and productive for the Genomic Science program, and this meeting will highlight research by individual investigators, interdisciplinary research teams, and the Bioenergy Research Centers (BRC). The BRCs are now well into their second 5-year phase of operations and continue to lead extremely productive research, paving the way for the conversion of cellulosic biomass to biofuels. They will report on exciting new results, highlighting the complementarity and integration within and between all three centers in a plenary session as well as the poster sessions. In addition, a report on a recent workshop, "DOE Bioenergy Research" (June 2014) assessing the current state of the science on lignocellulosic biofuels and highlighting research challenges for establishing viable advanced biofuel and bioproduct production systems will be presented.

The DOE Systems Biology Knowledgebase (KBase), a platform to share and integrate data and analytical tools for the acceleration of predictive biology, is now in its fourth year of operations. This year the meeting will feature the "KBase Experience,"an interactive, hands-on workshop where users can"test-drive" the system withtheir own microbial data or by using microbial, plant, and microbial communities data already stored in the system. Members of the KBase team will be present to provide assistance in using KBase on your own laptop. The program meeting also provides an opportunity for researchers to learn about capabilities available at DOE user facilities such as the Joint Genome Institute (JGI), the Environmental Molecular Sciences Laboratory (EMSL), and structural biology resources at DOE's synchrotron light sources and neutron facilities.

Remarkable advances in DNA sequencing technologies now allow the integration of'omics with phenotypic and other data, significantly enhancing the translation of genomics tools to plant breeding research and thus accelerating development of dedicated bioenergy crops.However a critical necessity for implementing such strategiesis the ability to phenotype hundreds to thousands of plants in the field, and current lack of technology to do so is creating a huge bottleneck to this effort. This year we are extremely pleased to have

Dr. Jan Leach, University Distinguished Professor at Colorado State University and Adjunct Scientist at the International Rice Research Institute (IRRI) give the Keynote presentation, who will present her groundbreaking work on high throughput phenotyping in the field and the use of genomics and plant-associated metagenomics to facilitate bioenergy feedstock improvement.

These same technological breakthroughs are allowing systems biology approaches to increase our functional understanding of microbes, including their roles in critical environmental processes such as carbon cycling and facilitating bioengineering of microbes for specific and targeted objectives. A plenary session on Microbial Communities and Biogeochemical Cycling as well as a breakout session focused on microbial systems biology will be held here this year.

As has become tradition in alternate years, this meeting once again convenes researchers supported by the Genomics Science Program and by the joint USDA-DOE Plant Feedstock Genomics for Bioenergy program, which focuses on genomics-based research leading to the improved use of biomass and plant feedstocks for producing biofuels. Bringing together these distinct but complementary research communities stimulates the exchange of ideas, sharing of expertise, and formation of new partnerships. Both a plenary session and a breakout session will be held to discuss some of the very important developments in feedstock genomics research that have taken place since these groups last met, and to highlight the most recent awards. In addition to covering research on a wide spectrum of bioenergy feedstocks, this year's breakout session will include an overview of the USDA NIFA AFRI Regional Bioenergy/Feedstock Systems Coordinated Agricultural Projects (CAP).

This year's breakout sessions will cover a broad range of BER mission-relevant topics, from fungal and microbial biology to computational biology and the application of new, advanced technologies within the various Genomic Science Program areas. A draft report on a recent workshop (May 2014) entitled "Molecular Science Challenges" will be presented, highlighting the scientific challenges and opportunities to understanding molecular processes underlying BER program goals, and how resources available at the DOE National Labs and Facilities can beleveraged to effectively tackle these challenges. Finally, we are proud to present a plenary session featuring the most recent recipients of the Office of Science Early Career program awards.

We look forward to an exciting and productive meeting and encourage you to exchange ideas and share your expertise with other researchers in discussions during the plenary presentations, breakout meetings, and poster sessions. We thank you for lending your knowledge, creativity, and vision to the Genomic Science program and wish you continued success in the coming year.

Sincerely,

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

Agenda

Genomic Science Program
2015 Contractor-Grantee Meeting
February 22-25, 2015

SUNDAY, FEBRUARY 22ND

5:00 PM – 8:00 PM	**Early Registration and Poster Setup**

MONDAY, FEBRUARY 23RD

7:00 AM – 8:30 AM	**Registration** Location: Fairfax Foyer
7:30 AM – 8:30 AM	**BREAKFAST ON YOUR OWN**
8:30 AM – 9:00 AM	**Welcome and Introduction to the Meeting** Location: Fairfax Ballroom
8:30 AM – 8:40 AM	**Sharlene Weatherwax**, Associate Director, DOE Office of Biological and Environmental Research Opening Remarks
8:40 AM – 8:50 AM	**Todd Anderson**, Director, Biological Systems Science Division, DOE BER Meeting Introduction
8:50 AM – 9:00 AM	**Joseph Graber**, Program Manager, DOE BER Genomic Science Program: Strategic Plan
9:00 AM – 10:30 AM	**Plenary Session: DOE Bioenergy Research Centers** Location: Fairfax Ballroom Moderator: Kent Peters
	Speakers:
9:00 AM – 9:30 AM	Blake Simmons, JBEI, "Driving the future: Advances in biomass to biofuel conversion enabled by integrated research and development at JBEI"
9:30 AM – 10:00 AM	Tim Donohue, GLBRC, "From fields to fuels & chemicals"
10:00 AM – 10:30 AM	Rick Dixon, BESC, "Progress in overcoming biomass recalcitrance in lignocellulosic feedstocks for biofuel production"
10:30 AM – 11:00 AM	**BREAK**
11:00 AM – 12:00 PM	**Keynote Presentation: Jan Leach, Colorado State University, "Field-scale high throughput phenotyping for biomass QTL"** Location: Fairfax Ballroom Moderator: Cathy Ronning
12:00 PM – 2:00 PM	**LUNCH ON YOUR OWN**

MONDAY, FEBRUARY 23RD

1:00 PM – 7:00 PM

KBase Experience Hands-on Session
Location: Wolftrap

Description of Session:
This session is intended to give you and your team an opportunity to test the latest tools available in KBase in an interactive setting. You can bring your own microbial data to analyze, or test KBase using microbial, plant, and microbial communities data already stored in the system. Members of the KBase team will be there to help you start using KBase on your laptop. To best prepare for this opportunity please sign up here: https://docs.google.com/forms/d/1qf-ClYh2BwLtIPav8FZfnZjLnhBvpATYP0Ts99Qd25w/viewform?c=0&w=1 or e-mail Meghan Drake (drakemm@ornl.gov) for more information.

2:00 PM – 5:00 PM

Breakout Session A: Microbial Systems Biology
Location: Fairfax A
Moderator: Joseph Graber

Description of Session:
New breakthroughs in our ability to dissect and predict integrated systems biology properties of microbes has significantly advanced our functional understanding of these organisms, their roles in higher scale processes, and our capacity to bioengineer novel capabilities. As we continue to refine our understanding of key model systems like E. coli, rapid progress is also being made in a broad suite of other organisms from across the tree of life, expanding our knowledge of diverse metabolic pathways and functional capabilities. In this session, researchers will present new findings on the systems biology on a diverse range of microbes and discuss how this work contributes to DOE missions in energy and the environment.

Speakers:

2:00 PM – 2:30 PM	Jim Liao, University of California, Los Angeles, “Ensemble Modeling for Pathway Construction”
2:30 PM – 3:00 PM	Aindrila Mukhopadhyay, Lawrence Berkeley National Laboratory, “Engineering Solvent Tolerant Bacteria: A Systems Biology Guided Effort”
3:00 PM – 3:30 PM	**BREAK**
3:30 PM – 4:00 PM	Himadri Pakrasi, Washington University in St. Louis, “Systems Level Analysis of the Photoautotrophic Lifestyle of a Cyanobacterium”
4:00 PM – 4:30 PM	Janelle Thompson, Massachusetts Institute of Technology, “Harnessing Microbes From the Deep Carbonated Biosphere”
4:30 PM – 5:00 PM	Yongqin Jiao, Lawrence Livermore National Laboratory, “Functional Understanding of Caulobacter crescentus for Uranium Detoxification and Mineralization”

2:00 PM – 5:00 PM

Breakout Session B: USDA-DOE Plant Feedstocks Genomics for Bioenergy
Location: Fairfax B
Moderator: Cathy Ronning

MONDAY, FEBRUARY 23RD

Description of Session:
The joint USDA-DOE Plant Feedstocks Genomics for Bioenergy program supports fundamental genomics-based research leading to the development of improved and sustainable plant feedstocks for the production of biofuels. The session will be introduced by Dr. Bill Goldner of USDA-NIFA with an overview of the AFRI Regional Bioenergy/Feedstock Systems Coordinated Agricultural Projects (CAPs), followed by presentations of recent accomplishments in poplar, switchgrass, sorghum, and willow research by the 2012 awardees. The 2014 awardees will follow with a series of "speed talks" highlighting newly funded research spanning a wide variety of candidate bioenergy feedstocks, including oilseed crops.

Speakers:

2:00 PM – 2:20 PM	Bill Goldner (USDA NIFA), "NIFA AFRI Regional Bioenergy/Feedstock Systems Coordinated Agricultural Projects – an Overview"
2012 Awardees:	
2:20 PM – 2:35 PM	Victor Busov, Michigan Technological University, "Discovery of genes affecting wood biomass yield and quality in poplar using activation tagging"
2:35 PM – 2:50 PM	Pankaj Jaiswal, Oregon State University, "Abiotic stress induced Poplar transcriptome analyses"
2:50 PM – 3:05 PM	Larry Smart, Cornell University, "Transcriptome analysis of diploid and triploid species hybrids of shrub willow"
3:05 PM – 3:20 PM	Andrew Paterson, University of Georgia, "Genomics of Bioenergy Grass Architecture"
3:20 PM – 3:40 PM	**BREAK**
3:40 PM – 3:55 PM	Tom Juenger, University of Texas, Austin, "Panicum hallii: a genomics enabled model for C4 perennial grass biology"
3:55 PM – 4:10 PM	Malay Saha, Noble Foundation, "Deciphering Natural Allelic Variation in Switchgrass for Biomass Yield and Quality Using a Nested Association Mapping Population"
2014 Awardees:	
4:10 PM – 4:15 PM	Todd Mockler, Donald Danforth Plant Science Center, "The Brachypodium ENCODE Project – From Sequence to Function: Predicting Physiological Responses in Grasses to Facilitate Engineering of Biofuel Crops"
4:15 PM – 4:20 PM	Maria Harrison, Boyce Thompson Institute, "Genetic Dissection of AM Symbiosis to Improve the Sustainability of Feedstock Production"
4:20 PM – 4:25 PM	Patrick Brown, University of Illinois, "Coordinated Genetic Improvement of Bioenergy Sorghum for Compositional and Agronomic Traits"
4:25 PM – 4:30 PM	John Mullet, Texas A&M University, "Genomics of Energy Sorghum's Water Use Efficiency / Drought Resilience"
4:30 PM – 4:35 PM	Erik Sacks, University of Illinois, "Quantifying Phenotypic and Genetic Diversity of Miscanthus sacchariflorus to Facilitate Knowledge of Directed Improvement of M.xgiganteus (M. sinensis x M. sacchariflorus) and Sugarcane

MONDAY, FEBRUARY 23RD

	4:35 PM – 4:40 PM	Robin Buell, Michigan State University, "Exploiting Natural Diversity to Identify Alleles and Mechanisms of Cold Adaptation in Switchgrass"
	4:40 PM – 4:45 PM	M. David Marks, University of Minnesota, "Advancing Field Pennycress as a New Oilseed Biodiesel Feedstock That Does Not Require New Land Commitments"
	4:45 PM – 4:50 PM	John McKay, Colorado State University, "Biofuels in the Arid West: Germplasm Development for Sustainable Production of Camelina Oilseed"
	4:50 PM – 4:55 PM	Amy Brunner, Virginia Polytechnic Institute and State University, "Abiotic Stress Networks Converging on FT2 to Control Growth in Populus"
	4:55 PM – 5:00 PM	Isabelle Henry (for Luca Comai), University of California, Davis, "A Novel Poplar Biomass Germplasm Resource for Functional Genomics and Breeding"

2:00 PM – 5:00 PM

Breakout Session C: Application of Advanced Technologies
Location: Ash Grove Ballroom AB
Moderator: Roland Hirsch

Description of Session:
Much research in the Genomic Science Program depends on the use of new technologies. This session includes presentations by GSP scientists on how they are developing and using new technologies for analysis and imaging of complex biological systems, including capabilities at DOE National User Facilities.

Speakers:

2:00 PM – 2:05 PM	Introductory Comments
2:05 PM – 2:25 PM	Robert Blake, Xavier University, Louisiana, "A Spectroscopic Device to Monitor Respiratory Electron Transfer in Suspensions of Live Organisms"
2:25 PM – 2:45 PM	Haw Yang, Princeton University, "3D multi-resolution imaging: development and applications"
3:45 PM – 3:05 PM	Andrzej Joachimiak, Argonne National Laboratory, "Crystallography at light sources enables progress in microbiology and plant science"
3:05 PM – 3:25 PM	Rhona Stuart, Lawrence Livermore National Laboratory, "Elucidating cyanobacterial recycling of microbial mat extracellular matrix"
3:25 PM – 3:40 PM	**BREAK**
3:40 PM – 4:00 PM	Hugh O'Neill, Oak Ridge National Laboratory, "Applications of Neutron Scattering in Studies of Complex Biological Systems"
4:00 PM – 4:20 PM	Paul Bohn, University of Notre Dame, "Correlated Confocal Raman and Mass Spectrometric Imaging - Spatiotemporal Signaling Relationships in Microbial Communities and Inter-kingdom Assemblies in the Rhizosphere"
4:20 PM – 4:50 PM	Shi-You Ding, Michigan State University, "Studying lipid accumulation mechanism in oleaginous yeast using hyperspectral SRS microscopy and RNA-Seq in single cells"

MONDAY, FEBRUARY 23RD

5:00 PM – 7:00 PM **Poster Session (odd-numbered posters)**
Location: Tysons Ballroom

TUESDAY, FEBRUARY 24TH

7:30 AM – 8:30 AM **BREAKFAST ON YOUR OWN**

8:30 AM – 10:00 AM **Plenary Session: Microbial Communities and the Terrestrial Carbon Cycle**
Location: Fairfax Ballroom
Moderator: Joe Graber

Speakers:

8:30 AM – 9:00 AM	Scott Saleska, University of Arizona, "From Archaea to the Atmosphere: How Genome-Scale Resolution Can Improve Earth System Models of Climate Change"
9:00 AM – 9:30 AM	Mary Lidstrom, University of Washington, "Methane Conversion as a Community Function"
9:30 AM – 10:00 AM	Dan Buckley, Cornell University, "A peek into the secret lives of soil microbes as revealed by high resolution stable isotope probing"

10:00 AM – 10:30 AM **BREAK**

10:30 AM – 12:00 PM **Plenary Session: USDA-DOE Plant Feedstocks Genomics for Bioenergy**
Location: Fairfax Ballroom
Moderator: Bill Goldner/Ed Kaleikau (USDA-NIFA)

Speakers:

10:30 AM – 11:00 AM	Kevin Childs, Michigan State University, "Gene Expression Differences in Tolerant and Sensitive Switchgrass during Abiotic Stress"
11:00 AM – 11:30 AM	Patrick Schnable, Iowa State University, "High-throughput, Field-based Phenotyping of Biomass Sorghum"
11:30 AM – 12:00 PM	Chung-Jui Tsai, University of Georgia, "Dual Effect of Tubulin Manipulation on Populus Wood Formation and Stomatal Behavior"

12:00 PM – 2:00 PM **LUNCH ON YOUR OWN**

1:00 PM – 7:00 PM **KBase Experience Hands-on Session**
Location: Wolf Trap

Description of Session:
This session is intended to give you and your team an opportunity to test the latest tools available in KBase in an interactive setting. You can bring your own microbial data to analyze, or test KBase using microbial, plant, and microbial communities data already stored in the system. Members of the KBase team will be there to help you start using KBase on your laptop. To best prepare for this opportunity please sign up here: https://docs.google.com/forms/d/1qf-ClYh2BwLtIPav8FZfnZjLnhBvpATYP0Ts99Qd25w/viewform?c=0&w=1 or e-mail Meghan Drake (drakemm@ornl.gov) for more information.

TUESDAY, FEBRUARY 24TH

2:00 PM – 5:00 PM

Breakout Session D: Computational Biology: Data Analysis and Modeling in the Genomic Sciences
Location: Fairfax A
Moderator: Pablo Rabinowicz

Description of Session:
The ever-increasing capacity of high throughput genomics technologies has transformed experimental biology, allowing researchers to conduct large-scale studies of biological systems at the whole-organism or even ecosystem level. These studies generate vast amounts of data and information that can only be organized and analyzed using powerful bioinformatics tools and resources. Computational biology approaches to predict the behavior and model biological systems are therefore an essential part of systems biology research. This breakout will focus on recent developments in bioinformatics and computational biology for data analysis and modeling. Some of the tools that will be presented in this session may become integrated within the DOE Systems Biology Knowledgebase (KBase).

Speakers:

2:00 PM – 2:30 PM	Daniel Segrè, Boston University, "Spatio-temporal models of engineered microbial communities"
2:30 PM – 3:00 PM	Brian Thomas, University of California, Berkeley, "Genomes from Metagenomes: rapid, accurate binning and metabolic analysis of complex metagenomic samples"
3:00 PM – 3:30 PM	Asa Ben-Hur, Colorado State University, "Differential splicing and long read transcriptome assembly using iDiffIR and TAPIS"
3:30 PM – 4:00 PM	**BREAK**
4:00 PM – 4:30 PM	Ronan Fleming, University of Luxembourg, "Multi-scale modelling of stationary biochemical network dynamics"
4:30 PM – 5:00 PM	Ivan Baxter, USD-ARS, Donald Danforth Plant Science Center, "High-throughput phenomic approaches to understanding drought adaptation in the C4 model Setaria"

2:00 PM – 5:00 PM

Breakout Session E: Fungal Biology
Location: Fairfax B
Moderator: Kent Peters

Speakers:

2:00 PM – 2:30 PM	Michelle O'Malley, University of California, Santa Barbara, "New Enzymes from Old Microbes: Exploiting Anaerobic Gut Fungi for Biomass Breakdown"
2:30 PM – 3:00 PM	Cheryl Kuske, Los Alamos National Laboratory, "Soil metatranscriptomes and rRNA surveys illustrate fungal community composition and responses to N deposition in forests"
3:00 PM – 3:30 PM	**BREAK**

TUESDAY, FEBRUARY 24TH

	3:30 PM – 4:00 PM	Jessy Labbé, Oak Ridge National Laboratory, "Investigating the fungal-bacterial interactions within the Populus microbiome"
	4:00 PM – 4:30 PM	Maria Harrison, Boyce Thompson Institute, "Insights into AM fungi through transcriptome analyses of Brachypodium distachyon AM symbioses"
	4:30 PM – 5:00 PM	Greg Stephanopoulos, Massachusetts Institute of Technology, Engineering organisms and processes for cost-effective lipid production"

5:00 PM – 7:00 PM **Poster Session (even-numbered posters)**
Location: Tysons Ballroom

WEDNESDAY, FEBRUARY 25TH

8:00 AM – 8:30 AM **BREAKFAST ON YOUR OWN**

8:30 AM – 8:50 AM **DOE Bioenergy Research Workshop Report**
Kent Peters, Program Manager, DOE-BER
Location: Fairfax Ballroom

8:50 AM – 9:10 AM **Molecular Science Challenges Workshop Report**
Roland Hirsch, Program Manager, DOE-BER
Location: Fairfax Ballroom

9:10 AM – 9:50 AM **DOE Early Career Research Awards**
Location: Fairfax Ballroom
Moderator: Pablo Rabinowicz

9:10 AM – 9:30 AM	Jennifer Pett-Ridge, LLNL, "Microbial carbon transformations in wet tropical soils: the importance of redox fluctuations"
9:30 AM – 9:50 AM	Trent Northen, LBNL, "Understanding microbial carbon cycling in soils using novel metabolomics approaches"

9:50 AM – 10:50 AM **Plenary Session: DOE User Facilities & Community Resources**
Location: Fairfax Ballroom
Moderator: Roland Hirsch

Speakers:

9:50 AM – 10:10 AM	Jeremy Schmutz, HudsonAlpha, "JGI Capabilities Enabling Biological Science"
10:10 AM – 10:30 AM	Ljiljana Paša-Tolić, Pacific Northwest National Lab, "EMSL Capabilities Enabling Biological Science"
10:30 AM – 10:50 AM	Adam Arkin, Lawrence Berkeley National Lab, "DOE Systems Biology Knowledgebase (KBase)"

11:00 AM **CLOSE–OUT AND ADJOURNMENT**

Table of Contents

1. BESC Switchgrass and Populus TOP Lines: Identifying and Analyzing the Best Candidate Lines for Enhanced Biofuel Production

Richard S. Nelson[1,2*] (rsnelson@noble.org), Jay Chen,[1,6] Erica Gjersing,[1,7] Susan Holladay,[1,6] Amy Mason,[1,2] Mitra Mazarei,[1,3] Luis Escamilla-Trevino,[1,4] Ajay Biswal,1,5 Guifen Li,[1,2] Desalegn Serba,[1,2] Hui Shen,[1,4] Avinash Srivastava,[1,2] Anthony Bryan,[1,6] Lee Gunter,[1,6] Yongil Yang,[1,6] Alan Darvill,[1,5] Jerry Tuskan,[1,6] Malay Saha,[1,2] Tim Tschaplinski,[1,6] Steve DiFazio,[1,8] Wellington Muchero,[1,6] Debra Mohnen,[1,5] Yuhong Tang,1,2 Richard A. Dixon,[1,4] Zeng Yu Wang,[1,2] Neal Stewart,[1,3] Brian Davison,[1,6] and Paul Gilna[1,6]

[1]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [2]The Samuel Roberts Noble Foundation, Ardmore, Oklahoma; [3]University of Tennessee, Knoxville; [4]University of North Texas, Denton; [5]University of Georgia, Athens; [6]Oak Ridge National Laboratory, Oak Ridge, Tennessee; [7]National Renewable Energy Laboratory, Golden, Colorado; [8]West Virginia University, Morgantown

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step (consolidated bioprocessing). BESC biomass formation and modification research involves working directly with two potential bioenergy crops (switchgrass and Populus) to develop varieties that are easier to break down into fermentable sugars. We are using both testing and generating of large numbers of natural and modified plant samples as well as developing genomics tools for detailed studies into poorly understood cell wall biosynthesis pathways.

Two goals of the BioEnergy Science Center (BESC) are to develop switchgrass and Populus lines with reduced recalcitrance and/or higher total cell wall sugar content and to understand the molecular and whole plant traits associated with this enhanced biofuel potential. BESC researchers have identified and vetted over 600 genes as targets to reduce recalcitrance and/or enhance total cell wall sugar content through transient or stable expression studies with switchgrass, Populus or model species. Switchgrass and Populus natural variants also were targeted for recalcitrance and cell wall sugar content evaluation. In total, thousands of transgenic and natural variant lines have been analyzed for enhanced biofuel potential and a process was developed to identify among these lines those with the greatest biofuel production potential (i.e., to identify TOP Lines). A TOP Line Selection Committee with experience in cell wall biochemistry, functional genomics, plant growth and development, and applied research evaluated the TOP Line candidates. A subset of lines with the best reduction in recalcitrance or enhanced cell wall sugar content while maintaining or enhancing growth traits were selected for detailed growth and wall analyses during large structured greenhouse and field trials. In addition to identifying high value biomass, the goal of this research is to identify a core set of biomass traits indicative of reduced recalcitrance (i.e., having increased biofuel potential). Teams of researchers were identified to grow, observe and harvest the lines and to evaluate TOP Line tissue using a variety of analytical approaches, with the goals of characterizing the growth morphology of each line under different conditions and determining the mechanisms of reduced recalcitrance. Specifically, a standardized protocol was developed for simultaneous growth of multiple TOP and comparator lines at the same location to enable meaningful comparisons between lines within the same growth experiment. The inclusion of internal reduced-recalcitrance control lines within each growth experiment allows normalization of the TOP Line findings between experiments. Comparisons across lines will allow researchers to identify, among the TOP Lines, those with the greatest potential to deliver improved biofuel production and to identify the number and

type of different "recalcitrance" mechanisms that impact production of biofuel from plant biomass. Data on TOP Line growth, biochemical and chemical characteristics are being incorporated into a Laboratory Information Management System (LIMS) to enable comparisons within and across lines. The entire process for the identification, growth and analysis of the BESC TOP lines is presented as a summary of the status of evaluation of some of the TOP Lines.

The BioEnergy Science Center 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.

2. Biological Lignocellulose Solubilization: Comparative Evaluation of Feedstock-Biocatalyst Combinations and Enhancement via Co-Treatment by Intermediate Milling

Julie Paye[1,2*] (JPaye@Dartmouth.edu), Anna Guseva,[1,2] Sarah K. Hammer,[1,2] Bryon S. Donohoe,[2,3] Erica Gjersing,[2,3] Jessica Olstad,[2,3] Mark F. Davis,[2,3] Brian Davison,[2] Lee Lynd,[1,2] and Paul Gilna[2]

[1]Dartmouth College, Hanover, New Hampshire; [2]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [3]National Renewable Energy Laboratory, Golden, Colorado

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step [consolidated bioprocessing (CBP)]. BESC research in biomass deconstruction and conversion targets CBP by studying model organisms and thermophilic anaerobes to understand novel strategies and enzyme complexes for biomass deconstruction.

We conducted comparative studies under controlled, but not industrial, conditions to inform how choice of feedstock, biocatalyst, and fermentation conditions impacts the ability to overcome recalcitrance. Several anaerobic bacteria are able to achieve substantial solubilization of minimally pretreated switchgrass, with solubilization yields up to two times higher than commercial fungal cellulase. Performance of fungal enzymes was not significantly improved by addition of yeast, higher enzyme loadings, increased hydrolysis temperature, or lower substrate loadings. For both Clostridium thermocellum and fungal cellulase, conversion is twice as high for switchgrass harvested at mid-season as compared to late season.

However, solubilization of late-season switchgrass can be improved by mechanical disruption of the substrate after biological attack has begun (as done in rumen and by termites). Such "co-treatment" may be contrasted to pretreatment, for which the lignocellulose matrix is disrupted prior to biological attack. Brief (5 minute) ball milling of solids remaining after fermentation of senescent switchgrass by C. thermocellum increased carbohydrate solubilization upon reinoculation as compared to a control without milling (68±2% and 41±2%, respectively). When the same mechanical disruption was applied as a pretreatment before the first fermentation stage, carbohydrate solubilization was 55±2%. Residual solids were characterized in order to gain insight into the mechanisms underlying the effectiveness of milling. The main effect of milling observed was decreased particle size, with final d50 (volume-mean particle size) values of 411 ± 42 um for no milling, 89 ± 6 um milling before the first fermentation, and 29 ± 2 um milling between the first and second fermentations. Milling had little effect on substrate cellulose crystallinity but pore volume increased with milling and fermentation. The substantial impact of brief milling observed here supports the possibility of mechanical disruption in a side vessel much smaller than the fermenter.

The BioEnergy Science Center 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.

3. Biomass Deconstruction by Members of the Genus Caldicellulosiruptor

Jenna Young,[1,5] Daehwan Chung,[1,5] Michael G. Hahn,[2,5] Debra Mohnen,[3,5] Yannick J. Bomble,[4,5] Michael E. Himmel,[4,5] Janet Westpheling[1,5*] (Janwest@uga.edu), and Paul Gilna[5]

[1]Department of Genetics, University of Georgia, Athens; [2]Department of Plant Biology, Complex Carbohydrate Research Center, University of Georgia, Athens; [3]Department of Biochemistry and Molecular Biology, Complex Carbohydrate Research Center, University of Georgia, Athens; [4]National Renewable Energy Laboratory, Golden, Colorado; [5]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step [consolidated bioprocessing (CBP)]. BESC research in biomass deconstruction and conversion targets CBP by studying model organisms and thermophilic anaerobes to understand novel strategies and enzyme complexes for biomass deconstruction.

A major obstacle, and perhaps the most important economic barrier to the effective use of plant biomass for the production of fuels, chemicals and bioproducts, is our current lack of knowledge of how to efficiently and effectively deconstruct plant cell wall polymers for their subsequent use as feedstocks. Plants represent a desired source of renewable energy and hydrocarbons because they fix CO2, making their use in such processes carbon-neutral. Plant structure, however, is a barrier to deconstruction and this characteristic is often referred to as recalcitrance. Members of the bacterial genus Caldicellulosiruptor are the most thermophilic cellulolytic organisms so far described and have the ability to grow on lignocellulosic biomass without conventional pretreatment. Here, we describe two studies using deletions in Caldicellulosiruptor bescii to examine biomass hydrolysis.

Different species vary in their ability to degrade cellulose, and the presence of CelA, a bi-functional glycoside hydrolase correlates well with cellulolytic ability in members of this genus. C. bescii, which produces CelA and expresses it constitutively, is among the most cellulolytic. In fact, CelA is the most abundant extracellular protein produced in C. bescii. In contrast, C. hydrothermalis, which does not contain a CelA homolog, or members of the common bacterial cellulases (GH48 or GH9 glycoside hydrolases), is the least cellulolytic of the Caldicellulosiruptor species so far described.

The enzyme contains two catalytic units, a Family 9A-CBM3c processive endoglucanase and a Family 48 exoglucanase joined by two Family 3b carbohydrate binding domains. While there are two GH9 and three GH48 glycoside hydrolases in C. bescii, CelA is the only protein that combines both enzymatic activities. A deletion of the celA gene in C. bescii resulted in a dramatic reduction in the microorganism's ability to grow on crystalline cellulose (Avicel) and diminished growth on lignocellulosic biomass. A comparison of the overall endoglucanase and exoglucanase activities of the mutant compared to the wild type suggests that the loss of the endoglucanase activity provided by the GH9 domain is perhaps compensated by other enzymes produced by the cell. In contrast, enzymatic activity on Avicel by the mutant strain resulted in a 15-fold decrease in sugar release compared to the parent and wild type strains. Taken together, these data suggest that the exoglucanase activity of the GH48 domain of CelA plays a major role in biomass degradation within the suite of C. bescii biomass degrading enzymes.

The combination of microbial digestion and plant biomass analysis provides an important platform to

identify plant wall structures whose presence reduces the ability of microbes to deconstruct plant walls and to identify enzymes that specifically deconstruct those structures. A deletion of a gene cluster encoding enzymes involved in pectin degradation was constructed in C. bescii and the resulting mutant was reduced in its ability to grow on both dicot and grass biomass, but not on soluble sugars. We believe that saccharification of plant biomass can be improved by modifying the structure of pectin. Most efforts targeting improved deconstruction of plant biomass for biofuel production have focused on crystalline cellulose, lignin, and xylan—the major hemicellulose in grass (e.g., switchgrass) walls and in dicot (e.g., Populus wood) secondary walls. Indeed, most models of biomass used in the biofuels field do not list pectin because of its low abundance in grass walls and in dicot secondary walls. Recent work, however, has shown that pectin is synthesized in secondary walls and that some pectin biosynthetic enzymes are amplified in grasses. Plant biomass from three phylogenetically diverse plants, Arabidopsis (an herbaceous dicot), switchgrass (a monocot grass), and Populus (a woody dicot) were used in the analysis. These biomass types have cell walls that are significantly different from each other in both structure and composition. Glycome profiling, using a large and diverse set of plant glycan-directed monoclonal antibodies, of the biomass remaining after growth of the bacterial mutant compared to the wild type revealed differences in the way the mutant utilizes these plants for growth. While pectin is a relatively minor component of the grass and woody dicot substrates, these analyses provide direct evidence that pectin plays an important role in the recalcitrance of all three types of biomass.

The BioEnergy Science Center 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.

4. Building the Cell Wall: Insights Into Xylan Synthesis Using Recombinant Enzymes

Breeanna Urbanowicz[1*] (breeanna@uga.edu), Maria Peña,[1] Jason Backe,[1] Malcolm O'Neill,[1] Heather Moniz,[1] Shuo Wang,[1] Kelley Moremen,[1] William York,[1] and Paul Gilna[2]

[1]University of Georgia, Department of Plant Biology, Complex Carbonhydrate Research Center, Athens; [2]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step (consolidated bioprocessing). BESC biomass formation and modification research involves working directly with two potential bioenergy crops (switchgrass and Populus) to develop varieties that are easier to break down into fermentable sugars. We are using both testing and generating of large numbers of natural and modified plant samples as well as developing genomics tools for detailed studies into poorly understood cell wall biosynthesis pathways.

Xylans are the dominant hemicellulosic polysaccharide found in the plant kingdom, second only to cellulose in abundance, and are present in load-bearing secondary cell walls of dicots and in both primary and secondary cell walls of grasses and cereals. These highly acetylated cell wall polysaccharides are a vital component of the plant cell wall that functions as a molecular scaffold, providing plants with mechanical strength and flexibility. Due to its abundance and complex interactions with cellulose and lignin, xylan has profound effects on the recalcitrance of biomass to saccharification, making it a prime target for genetic manipulation. However, the mechanisms by which plants synthesize xylan are poorly understood. A major impediment to the structural and functional characterization of plant carbohydrate active enzymes (CAZYmes) has been the lack of highly purified and active enzymes. We have recently overcome these limitations by carrying out heterologous expression of plant CAZYmes in Human Embryonic Kidney cells. This breakthrough technology has allowed us to successfully express and purify several different enzymes involved in hemicellulose biosynthetic pathways. Our ability to consistently obtain high-level expression of functional plant glycosyltransferases is a major milestone for BESC. We combined purified, recombinant CAZYmes with high-throughput glycosyl-, methyl- and acetyltransferase assays to investigate their biochemical activities. Using this approach, we have demonstrated that Arabidopsis thaliana IRREGULAR XYLEM 10-L has UDP-Xyl: β-(1,4)-xylosyl transferase activity and is a xylan synthase that elongates the xylan backbone. Further, we show that ESKIMO1/TRICOME BIREFRINGENCE 29 (TBL29/ESK1) catalyzes the subsequent addition of O-acetyl groups from acetyl-CoA to the 2-position of xylosyl backbone residues, making it the only plant polysaccharide O-acetyl transferase that has been biochemically characterized to date. These two enzymes can be used in combination to synthesize and O-acetylate xylan in vitro. These advances represent a significant achievement for the field of cell wall polysaccharide biosynthesis, but also provide the plant science community with a new and powerful tool-box to produce recombinant proteins for biochemical analysis

The BioEnergy Science Center 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.

5. Characterizing Cell Wall Crosslinking and/or Entanglement in Reduced Recalcitrance Feedstocks

Erica Gjersing[1*] (Erica.gjersing.nrel.gov), Bryon Donohoe,[2] William Rottmann,[3] Maud Hinchee,[3] Robert Sykes,[1] Geoffrey B. Turner,[2] Rui Katahira,[1] Angela Ziebell,[1] David Johnson,[2] Stephen Decker,[2] Mark Davis,[1] and Paul Gilna[4]

[1]National Bioenergy Center, National Renewable Energy Laboratory, Golden, Colorado; [2]Biosciences Center, National Renewable Energy Laboratory, Golden, Colorado; [3]ArborGen, Inc., Summerville, South Carolina; [4]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step (consolidated bioprocessing). BESC researchers provide enabling technologies in characterization, ′□omics, modelin
in order to understand chemical and structural changes within biomass and (2) to provide insights into biomass formation and conversion.

The cell wall of plants is composed up of a variety of biopolymers including lignin, hemicellulose, pectin, and cellulose. When these polymers are deconstructed to monomeric sugars, they can be utilized to produce fuels and chemicals. However, this process of breaking down these polymeric components to their monomers, can be difficult due to the recalcitrant nature of lignocellulosic material. Understanding how the properties of the biopolymer matrix, such as crosslink density and/or entanglement, are affecting recalcitrance is vital to engineering plants for improved biofuels production and performance.
Lignin has been shown to play a major role in recalcitrance.1,2 Therefore, plant systems with changes in their lignin content and structure have been selected for these studies. Specifically, we have studied Eucalyptus urophylla x grandis antisense downregulated in p-coumarate 3'-hydroxylase (C3H) or cinnamate-4-hydroxylase (C4H). These transgenic lines have been identified with the properties of low recalcitrance, low lignin and improved biofuel processing efficiency. In addition, both the C3H and C4H downregulated lines were found to have significantly higher carbohydrate extractability when exposed to NaOH base extraction, indicating altered cell wall construction. The molecular weight of isolated lignin and lignin structure determined by HSQC NMR based lignin subunit analysis for control and the C3H and C4H downregulated lines were also examined. The slight reductions in weight average molecular weights of the lignins isolated from the transgenic lines does not appear to explain the difference in extractability while the HSQC NMR based lignin subunit analysis showed only slight differences in the lignin structure of the C3H and C4H downregulated lines when compared to the control.

To further investigate the role of the lowered lignin content on the reduced recalcitrance and increased extractability of the transgenic lines, we performed solvent swelling experiments and T2 NMR relaxation measurements. Swelling samples in water and measuring the difference between the wet and dry volume is an experiment that has been used to measure the dimensional stability of wood samples. For the eucalyptus samples, the control had a 14% volume change (ΔV), compared to the C4H transgenic lines with a 54% ΔV and the C3H line with 31%ΔV. These results indicate that the transgenic lines have increased mobility in their cell wall matrix that can expand to take up more water than the controls. To further investigate changes in the biopolymer matrix mobility, T2 NMR relaxation measurements were performed. These experiments are commonly used to measure mobility in synthetic polymer systems where an increase in T2 relaxation time indicates an increase in the mobility of the protons within the

polymer matrix and has been correlated to decreases in crosslink density and/or entanglement. For the C4H transgenic, there was a 68% increase in the T2 relaxation time while the C3H line displayed a 34% increase compared to the control. These experiments indicate that the reduced recalcitrance of these eucalyptus samples is correlated to the cell wall polymer mobility that is dependent on the crosslink density and/or entanglement of the biopolymer matrix.

References

1. Chen, F. and R. A. Dixon (2007), Nature Biotechnology, "Lignin Modification Improves Fermentable Sugar Yields for Biofuel Production," 25, 759−761.
2. Ziebell, A., K. Gracom, R. Katahira, F. Chen, Y. Q. Pu, A. Ragauskas, R. A. Dixon, and M. Davis (2010), The Journal of Biological Chemistry, "Increase in 4-Coumaryl Alcohol Units During Lignification in Alfalfa (Medicago sativa) Alters the Extractability and Molecular Weight of Lignin," 285(50), 38961−38968.

The BioEnergy Science Center 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.

6. Consolidated Bioprocessing of Cellulose to Isobutanol in Clostridium thermocellum

Paul P. Lin[1,2*] (phlin@ucla.edu), Lemuel M. Soh,[1,2] Amy H. Morioka,[1,2] Kouki M. Yoshino,[1] Sawako Konishi,[1] Sharon C. Xu,[1] Jennifer L. Takasumi,[1,2] James C. Liao,[1,2] and Paul Gilna[2]

[1]Department of Chemical and Biomolecular Engineering, University of California, Los Angeles;
[2]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee

http://bioenergycenter.org

Project Goals: BESC biomass formation and modification research involves working directly with two potential bioenergy crops (switchgrass and Populus) to develop varieties that are easier to break down into fermentable sugars. We are using both testing and generation of large numbers of natural and modified plant samples as well as developing genomics tools for detailed studies into poorly understood cell wall biosynthesis pathways. BESC research in biomass deconstruction and conversion targets consolidated bioprocessing (CBP) by studying model organisms and thermophilic anaerobes to understand novel strategies and enzyme complexes for biomass deconstruction.

Biofuels from cellulose could effectively lower cost compared to those from corn or sugar sources. However, biomass recalcitrance currently limits the use of lignocelluloses. CBP is a potential solution in which biomass hydrolysis and fermentation occur simultaneously. Clostridium thermocellum is an attractive thermophilic CBP host because of its high cellulose utilization rate. In addition, isobutanol is an emerging biofuel with comparable energy density to gasoline and compatibility with existing infrastructures. Here, we seek to improve isobutanol titer using an engineered C. thermocellum. However, C. thermocellum cellulosic isobutanol production includes several challenges: (1) lack of suitable overexpression system and (2) limited enzyme stability at elevated temperatures. We first developed a small scale fermentation condition to quickly test isobutanol production. Then we applied a medium-throughput workflow to design, assemble, transform different plasmids and directly screen C. thermocellum recombinants for isobutanol production. Our best-engineered strain, CT24, produced 2 g/L isobutanol from cellulose at 50°C.

In addition, we developed a high-throughput enzyme assay to screen for improved thermostability of keto-isovalerate decarboxylase (KIVD). The evolved enzyme, LLM2, remains active after 20 minute incubation at 56°C. With LLM2 overexpressed, the engineered strain CT230 achieves a 0.6 g/L isobutanol titer at 55°C.

The BioEnergy Science Center 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.

7. Development of a LC-MS/MS-Based Cell Density Assay for Assessing Microbial Growth on Complex Solid Biomass Substrates

Richard J. Giannone[1,2*] (giannonerj@ornl.gov), Evert Holwerda,[1,3] Lee Lynd,[1,3] Robert Hettich[1,2] and Paul Gilna[1]

[1]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [2]Oak Ridge National Laboratory, Oak Ridge Tennessee; [3]Dartmouth College, Hanover, New Hampshire

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step (consolidated bioprocessing). BESC researchers provide enabling technologies in characterization, 'omics, modeling and data management in order to (1) understand chemical and structural changes within biomass and (2) to provide insights into biomass formation and conversion.

The use of cellulolytic microbes to deconstruct plant-derived lignocellulosic biomass attempts to address the increased global demand for renewable transportation fuels, ultimately through the production of ethanol. One such microbe, the anaerobic thermophile Clostridium thermocellum, is an ideal candidate for further study due to its cellulolytic and ethanologenic capabilities. Though C. thermocellum is naturally efficient at degrading complex plant biomass, current efforts aim to enhance this microbe's catalytic capacity by better understanding the unique enzymatic system it employs. To this end, extensive work using soluble and insoluble model substrates like cellobiose and crystalline cellulose have been used to study various aspects of cellulose utilization. However, lignocellulosic substrates like switchgrass (SWG) have received considerably less attention due in part to difficulties in measuring seemingly routine fermentation metrics such as the C. thermocellum growth dynamics, substrate conversion per unit cell and product yield/titer—all of which are industrially relevant and generally require knowledge of the amount of catalyst (i.e., microbes) present at a given time. Unfortunately, complex solid biomass substrates like SWG often preclude accurate cell count determination both by visual- and protein-based counting assays—the latter of which can be skewed by residual biomass-derived proteins. To address this problem, we developed a novel, mass spectrometry-based proteomic cell density assay that employs a quantitative metric called Matched-Ion inTensity (MIT). This assay capitalizes on proteomic's ability to uniquely assign peptides to specific organisms/species while concurrently providing protein abundance information useful for organism quantitation. Initial work examining a dilution series of C. thermocellum on a constant amount of SWG demonstrated the power of this method whereby total organism MIT (tMIT) provided an almost perfect match to the C. thermocellum dilution series, regardless of interfering SWG-derived proteins or small molecules. Applied to real-world samples, tMIT specifically and accurately tracked cell density across both simple (cellobiose or Avicel) and complex SWG carbon sources, thereby alleviating the above-mentioned analytical challenge. Furthermore, tMIT shows promise in accurately determining organism ratios in simple co-cultures and could become a necessary tool in the field of community proteomics.

The BioEnergy Science Center 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.

8. Engineering a New Alcohol Production Pathway Into the Cellulolytic Thermophile Caldicellulosiruptor Bescii

Gabriel M. Rubinstein[1,2*] (gabe.m.rubinstein@gmail.com), Mirko Basen,[1,2] Israel M. Scott,[1,2] Gerrit J. Schut,[1] Amanda M. Rhaesa,[1,2] Diep M. Nguyen,[1] Gina L. Lipscomb,[1] Robert M. Kelly,[2] Michael W. Adams,[1,2] and Paul Gilna[2]

[1]University of Georgia, Athens; [2]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step [consolidated bioprocessing (CBP)]. BESC research in biomass deconstruction and conversion targets CBP by studying model organisms and thermophilic anaerobes to understand novel strategies and enzyme complexes for biomass deconstruction.

A novel metabolic pathway for bioalcohol production was recently engineered into the hyperthermophilic archaeon Pyrococcus furiosus. Insertion of a bacterial alcohol dehydrogenase gene (AdhA) into P. furiosus allows the conversion of glucose to ethanol in a pathway that proceeds through acetate, rather than acetyl-CoA. Acetate is reduced to acetaldehyde, catalyzed by the native aldehyde ferredoxin oxidoreductase (AOR) in P. furiosus. Acetate reduction is driven thermodynamically by the low redox potential of the electron carrier, a 4Fe-ferredoxin. The heterologously expressed AdhA from Thermoanaerobacter strain X514 catalyzes acetaldehyde reduction to ethanol using nicotinamide adenine dinucleotide phosphate (NADPH) as the electron donor. Interestingly, due to the broad substrate specificities of both AOR and AdhA, a variety of exogenously added carboxylic acids are also converted to their corresponding alcohols.

The goal of this project is to engineer the new pathway into Caldicellulosiruptor bescii. This thermophilic, anaerobic, gram-positive bacterium utilizes unpretreated plant biomass as a sole carbon source. C. bescii does not contain functional alcohol dehydrogenases but it does possess an AOR homolog. However, this enzyme does not utilize acetaldehyde as a substrate and is of unknown function. AOR is a complex tungstopterin-containing enzyme that requires a plethora of processing proteins for cofactor assembly, all of which are encoded in the C. bescii genome. In a first step, we have expressed the gene encoding P. furiosus AOR in C. bescii and active AOR is produced, demonstrating C. bescii has a functional tungstopterin cofactor biosynthesis pathway. We have now constructed several plasmid vectors containing both AOR and AdhA from a variety of different thermophilic microorganisms that take into account possible differences in their specificities for nicotinamide adenine dinucleotide, NADPH and 4Fe- and 8Fe-ferredoxins. Characterization of the recombinant strains is currently underway. The novel AOR/AdhA pathway may provide benefits over the previously published engineered route to ethanol from acetyl-CoA in bescii because this pathway is potentially redox balanced and energy conserving.

9. Field Experiments of Seven Switchgrass TOP Lines

Holly Baxter,[1,2] Mitra Mazarei[1,2*] (mmazarei@utk.edu), Charleson Poovaiah,[1,2] Chunxiang Fu,[2,3] Hui Shen,[2,4] Ajay Biswall,[2,5] Kelsey Yee,[2] Miguel Rodriguez,[2] Olivia Thompson,[2] Guifen Li,[2,3] Desalegn Serba,[2,3] Geoffrey Turner,[2,6] Robert Sykes,[2,6] Steve Decker,[2,6] Mark Davis,[2,6] Jonathan Mielenz,[2] Brian Davison,[2] Malay Saha,[2,3] Yuhong Tang,[2,3] Debra Mohnen,[2,5] Richard Dixon,[2,4] Zeng-Yu Wang,[2,3] C. Neal Stewart, Jr.,[1,2] and Paul Gilna[2]

[1]University of Tennessee, Knoxville; [2]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [3]Samuel Roberts Noble Foundation, Ardmore, Oklahoma; [4]University of North Texas, Denton; [5]University of Georgia, Athens; [6]National Renewable Energy Laboratory, Golden, Colorado
http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step (consolidated bioprocessing). BESC biomass formation and modification research involves working directly with two potential bioenergy crops (switchgrass and Populus) to develop varieties that are easier to break down into fermentable sugars. We are using both testing and generating of large numbers of natural and modified plant samples as well as developing genomics tools for detailed studies into poorly understood cell wall biosynthesis pathways.

Switchgrass (Panicum virgatum L.) is a perennial warm-season C4 grass that has been identified as a candidate lignocellulosic bioenergy crop because of its rapid growth rate, nutrient use efficiency, and widespread adaptation throughout eastern North America. Cell wall recalcitrance in switchgrass and other lignocellulosic feedstocks is a major economic barrier for enabling efficient enzymatic, microbial, or chemical breakdown of cell wall carbohydrates into fermentable sugars. Recent research has focused on developing switchgrass lines that are more amenable to the fermentation process, either by genetic engineering or by selection of low-recalcitrant lines through association analyses.

Greenhouse studies have identified several successful routes for reducing cell wall recalcitrance and improving sugar release efficiency in switchgrass, some of which can also improve biomass yield. Down-regulation of COMT (caffeic acid O-methyltransferase), a lignin biosynthetic gene, reduces lignin and the S/G ratio, thereby improving sugar release and ethanol yield. Over-expression of MYB4, a transcription factor that represses the expression of multiple lignin biosynthetic genes simultaneously, also reduces recalcitrance and improves sugar release efficiency and ethanol yield. Over-expression of miRNA156, a regulator of plant developmental processes, also improves biomass yield and sugar release efficiency. Down-regulation of CWG-1 (cell wall gene) improves biomass yield, sugar release efficiency, and ethanol yield. Down-regulation of CWG-2 improves sugar release efficiency. Natural variants of switchgrass with higher biomass yields and improved sugar release efficiency have been identified through association analysis.

The goal of the current study is to validate previous greenhouse-conducted research with BESC "TOP" Lines in an agronomic field setting. Transgenic and natural variant plants with reduced recalcitrance are being evaluated for two-to-three growing seasons in a University of Tennessee field for: (1) agronomic performance (morphology and biomass yield), (2) lignin content and composition by high-throughput py-MBMS, (3) sugar release by high-throughput enzymatic assays, (4) ethanol yield by SHF and/or SSF, and (5) susceptibility to the rust pathogen (Puccinia emaculata). The transgenic plants and their controls are being grown under USDA APHIS BRS release into the environment permits.

COMT, MYB4, miRNA, and CWG-1 fields are established and have been fully characterized for at least one growing season. COMT down-regulated switchgrass grown in the field for three growing seasons (2011–2013) had reduced lignin content and lower S/G ratios, up to 34% higher sugar release, up to 28% higher ethanol yield, and maintained biomass yields similar to or higher than that of the controls. A two-year (2012–2013) field study with MYB4 over-expressing plants showed that two of three transgenic events had ethanol yields of up to 43% higher than the control, one of which also produced 63% more biomass. Third year (2014) field data analysis of MYB4 over-expressing plants is in progress. First year (2013) field results from miRNA over-expressing plants showed improved biomass yields of up to 149% in some transgenic lines relative to the control. First year (2013) field results from plants with down-regulated CWG-1 showed an improved biomass yield of 50% in one transgenic line relative to the control. Analysis of second year (2014) data for the miRNA and CWG-1 fields is in progress. Data are currently being collected and analyzed for CWG-2 down-regulated plants and natural variant selections, both of which were planted in late spring of 2014 and have just completed their first growing season.

The BioEnergy Science Center 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.

10. Improving Cellulolytic Activity of the C. thermocellum Secretome by Engineering the CipA Scaffoldin

Qi Xu1* (Qi.Xu@nrel.gov), John O. Baker,[1] Daniel G. Olson,2 Lee R. Lynd,[2] Yannick J. Bomble,[1] Michael E. Himmel,[1] and Paul Gilna[3]

[1]National Renewable Energy Laboratory, Golden, Colorado; [2]Dartmouth College, Hanover, New Hampshire; [3]Bioenergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step [consolidated bioprocessing CBP)]. BESC research in biomass deconstruction and conversion targets CBP by studying model organisms and thermophilic anaerobes to understand novel strategies and enzyme complexes for biomass deconstruction.

The extremely high cellulolytic activity of Clostridium thermocellum is based on a cellulase system that is composed of two sub-systems, one being a consortium of free, individual enzymes and the other being composed of large, multi-enzyme complexes (cellulosomes) in which individual enzymes are organized into supramolecular machines for the depolymerization of crystalline cellulose. The most basic step in the organization of cellulosomes is the binding of multiple individual enzymes, through specific cohesin-dockerin interactions, to the primary scaffoldin protein CipA. The enzyme-laden CipA molecules are in turn bound, also through specific modular interactions, to secondary scaffoldins to generate even larger protein complexes, or cellulosomes. Our studies and those of others have demonstrated that CipA plays an essential role in efficient cellulose degradation by C. thermocellum, whereas the secondary scaffoldins are considerably less important. In this study, we present a novel approach to improve C. thermocellum cellulose-degradation performance via modifications of CipA. By characterizing these mutants, we found that the secretome activities of some mutants, both on Avicel and on process-relevant biomass (pretreated corn stover) were greatly enhanced (30 to 80% higher than that of the parent on Avicel and 3 to 25% higher on pretreated corn stover), paving the road for further improvements in the cellulolytic performance of C. thermocellum.

The BioEnergy Science Center 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.

11. Integrating Conventional and Novel Pretreatments with C. thermocellum CBP to Overcome Biomass Recalcitrance

Vanessa A. Thomas[1, 2*] (vanessalutzke@gmail.com), Thanh Nguyen,[1] Rachna Dhir,[1,2] Rajeev Kumar,[1,2] Charles E. Wyman[1,2] and Paul Gilna[2]

[1]University of California-Riverside, Bourns College of Engineering, Riverside; [2]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step [consolidated bioprocessing (CBP)]. BESC researchers provide enabling technologies in characterization, 'omics, modeling and data management in order to (1) understand chemical and structural changes within biomass and (2) to provide insights into biomass formation and conversion.

The major barrier to low-cost utilization of plant biomass as a resource for renewable liquid fuels production is the high costs associated with overcoming plant cell wall recalcitrance, i.e., breaking down the plant cell wall into component sugars. This challenge is particularly great for conversion of plants with high lignin content such as hardwoods, and biomass pretreatment is generally needed to achieve the high yields vital to low costs. In order to cost effectively release and recover plant cell wall sugars, biomass pretreatment should maximize biological digestibility of the residual solid fraction while minimizing degradation of the more readily released sugars that pass into the liquid fraction. As such, pretreatment is a pivotal process step for successful biological production of fuels from biomass as it controls this trade-off between high sugar conversion (residual solids digestibility) and high sugar yields (sugar conservation).

In a typical process for biological conversion of cellulosic biomass to fuels mediated by externally secreted enzymes that are typically derived from the fungi Trichoderma reesei, biomass is subjected to either a physical or chemical pretreatment to open up the cell wall structure by removing and/or disrupting hemicellulose and lignin so that enzymes can access cellulose fibers and other polysaccharides. After pretreatment, biomass polysaccharides remaining in the solids are converted into individual sugars by enzymatic hydrolysis and then metabolized into biofuels, such as ethanol, by fermentation. However, the loadings of fungal cellulase and hemicellulase enzymes needed to realize high sugar yields from hydrolysis of the carbohydrates left in the solids produced by conventional pretreatments have been estimated to cost $0.68 to $1.47 per gallon of ethanol produced.1 CBP, on the other hand, has emerged as the most promising biological pathway to avoid this high cost of enzymes by utilizing a single organism for both enzyme production and fermentation, thus eliminating the need for added enzymes and streamlining multiple processing steps into a single unit operation. In particular, combining Clostridium thermocellum's powerful and extensive native enzyme system with appropriate pretreatment conditions has the potential to overcome biomass recalcitrance as the major economic barrier to large-scale production of fuels from cellulosic biomass.

The focus of this work was to understand the influence of pretreatment on C. thermocellum based CBP performance compared to that by fungal enzymes when applied to feedstocks with relatively low and high recalcitrance. Corn stover was chosen to represent the former because of its established greater susceptibility to breakdown into sugars and near term abundance. Poplus was employed as well because

of its much greater resistance to sugar release and position as a leading hardwood for fuels production. Biological deconstruction of solids produced by hydrothermal pretreatment of natural poplar variants that had been previously identified to have lower recalcitrance as well as one line with high recalcitrance showed C. thermocellum to be more effective in releasing glucan and xylan from pretreated poplar solids than adding fungal enzymes even at very high loadings of up to 65 mg protein/g glucan for the latter. In addition, Co-solvent Enhanced Lignocellulosic Fractionation (CELF), a novel pretreatment technology recently invented by our team, and conventional dilute sulfuric acid (DSA) were used to pretreat corn stover and poplar followed by deconstruction of the resulting solids by both fungal enzymes and C. thermocellum. In this study, an aqueous mixture of 0.5 % w/w sulfuric acid and tetrahydrofuran at a 1:1 ratio that form a single phase was employed for CELF pretreatment at a 5-10 % w/w biomass solids loading and a temperature of 150-160 °C. Comparing the results demonstrated that combining CELF pretreatment with C. thermocellum fermentation resulted in nearly complete conversion (~100%) of the available glucan plus xylan for both poplar and corn stover within 48 hours while conversion was limited to approximately 71% and 85% for DSA pretreated poplar and corn stover, respectively, when followed by C. thermocellum fermentation for 7 days. While enzymes were able to achieve similar yields for CELF pretreated poplar and corn stover, either much higher enzyme loadings (>50 mg/g glucan) or very long incubation times (7 or more days) were required.

References

1. Klein-Marcuschamer, D., P. Oleskowicz-Popiel, B. A. Simmons, and H. W. Blanch, “The challenge of Enzyme Cost in the Production of Lignocellulosic Biofuels,” Biotechnology and Bioengineering 109(4), 2012, pp. 1083–7.

The BioEnergy Science Center 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.

12. Integration of Multi-Omic Data for Advanced Consolidated Bioprocesses

Intawat Nookaew[1] (nookaewi@ornl.gov), Dan Olsen,[2] Mirko Basen,[3] Manesh Shah,[1] Visanu Wanchai,[1] Cong Trinh[4], Philip D. Hyatt,[1] Steve D. Brown,[1] Miriam L.L and,[1] Michael R Leuze,[1] Robert L. Hettich,[1] Robert M. Kelly[5], Mike Adams,[3] Lee Lynd,[2] David W. Ussery,1 and Paul Gilna[1]

[1]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [2]Dartmouth College, Hanover, New Hampshire; [3]University of Georgia, Athens; [4]University of Tennessee, [5]North Carolina State University

http://bioenergy.org

Project Goals: The BioEnergy Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step [consolidated bioprocessing (CBP)]. BESC researchers provide enabling technologies in characterization, 'omics, modeling and data management in order to (1) understand chemical and structural changes within biomass and (2) to provide insights into biomass formation and conversion.

Within the BESC project, several multi-level '□sequencing datasets, such as gen transcriptome, proteome and metabolome data have been generated across different laboratories. In general, the aim of these studies is to better understand the cellular processes of BESC organisms used in CBP, which include Clostridium thermocelleum and Caldicellulosiruptor bescii. We have applied computational and systems biology approaches to quantitatively integrate multi-level 'omic data from various experiments. These approaches provide the important frameworks and strategies for further improvement of the capabilities of the organisms for advancing biofuel production. We present three examples of collaborative efforts within BESC between experimentalists and computational biologists to accelerate the development BESC's CBP organisms.

First, we present a metabolic model for the well-curated central metabolism of C. thermocellum DSM 1313. The model was supported by available transcriptomics and proteomic data, used rigorous fermentation data to validate the model, and provided comprehensive metabolic flux analysis under various environmental and genetic perturbations. With this central metabolic network in hand, we can identify targets of genetic modifications to design optimal networks for enhanced ethanol production, and also hypothesize potential bottlenecks. The central metabolism has been extended to a genome-scale metabolic network (GeM) that will be strengthened by incorporating physiological data and '□omics data for network curation as well as metabolic flux constraints derived from experimental data from the laboratories across from BESC. The developed genome- scale metabolic model takes into account growth media optimized for C. thermocellum and predicts complex cellular phenotypes, such as utilization of lignocellulosic biomass for biofuel production. These predictions can lead to new metabolic engineering targets.

Second, C. thermocellum has been engineered for increased ethanol production by different sequential deletion strategies. After direct genetic engineering on identified metabolic targets was performed, the mutants were re-sequenced and non-targeted genetic mutations were found to occur. Therefore, in-depth analysis of genome re-sequencing of different lineages was used to uncover additional mutations that may play effect on changing on cellular physiology. A computational analysis pipeline tailor-made for genome re-sequencing was used to identify desired and undesired genetic changes, such as single nucleotide variation, transposon changes, and the presence of transformation markers. The genetic changes within

individual strains will be further evaluated with fermentation data as well as transcriptome data to evaluate the impact of complex mutations on ethanol production and undesired fermentation products. We found a non-targeted point mutation in a developed strain impacted on cofactor specificity of AdhE that is beneficial for ethanol production.

Third, we present a predicted transcript-based annotation of the C. bescii genome, based on transcript RNA-seq, 5'RACE RNA-seq and extracellular proteomics for a variety of growth conditions in different carbon sources to determine all transcribed genes, identity of small RNAs/CRISPR, and map the intensity of transcription along the chromosome. Those experimental data will capture the missing features from standard computational gene annotation. Global transcriptional landscapes of C. bescii can be used to design strategies for further strain improvements. Mapping of transcript expression levels and gene regulatory patterns will be identified through the integration of transcriptome, extracellular proteome using network analysis across different growth conditions to inform metabolic modeling and metabolic engineering efforts.

The BioEnergy Science Center 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.

13. Metabolic Engineering of *Clostridium Thermocellum* for High-Yield Ethanol Production

Beth Papanek,[1,2,3] Ranjita Biswas,[1,2] Daniel G. Olson,[2,4] Lee R. Lynd,[2,4] Adam M. Guss[1,2,3*] (gussam@ornl.gov), and Paul Gilna[2]

[1]Oak Ridge National Laboratory, Oak Ridge, Tennessee; [2]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [3]University of Tennessee, Knoxville; [4]Dartmouth College, Hanover, New Hampshire

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step [consolidated bioprocessing (CBP)]. BESC research in biomass deconstruction and conversion targets CBP by studying model organisms and thermophilic anaerobes to understand novel strategies and enzyme complexes for biomass deconstruction.

Clostridium thermocellum is a leading candidate organism for implementing a CBP strategy for biofuel production due to its native ability to rapidly consume cellulose and its existing ethanol production pathway. C. thermocellum converts cellulose and soluble cellodextrins such as cellobiose to lactate, formate, acetate, H2, ethanol, amino acids, and other products. A mutant strain of C. theromcellum was constructed to remove major side product formation, resulting in C. theromcellum ΔhydG Δldh Δpfl Δpta-ack. This strain no longer produced formate, acetate and lactate; hydrogen production decreased by four fold; and the ethanol yield doubled compared with the wild type on cellobiose, crystalline cellulose Avicel, and pretreated biomass. As Avicel loadings increased from 5 g/L to 50 g/L in batch serum bottles, product titers did not increase in the wild-type C. thermocellum beyond that achieved in 5 g/L Avicel, presumably limited by the drop in pH associated with production of organic acids. C. theromcellum ΔhydG Δldh Δpfl Δpta-ack, on the other hand, continued to show increased ethanol titers to a loading of 20 g/L Avicel. While this mutant exhibited a higher ethanol yield, the growth rate was negatively impacted. Therefore, faster growing mutants were enriched by serial transfer. After 1500 generations, individual evolved strains were isolated from the population via single colony purification and were resequenced.

Of the evolved mutants, all produced ethanol at higher yield than the parent C. theromcellum ΔhydG Δldh Δpfl Δpta-ack, including one that produced ethanol at >80% of theoretical yield. Correlations between genotypes and phenotypes for these strains will be discussed.

The BioEnergy Science Center 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.

14. Populus Natural Variation Study, from Single Nucleotide Polymorphisms to Reduced Recalcitrance Biomass and Elucidation of New Function for an EPSP Gene

Wellington Muchero[1*] (mucherow@ornl.gov), Anthony Bryan,[1] Jin-Gui Chen,[1] Kelsey Yee,[1] Hao-Bo Guo,[2] Olivia Thompson,[1] Miguel Rodriquez, Jr.,[1] Sara Jawdy,[1] Lee Gunter,[1] Priya Ranjan,[1] Kai Feng,[1] Stephen DiFazio,[3] Luke Evans,[3] Gancho Slavov,[3] Robert Sykes,[4] Erica Gjersing,[4] Wendy Schackwitz,[5] Joel Martin,[5] Erika Lindquist,[5] Vasanth Singan,[5] Chris Daum,[5] Gerald A. Tuskan,[1] and Paul Gilna[1]

[1]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [2]University of Tennessee; Knoxville; [3]West Virginia University, Morgantown; [4]National Renewable Energy Laboratory, Golden, Colorado; [5]DOE Joint Genome Institute; Walnut Creek, California

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step (consolidated bioprocessing). BESC biomass formation and modification research involves working directly with two potential bioenergy crops (switchgrass and Populus) to develop varieties that are easier to break down into fermentable sugars. We are using both testing and generating of large numbers of natural and modified plant samples as well as developing genomics tools for detailed studies into poorly understood cell wall biosynthesis pathways.

Quantitative trait loci cloning for the discovery of genes underlying polygenic traits has historically been cumbersome in long-lived perennial plants like Populus. Linkage disequilibrium-based association mapping has been proposed as a cloning tool, and recent advances in high-throughput genotyping and whole-genome re-sequencing enable marker saturation to levels sufficient for association mapping with no a priori candidate gene selection. Here we illustrate the successful utilization of this technique to identify a novel isoform of a shikimate biosynthesis enzyme with hitherto unknown transcriptional regulatory activity. The 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase catalyzes the sixth step in the shikimate pathway and is functionally conserved across prokaryotes and eukaryotes. Despite observing protein sizes ranging from 317 to 675 amino acids in 41 sequenced plant genomes, no functional difference has been suggested for these various isoforms. Here, we report the identification of a Populus EPSP that functions as a transcriptional regulator of genes in the phenylpropanoid, flavonoid and tryptophan pathways. This isoform encoded a 518-residue protein with an N-terminus region showing structural similarities to the DNA-binding helix-turn-helix motif and it also exhibited nuclear-localization capability. Based on these observations, we propose a new model for transcriptional regulation of secondary cell wall and flavonoid biosynthesis in Populus.

The BioEnergy Science Center 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.

15. Progress towards Molecular Breeding of Switchgrass for Feedstock Yield and Quality

Desalegn D. Serba[1,2*] (ddserba@noble.org), Malay C. Saha,[1,2] Ananta Acharya,[1,2,3] Jeremy Schmutz,[4,5] Angela Ziebell,[1,6†] Erica Gjersing,[1,6] Robert Sykes,[1,6] Joe H. Bouton,[1,2,3] Katrien M. Devos,[1,3] E. Charles Brummer,[1,2‡] and Paul Gilna[1]

[1]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [2] The Samuel Roberts Noble Foundation, Ardmore, Oklahoma; [3]University of Georgia, Athens; [4]Joint Genome Institute, Walnut Creek, California; [5]Hudson Alpha Institute for Biotechnology, Huntsville, Alabama; [6]National Renewable Energy Laboratory, Golden, Colorado

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step (consolidated bioprocessing). BESC biomass formation and modification research involves working directly with two potential bioenergy crops (switchgrass and Populus) to develop varieties that are easier to break down into fermentable sugars. We are using both testing and generating tools for detailed studies into poorly understood cell wall biosynthesis pathways.

Improved biomass yield and reduced recalcitrance are the two key breeding goals for switchgrass (Panicum virgatum L.), a second-generation biofuel feedstock. Application of molecular techniques facilitates the breeding process and also improves the selection efficiency per cycle. Linkage map of AP13 (lowland) consisting of 515 loci arranged in 18 linkage groups spanning 1,733 cM; and VS16 (upland) consisting 363 loci arranged in 17 linkage groups spanning 1,508 cM were constructed. Quantitative trait loci (QTL) analysis using phenotypic data from multiple environments identified 11 genomic regions controlling both biomass yield and plant height traits. Similarly, 9 and 14 genomic regions were identified controlling sugar release and lignin content, respectively. Both the yield and quality traits are controlled by several QTL of minor effects each. An association mapping population (AMP) consisting of 480 genotypes from 36 accessions mainly collected from the southern United States has been phenotyped at two environments in Georgia and Oklahoma. Significant variation was observed for lignin content (18.3 to 28.0% of dry matter) and total sugar (glucose + xylose) release (245–571 mg g-1). Genotyping by sequencing of 352 of the AMP panels with FseI-MspI identified over 65,000 single nucleotide polymorphisms (SNPs).

Structure analysis distinguished distinct ecotype groups and sub-groups within each ecotype. A subset of 3,196 SNPs that could be genotyped in at least 80% of population was used in preliminary marker-trait associations for biomass yield, plant height, stem thickness, and days to flower. After correction for false discovery, 51 SNPs were found associated with the four traits. Three of the SNPs were associated with both plant height and biomass yield. To generate more markers, genotyping of the population with exome capture is in progress. Based on sugar release, biomass yield, and lignin content six genotypes were identified from the AMP. Two of these genotypes that exhibit both reduced recalcitrance and high biomass yield are currently under further field and greenhouse evaluations.

The BioEnergy Science Center 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.

16. Selection of Superior Transgenic Switchgrass Through Microbial Bioconversion for Enhanced Ethanol Production

Steven D. Brown[1,2,*] (brownsd@ornl.gov), Kelsey L. Yee,[1,2] Miguel Rodriguez Jr.,[1,2] Olivia A Thompson,[1,2] Alexandru Dumitrache,[1,2] Charleson R. Poovaiah,[1,5] Hui Shen,[1,6] Mitra Mazarei,[1,5] Holly L. Baxter,[1,5] Nancy L. Engle,[1,2] Angela Ziebell,[1,3] Robert W. Sykes,[1,3] Erica Gjersing,[1,3] Chunxiang Fu,[1,4] Zeng-Yu Wang,[1,4] Mark F. Davis,[1,3] Timothy J. Tschaplinski,[1,2] C. Neal Stewart, Jr.,[1,5] Richard A. Dixon,[1,6] Jonathan R. Mielenz,[1,2] Brian H. Davison,[1,2] and Paul Gilna[1,2]

[1]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [2]Oak Ridge National Laboratory, Oak Ridge, Tennessee; [3]National Renewable Energy Laboratory, Golden, Colorado; [4]The Samuel Roberts Noble Foundation, Ardmore, Oklahoma; [5]University of Tennessee, Knoxville; [6]University of North Texas, Denton

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step (consolidated bioprocessing). BESC researchers provide enabling technologies in characterization, ′□omics, modelin in order to (1) understand chemical and structural changes within biomass and (2) to provide insights into biomass formation and conversion.

Switchgrass is a potential bioenergy feedstock but its inherent recalcitrance to bioconversion is an economic hurdle for biofuel production. Improvement to these feedstocks through the use of transgenic technology is a continuing goal, with efforts focused on generating enhanced phenotypes that display increased fermentation yields while maintaining normal growth characteristics in the field. We evaluated field-grown transgenic switchgrass lines with reduced lignin content due to genetic modifications in the lignin biosynthetic pathway. Microbial bioconversion assays of switchgrass feedstocks (unpretreated and pretreated) were performed with traditional enzymatic hydrolysis and yeast-based fermentation as part of an assessment of plant lines with reduced recalcitrance. In addition, the top lines were subjected to a consolidated bioprocessing platform using Clostridium thermocellum. Previously reported, caffeic acid O-methyltransferase (COMT) knockdown and MYB4 transcriptional regulator overexpression switchgrass lines, field-grown through years one, two and three were screened using microbial bioconversion fermentation assays. Several lines from year one displayed up to 30% increased fermentation ethanol yields over the control lines, a trend which continued throughout subsequent plant generations. However, there was a significant, generalized decrease in bioconversion performance between the plants that spent one year in the field and their subsequent years of growth; a phenomenon that is widely observed but not fully explained. Therefore, for field grown switchgrass, only plants in their second or more years of growth are considered to provide a representative bioconversion profile. We also report the first use of a microbe engineered to produce increased amounts of biofuels on a transgenic energy crop that has been modified for the same purpose, and show a cumulative increase in bioconversion performance. The overall results demonstrate lignin pathway modifications can lead to reduced recalcitrance with improved hydrolysis and fermentation yields under multiple microbial conversion processes.

The BioEnergy Science Center 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.

17. Solid Residuals from Populus trichocarpa Demonstrate that Recalcitrance Persists During Consolidated Bioprocessing

Hannah Akinosho[1,2*] (akinoshoho@ornl.gov), Kelsey Yee,[1] Miguel Rodriguez,[1] Wellington Muchero,[1] Yunqiao Pu,[1] Arthur Ragauskas,[1,3] and Paul Gilna[1]

[1]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [2]Georgia Institute of Technology, Renewable Bioproducts Institute, Atlanta; [3]Department of Chemical and Biomolecular Engineering and Department of Forestry, Wildlife, and Fisheries, University of Tennessee, Knoxville

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step (consolidated bioprocessing). BESC researchers provide enabling technologies in characterization, 'omics, modeling and data management in order to (1) understand chemical and structural changes within biomass and (2) to provide insights into biomass formation and conversion.

We seek to determine the structural features of biomass that may influence the hydrolysis patterns of Clostridium thermocellum during the deconstruction of biomass. This study will provide a snapshot of features that have been associated with recalcitrance and their relationship to cellulose hydrolysis.

Consolidated bioprocessing (CBP) is a competitive alternative to current ethanol producing technologies for biomass, such as separation saccharification and fermentation whose cost has rendered them too costly for widespread industrial production. C. thermocellum is a CBP microorganism that localizes over 20 identified enzymes in its cellulosome, or multienzyme complex— a feature that starkly contrasts with conventional processing methods. Cellulases and hemicellulases are two groups of enzymes positioned in the cellulosome, and according to this study, display preferences towards the structural features of P. trichocarpa. Natural variants of autoclaved Populus trichocarpa were treated with C. thermocellum (ATCC 27405) at for five days at 58ºC and analyzed for structural properties that may influence enzymatic deconstruction. The remaining materials, or solid residuals, were compared against controls, or untreated natural variants, for properties including the degree of polymerization (DP) and crystallinity of cellulose (CrI), and the molecular weight of hemicellulose. The cellulases housed in C. thermocellum preferentially hydrolyze low DP (<4000) and higher CrI cellulose (~55%). Changes to hemicellulose structure were not definitive between the controls and the solid residuals.

These results may demonstrate that disruption of the hemicellulose structure precedes cellulose hydrolysis of short DP and higher CrI cellulose. These results suggest that the recalcitrance to complete biomass degradation exists in C. thermocellum as well as fungal cellulases.

The BioEnergy Science Center 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.

18. Understanding Why Knockdown Expression of Populus GAUT12.1 Results in Decreased Recalcitrance and Increased Growth in This Woody Feedstock

Ajaya K. Biswal[1,3,4*] (biswalajaya@ccrc.uga.edu), Zhangying Hao,[2,3,4] Sivakumar Pattathil,[3,4] Xiaohan Yang,[4,5] Kim Winkeler,[4,6] Cassandra Collins,[4,6] Sushree S. Mohanty,[3,4] Elizabeth A. Richardson,[2] Ivana Gelineo-Albersheim,[3,4] Kimberly Hunt,[3,4] David Ryno,[3,4] Robert W. Sykes,[4,7] Geoffrey B. Turner,[4,7] Angela Ziebell,[4,7] Erica Gjersing,[4,7] Wolfgang Lukowitz,[2] Mark F. Davis,[4,7] Stephen R. Decker,[4,7] Michael G. Hahn,[2,3,4] Debra Mohnen,[1,3,4] and Paul Gilna[4]

[1]Department of Biochemistry and Molecular Biology, University of Georgia, Athens; [2]Department of Plant Biology, University of Georgia, Athens; [3]Complex Carbohydrate Research Center, University of Georgia, Athens; [4]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [5]Bioscience Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [6]ArborGen, Inc., Ridgeville, South Carolina; [7]National Renewable Energy Laboratory, Golden, Colorado

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves (1) designing plant cell walls for rapid deconstruction and (2) developing multi-talented microbes or converting plant biomass into biofuels in a single step (consolidated bioprocessing). BESC biomass formation and modification research involves working directly with two potential bioenergy crops (switchgrass and Populus) to develop varieties that are easier to break down into fermentable sugars. We are using both testing and generating tools for detailed studies into poorly understoond cell wall biosynthesis pathways.

The major challenge to use bioenergy feedstocks like woody Populus as a source of second-generation biofuel is the recalcitrance of the biomass to facile deconstruction. Populus biomass is rich in cellulose, xylan and lignin whose degree of cross-linking and interactions result in highly recalcitrant woody biomass that is difficult to bioconvert into ethanol without extensive chemical and enzymatic treatment. GAUT12 (GAlactUronosylTransferase12)/IRX8 (irregular xylem8) is a putative glycosyltransferase proposed to be involved in secondary cell wall glucuronoxylan and/or pectin biosynthesis based on concomitant reductions of both xylan and a subfraction of the pectin homogalacturonan (HG) in Arabidopsis irx8 mutants. Two GAUT12 homologs exist in Populus trichocarpa, PtGAUT12.1 and PtGAUT12.2. Here we down-regulated GAUT12.1 expression in Populus deltoides using an RNAi approach to determine the effects of reduced expression on recalcitrance. The 50–67% knockdown expression of GAUT12.1 in Populus deltoides yielded 4–8% greater glucose release upon enzymatic saccharification than the controls. Unexpectedly, the transgenic lines also displayed 12–52% increased plant height and 12–44% radial stem diameter compared to the controls. Knock-down of GAUT12.1 resulted in 25–47% reduced galacturonic acid (GalA) and 17–30% reduced xylan without affecting the total lignin content, revealing that in Populus wood as in Arabidopsis stems, GAUT12 affects both pectin and xylan formation. Our results suggest that targeting xylan and pectin in woody feedstock genetic improvement programs may be an effective means to increase biomass production and simultaneously improve saccharification efficiency in biofuel biorefineries.

The BioEnergy Science Center 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.

19. Gene editing, artificial chromosomes, and highly efficient conjugal gene transfer in diatoms

*C. L. Dupont[1], cdupont@jcvi.org, P. Weyman[1], B. Karas[1], R. Diner[1,2], S. Lefebvre[1], K. Beeri[1], G. Peers[1], A. E. Allen[1]

[1]J. Craig Venter Institute, La Jolla, CA 92037
[2]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92037
[3]Colorado State University, Fort Collins, CO, 80523

Project Goals: Genome-scale metabolic models are fundamental for the analysis of cellular processes at a systems level and represent an ideal organizational framework for analyses of functional genomics, experimental work and computational studies. In recent years, there has been an increasing interest in high- quality metabolic reconstructions of phototrophic organisms and robust computational tools to integrate 'omic' data from these organisms within genome- scale models. The approach of the project is to combine cutting-edge genome manipulation and physiological characterization with metabolic modeling. The ultimate goal is the exploration of next-generation biofuels through a comprehensive understanding of light-driven lipid metabolism in the model marine diatom Phaeodactylum tricornutum.

The eukaryotic microalgae diatoms hold great promise for bioproduction of fuels and higher value chemicals. However, compared to model genetic organisms such as Escherichia coli and Saccharomyces cerevisiae, functional characterization of diatom genes and strain improvement has been hampered by the inefficient genetic tools. To date, diatoms are transformable only via particle bombardment, where the introduced DNA is integrated randomly into the nuclear genome in a (n=1 - many) copy number. The copy number and location effects introduce a wide range of efficiency in the resulting transformants, and the screening of transformants constitutes substantial labor. Similarly, the methods for gene silencing in diatoms depend upon the introduction of RNAi constructs via particle bombardment, with labor intensive screening of transformants using gene and protein expression, along with an a priori hypothesized phenotype. Here we describe improved methods for gene-specific genome editing using Transcription Activator-like Effector Nucleases (TALENs) and demonstrate its application in the diatom Phaeodactylum tricornutum by knocking out the Ni metalloenzyme urease and examining the phenotype by metabolomics. Further, we have engineered the first nuclear episomal vector for diatoms with a parallel delivery method via conjugation from Escherichia coli to the diatoms P. tricornutum and Thalassiosira pseudonana. We identify a yeast-derived sequence that enables stable episome replication in these diatoms even in the absence of antibiotic selection and show that episomes are maintained as closed circles at copy number equivalent to native chromosomes. This episome delivery system is several orders of magnitude more efficient than particle bombardment and lacks positional or copy number effects, thereby facilitating millions of experimental manipulations in reasonable time periods. The system also allows for efficient complementation of deletional mutants and examples of experimental cycles of forward and reverse genetics for functional gene identification in diatoms at the whole genome scale will be discussed.

Funding: DE-SC0008593-Optimization of energy flow through synthetic metabolic modules and regulatory networks in a model photosynthetic eukaryotic microbe

20. Systems biology of nitrogen assimilation in the diatom Phaeodactylum tricornutum

*C. L. Dupont[1], cdupont@jcvi.org, S. Kleesen[2], J. McCarthy[1], K. Beeri[1], S. Lefebvre[1], J. Gillard[1], G. Peers[3], A. Fernie[2], Z. Nikoloski[2], A. E. Allen[1, 4]

[1]J. Craig Venter Institute, La Jolla, CA 92037, [2]Max Planck Institute of Molecular Plant Physiology, Potsdam, 14476, [3]Colorado State University, Fort Collins, CO, 80523, [4]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92037

Project Goals: Genome-scale metabolic models are fundamental for the analysis of cellular processes at a systems level and represent an ideal organizational framework for analyses of functional genomics, experimental work and computational studies. In recent years, there has been an increasing interest in high- quality metabolic reconstructions of phototrophic organisms and robust computational tools to integrate 'omic' data from these organisms within genome- scale models. The approach of the project is to combine cutting-edge genome manipulation and physiological characterization with metabolic modeling. The ultimate goal is the exploration of next-generation biofuels through a comprehensive understanding of light-driven lipid metabolism in the model marine diatom Phaeodactylum tricornutum.

The eukaryotic microalgae diatoms hold great promise for bioproduction of fuels and higher value chemicals and recent advances by our group have greatly improved the ability to genetically engineer diatoms. Given their evolutionary history of serial endosymbiotic events with increasingly evolved exosymbionts, diatoms contain novel combinations of the biochemical pathways, making them alluring alternative bioproduction systems. However, this also means our understanding of central metabolism, including nitrogen assimilation, is lacking. To address this knowledge gap, we used physiological characterizations, genetic manipulations, transcriptomics, metabolomics, and proteomics to examine the distribution and expression of proteins involved in nitrogen assimilation and metabolism in the pennate diatom Phaeodactylum tricornutum. Cellular measurements of carbon and nitrogen and fourier transform infrared spectrometry profiles revealed dramatic and rapid shifts in cellular nitrogen contents, specifically proteins, with changes in nitrogen availability. They also highlighted a remarkable rapidity in protein synthesis and growth recovery upon the addition of nitrogen to nitrogen-limited cultures, a phenotype of diatoms with ecological implications. Protein localizations revealed both mitochondrial and chloroplast localizations for complete Glutamine-synthase-Glutamate synthase (GS-GOGAT) pathways for ammonia assimilation. The localizations of urease and nitrite reductase suggest that urea and nitrite are assimilated in the mitochondria and chloroplast respectively. Fluxomics with 15urea and 15nitrate confirmed this hypothesis but also highlighted variations in the metabolic profiles of diatoms grown on different nitrogen sources. Short term (<1 hours) multi-time point transcriptomic and proteomic experiments show that the addition of nitrate, nitrite, urea, and ammonia result in divergent expression profiles of proteins centrally involved in nitrogen assimilation. In summary, diatoms have specialized systems for the assimilation of different nitrogen compounds that are very rapidly regulated (<15 minutes) by the availability of the various compounds. All of these results were incorporated into a constraint-based metabolic model of nitrogen assimilation in P. tricornutum, which reveals key metabolite exchanges between subcellular compartments and carbon metabolism while guiding future experiments.

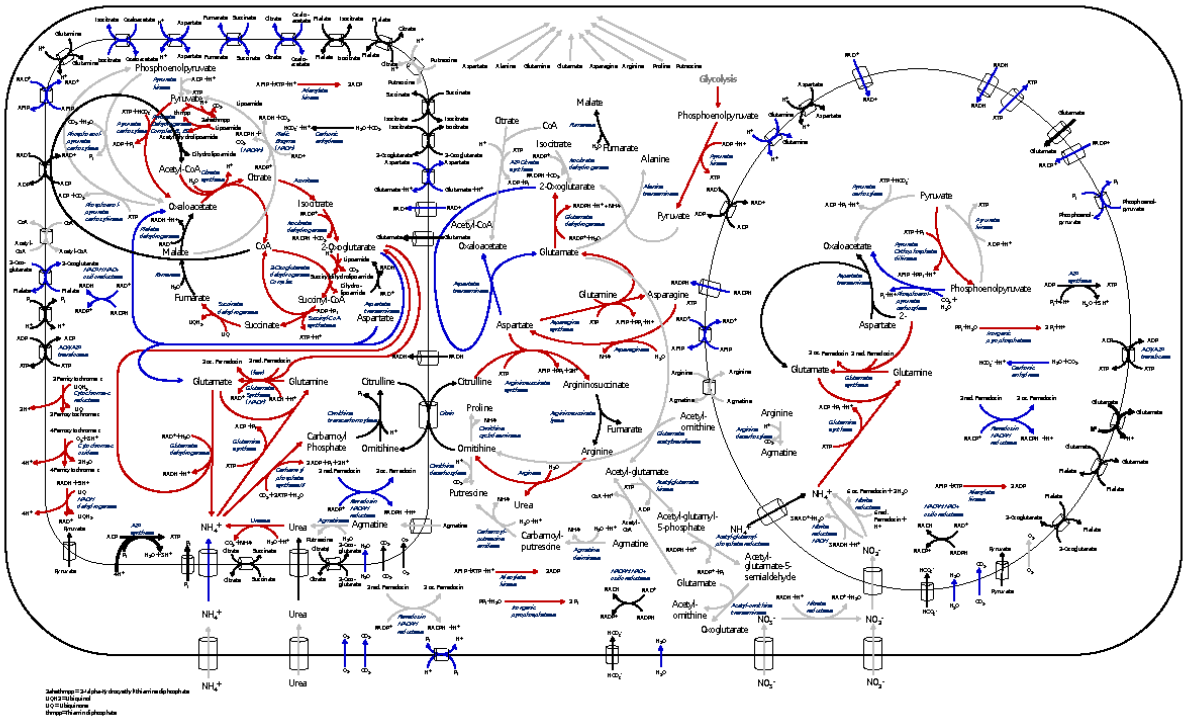


Figure: Modeled metabolic flux and protein expression in nitrogen starved diatoms

Funding: DE-SC0008593-Optimization of energy flow through synthetic metabolic modules and regulatory networks in a model photosynthetic eukaryotic microbe

21. Towards a Comprehensive Knowledge Base for the Marine Diatom Phaeodactylum tricornutum

Jennifer Levering[1*](jlevering@ucsd.edu), Jared Broddrick,[1] Philip Miller,[1] Bernhard Ø Palsson,[1] Graham Peers,[2] Christopher L. Dupont,[3] Karsten Zengler,[1] Andrew E Allen (PI)[3,4]

[1]Bioengineering Department, University of California San Diego, La Jolla, CA, USA, [2]Colorado State University, Fort Collins, CO, USA, [3]J. Craig Venter Institute, La Jolla, CA, USA, [4]Integrative Oceanography Division, Scripps Institute of Oceanography, University of California San Diego, La Jolla, CA, USA

Project Goals: Genome-scale metabolic models are fundamental for the analysis of cellular processes at a systems level and represent an ideal organizational framework for analyses of functional genomics, experimental work and computational studies. In recent years, there has been an increasing interest in high-quality metabolic reconstructions of phototrophic organisms and robust computational tools to integrate 'omic' data from these organisms within genome-scale models. The approach of the project is to combine cutting- edge genome manipulation and physiological characterization with metabolic modeling. The ultimate goal is the exploration of next-generation biofuels through a comprehensive understanding of light-driven lipid metabolism in the model marine diatom Phaeodactylum tricornutum.

Bottom-up reconstructions are biochemically, genetically and genomically structured knowledge-bases that contain information such as reaction stoichiometry, reaction reversibility, and the association between genes, proteins and reactions. The first step in the genome-scale metabolic network reconstruction process involves the generation of a draft reconstruction based on the organism's genome and manually curated reference models. In the second step the draft reconstruction is manually refined using several resources such as genome annotation, primary organism-specific literature, experimental data and databases as KEGG and UniProt. During the manual curation phase the given information for each pathway is reviewed and added if missing.

One challenge of building metabolic network reconstructions for eukaryotic organisms is the prediction of subcellular protein localization. Based on previous work by Sunaga et al. (2014) we implemented a pipeline that combines several bioinformatic tools such as TargetP, SignalP, HECATR, and Mitoprot to generate a putative localization for each protein in the reconstruction. It is well known that protein localization in Phaeodactylum tricornutum (Pt) differs from other algae, such as Chlamydomonas reinhardtii, and pathways are spread among multiple compartments. Comparing the pipeline results with literature information facilitated the assignment of pathway localization.

The reconstructed metabolic network accounts for roughly 1,000 genes associated with approximately 4,800 reactions including over 200 intracellular transport reactions and about 2,500 metabolites distributed across six compartments, namely cytoplasm, extracellular space, chloroplast including thylakoid lumen, mitochondria and peroxisome. The reactions are mass and charge balanced and roughly 96% of the reactions have associated gene-reaction rules. Different identifiers such as KEGG compound and reactions IDs, ChEBIs and InChIs were used to annotate metabolites and reactions. These identifiers are also used in other metabolic network reconstructions and enable the comparison of the metabolic network of Pt to the ones of other microalgae, e.g. C. reinhardtii and Synechocystis sp. PCC6803.

Finally, using the COBRA Toolbox the manually curated reconstruction is converted into a mathematical model. This model will be evaluated and tested against well-known metabolic capabilities of Pt such as growth rate, by-product formation and secretion. An objective function is necessary to compute an optimal network state. To be able to predict growth rates the biomass objective function was chosen. This

objective accounts for both the composition of the cell and the energetic requirements necessary to generate biomass such as amino acids, lipids, carbohydrates, pigments and nucleotides. Currently, we are testing the models capabilities to produce all biomass components and fill gaps if necessary. The next step includes the exploration of the exact experimental derived biochemical composition of the cell and growth rates under different conditions to constrain the model and simulate cell growth under different conditions.

The database support for the model construction was based on two components. The first is the database repository of BiGG (bigg.ucsd.edu) in which highly curated metabolic models are stored within an Oracle system. Metabolite, reaction and gene tables allow cross model comparisons, as well as, links to external references such as KEGG and CAS. In development is an extensive revision and upgrade to be known as BiGG 2.0 and be based on a new Postgres database along with a python language backend system. Within this release, the number of curated models will be increased by an order of magnitude. New functionality will also be introduced. These include an open web service interface or API to allow binding to model, reaction and metabolite SQL queries. A versioning system will allow tracking of model changes over the development cycle. Lastly, the webpage has been redesigned to promote a user friendly, but powerful interface.

Besides reconstructing the metabolic network, a regulatory network is being built based on RNA sequencing data under multiple conditions. After normalizing the raw counts, the data set was filtered for all the genes present in the metabolic reconstruction. As a first step in the network inference, the dimensionality of the input data was reduced by grouping genes that are co- regulated into clusters. The resulting clustering is used to obtain a global regulatory network for Pt suggesting regulatory interactions. This network is used to explore Pt's global expression under novel perturbations.

Supported by DOE-DE-SC0008593, Optimization of Energy Flow Through Synthetic Metabolic Modules

22. The Ecology of Algal Polysaccharide Degradation: Characterizing Novel Fucoidan-Degrading Bacteria

Christopher Corzett[1*] (corzett@mit.edu), Jan-Hendrik Hehemann[1], Martin Polz[1], and Eric Alm[1]

[1] Massachusetts Institute of Technology, Cambridge, MA

Project Goals: Marine algae hold great promise for biofuel production and have advantages over terrestrial biomass and freshwater algae. Despite this potential, little effort has been made to date to harness the enzymatic machinery that bacteria use to convert marine algal carbohydrates into bioenergy substrates. Our project capitalizes on this unexplored opportunity via three distinct activities: bioprospecting for novel algal polysaccharide-degrading genes, functional screening for enzymes with desired biochemical properties, and repackaging pathways in reusable genetic modules.

Marine macroalgae are vital players in the global carbon cycle, and polysaccharides represent a significant output of their primary production (~90% dry mass). Identifying the microbes and metabolic pathways responsible for degrading these sugars is not only crucial to understanding marine carbon flow, but also offers vast potential for biofuel production utilizing seaweed feedstocks. Fucoidans are an important class of structurally heterogeneous sulfated polysaccharides found in brown seaweeds, yet few organisms have been shown to metabolize this abundant carbohydrate. Using environmental samples from coastal waters we have isolated numerous representatives from diverse genera (Vibrio, Lentimonas, Stappia, Neptunomonas, Alteromonas, Tenacibaculum) capable of using fucoidan as a sole carbon source, and demonstrated enzymatic degradation of fucoidan polysaccharides using cellular extracts.

Furthermore, draft genomes of novel fucoidan-degrading Verrucomicrobia sp. have revealed Polysaccharide Utilization Loci (PULs) enriched with numerous and diverse Carbohydrate- Active Enzymes (CAZymes), with some isolates encoding as many as 60 genes with homology to established fucosidases. Preliminary findings reveal differences in the dynamics and extent of fucoidan degradation among closely related isolates, suggesting variation in enzymatic capabilities may reflect metabolic specialization and resource partitioning within natural populations. Specific combinations of natural isolates also appear to complement one another and yield greater overall biomass accumulation, raising the possibility that engineered organisms or communities with a full repertoire of enzymatic machinery and metabolic modules may facilitate the efficient conversion of algal biomass.

This work is supported by the Office of Biological and Environmental Research in the Department of Energy Office of Science (DE-SC0008743).

23. Expanding the breeder's toolbox for perennial grasses: The use of the model perennial grass Brachypodium sylvaticum to identify combinations of transgenes conferring tolerance to multiple abiotic stresses.

Maria Reguera[1], Nir Sade[1], Sean Gordon[3], Christian Tobias[2], Roger Thilmony[2], John Vogel[3], Eduardo Blumwald[1*] (eblumwald@ucdavis.edu)

[1]Dept. of Plant Sciences, University of California, Davis, CA 95616; [2]USDA-ARS Western Regional Research Center, Albany CA94710; [3]DOE Joint Genome Institute, Walnut Creek, CA94598.

Project Goals: The project aims at using a systems-based approach to develop new breeding tools for perennial grasses using the perennial grass Brachypodium sylvaticum as a model, and apply these tools towards the improvement of switchgrass (Panicum virgatum L.).

To further develop Brachypodium sylvaticum as a perennial model grass and facilitate our planned transcriptional profiling, we are sequencing and annotating the genome. We have generated ~40x genome coverage using PacBio sequencing technology using the largest possible size selected libraries (18, 22, 25kb) and are currently assembling the data. In addition, we generated ~100x genome coverage using Illumina technology for SNP polishing. A preliminary assembly of the Illumina data was created to allow us to identify genes of interest before the PacBio assembly was available. The Illumina-only assembly contained 226Mb of sequence in 19k contigs. Using 20x PacBio sequence to scaffold and gap fill this assembly produced an improved assembly of 319Mb in 3k scaffolds. While more complete, this assembly is still missing approximately 21Mb based on c-value estimates of genome size. This underscores the need for the full PacBio assembly which should span repetitive regions to increase contiguity. To aid the assembly of the scaffolds into chromosome-scale assemblies we produced an F2 mapping population and have genotyped 480 individuals using a genotype by sequence approach.

One of the reasons for using B. sylvaticum as a model system is to determine if the transgenes adversely affect perenniality and winter hardiness. Toward this goal, we examined the freezing tolerance of wild type B. sylvaticum lines to determine the optimal conditions for testing the freezing tolerance of the transgenics. A survey of seven accessions noted significant natural variation in freezing tolerance. Seedling or adult Ain-1 plants, the line used for transformation, survived an 8 hour challenge down to -6 deg C and 50% survived a challenge down to -9 deg C. Thus, we will be able to easily determine if the transgenes compromise freezing tolerance.

In the effort to develop biotechnological tools for perennial grass improvement, we have completed the transformation of B. sylvaticum with constructs containing 20 genes shown to be associated with enhanced abiotic stress tolerance in monocots. In addition, we have transformed plants with constructs containing a combination of genes (i.e. SARK::IPT- Ubi::HSR1::Ubi::NHX1) in order to simultaneously overexpress genes associated with drought + heat tolerance + salt tolerance. We generated single copy insert T1 lines for all constructs and the generation and bulking of homozygous T2 lines is well underway. In addition to our B. sylvaticum transgenics, we transformed B. distachyon with many of the same genes. Some of the transgenic B. distachyon plants subjected to a combined stress of both drought and salinity were able to produced higher yields than wild type plants. Our results indicate a great potential for the development of grasses with improved performance and yield in water-limited areas.

24. A versatile phenotyping system and analytics platform reveals diverse temporal responses to water limitation in Setaria

Noah Fahlgren[1], Maximilian Feldman[1], Malia A. Gehan[1], Melinda S. Wilson[1], Christine Shyu[1], Douglas W. Bryant[1], Steven T. Hill[1], Colton J. McEntee[1], Sankalpi N. Warnasooriya[1], Indrajit Kumar[1], Tracy Ficor[1], Stephanie Turnipseed[1], Kerrigan B. Gilbert[1], Thomas P. Brutnell[1], James C. Carrington[1], Todd C. Mockler[1], Ivan Baxter[2 *](ivan.baxter@ars.usda.gov) and Thomas Brutnell[2]

[1]Donald Danforth Plant Science Center, 975 N. Warson Road, St. Louis, MO, 63132 USA [2]USDA-ARS, Donald Danforth Plant Science Center, United States (IB)

http://sviridis.org/
http://foxmillet.org/

Project Goals: "A Systems-Level Analysis of Drought and Density Response in the Model C4 Grass Setaria viridis". The specific aims of the proposal are to: 1) Identify QTL for the effect of drought and density on biomass and seed yield components of Setaria. 2) Conduct in–depth physiological profiles in roots and leaves of a subset of selected lines 3) Integrate datasets and develop metabolic and gene networks for Setaria 4) Develop transformation technologies for Setaria viridis 5) Functionally examine the role of candidate genes deduced by network models; and 6) Develop protocols and best practices for monitoring gene flow in transgenic Setaria. To achieve these aims we will produce one of the most extensive molecular characterizations of plant growth in the field to date, generating several million data points that will be collected from physiological and molecular genetic studies. We will develop novel informatics models and network tools that will guide future molecular characterization in S. viridis and guide breeding efforts in major feedstock targets.

Phenotyping has become the rate-limiting step in using large-scale genomic data to understand and improve agricultural crops. Here, the Bellwether Phenotyping platform for controlled- environment plant growth and automated, multimodal phenotyping is described. The system has capacity for 1,140 plants, which pass daily through stations to record fluorescence images, near infrared images, and visible images. Plant Computer Vision (PlantCV) was developed as an open-source, platform independent quantitative image analysis community resource. In a four week experiment, wild Setaria viridis, and domesticated Setaria italica had fundamentally different temporal responses to water availability. While the lines produced similar levels of biomass under limited water conditions, Setaria viridis maintained the same water use efficiency under water replete conditions, while Setaria italica shifted to less efficient growth. Overall, the Bellwether Phenotyping platform and PlantCV software detected significant effects of genotype, and environment on height, biomass, water-use efficiency, color, plant architecture, and near-infrared traits. All ~79,000 images acquired during the course of the experiment are publically available. We have also collected data using this system on Setaria RIL and association panels.

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

25. Geminivirus replicons for targeted modification of cereal genomes

Javier Gil-Humanes[1], Nicholas J. Baltes[1], Daniel F. Voytas[1*] (voytas@umn.edu) and Thomas P. Brutnell[2]

[1]University of Minnesota, St. Paul, Minnesota; and [2]Donald Danforth Plant Science Center, St. Louis, Missouri

http://sviridis.org/
http://foxmillet.org/

Project Goals: "A Systems-Level Analysis of Drought and Density Response in the Model C4 Grass Setaria viridis". The specific aims of the proposal are to: 1) Identify QTL for the effect of drought and density on biomass and seed yield components of Setaria. 2) Conduct in–depth physiological profiles in roots and leaves of a subset of selected lines 3) Integrate datasets and develop metabolic and gene networks for Setaria 4) Develop transformation technologies for Setaria viridis 5) Functionally examine the role of candidate genes deduced by network models; and 6) Develop protocols and best practices for monitoring gene flow in transgenic Setaria. To achieve these aims we will produce one of the most extensive molecular characterizations of plant growth in the field to date, generating several million data points that will be collected from physiological and molecular genetic studies. We will develop novel informatics models and network tools that will guide future molecular characterization in S. viridis and guide breeding efforts in major feedstock targets.

We developed a deconstructed version of the wheat dwarf virus (WDV) for use as a delivery vector for sequence-specific nucleases and DNA donor templates to enable efficient genome engineering of cereal species. These deconstructed WDV vectors lack the movement and coat protein coding sequences, and consequently cannot spread from cell-to-cell or to other plants. The WDV vectors replicate in transformed wheat scutella, calli, and protoplasts, leading to up to a 110-fold increase in the expression of a GFP reporter gene compared to the non-replicating control. CRISPR/Cas9 sequence-specific nucleases were cloned into the WDV vectors. The increased copy number of the nucleases coding sequences and the donor template led to enhanced levels of gene targeting efficiencies in wheat cells. Other important crops such as corn and rice, as well as green millet, were also found to be suitable hosts of the WDV replicon. The results demonstrate a new platform for genome engineering of cereal crops.

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

26. Rapid Hemispherical Photographic Phenotyping of Productivity and Canopy Dynamics in a Setaria RIL Population

Darshi Banan[1]; Mark Holmes[1]; Hannah Schlake[1]; Rachel E Paul[1]; Max J Feldman[2]; Ivan Baxter[2] *(ivan.baxter@ars.usda.gov); Andrew DB Leakey[1] and Thomas Brutnell[2]

[1] University of Illinois Urbana Champaign, 1402 IGB 1206 W Gregory Drive, Urbana, IL 61801 USA
[2] USDA-ARS, Donald Danforth Plant Science Center, 975 North Warson Road, St. Louis, MO 63132 USA

http://sviridis.org/ http://foxmillet.org/

Project Goals: "A Systems-Level Analysis of Drought and Density Response in the Model C4 Grass Setaria viridis". The specific aims of the proposal are to: 1) Identify QTL for the effect of drought and density on biomass and seed yield components of Setaria. 2) Conduct in–depth physiological profiles in roots and leaves of a subset of selected lines 3) Integrate datasets and develop metabolic and gene networks for Setaria 4) Develop transformation technologies for Setaria viridis 5) Functionally examine the role of candidate genes deduced by network models; and 6) Develop protocols and best practices for monitoring gene flow in transgenic Setaria. To achieve these aims we will produce one of the most extensive molecular characterizations of plant growth in the field to date, generating several million data points that will be collected from physiological and molecular genetic studies. We will develop novel informatics models and network tools that will guide future molecular characterization in S. viridis and guide breeding efforts in major feedstock targets.

Crop genetics and breeding is limited by the ability to accurately gather useful phenotypic information from large, diverse populations of crop genotypes. Most significantly, rapidly and non-destructively assessing the productivity and allometry of crop plants at high frequency during a growing season in the field remains a significant challenge. Hemispherical photography has proven utility in forestry research as a tool to evaluate the growth, structure, and light interception of a canopy. But, high resolution digital cameras with fish-eye lens have been large and relatively expensive. This study tested the use of a small digital camera (GoPro Hero3+) customized with a fully hemispherical lens and miniature self-leveling gimball to rapidly assess leaf area index, biomass production and radiation interception efficiency for quantitative trait loci (QTL) analysis in a diverse population of 186 recombinant inbred lines (RILs) generated by crossing Setaria viridis with S. italica. Plant area index estimated from hemispherical photographs correlated strongly with leaf area index ($r2 = 0.85$), stem biomass ($r2 = 0.76$) and total vegetative biomass ($r2 = 0.84$) in an initial validation experiment where LAI varied over time and genotypes from 0.3 – 3.6. Using these trait regressions, canopy hemispherical photography was then analyzed from the full population of RILs. QTL analysis will be presented comparing loci identified from conventional biomass harvests and the rapid hemispherical photographic phenotyping.

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

27. Screening for Water-Use Efficiency with Stable Isotopes in the Model C4 Grass Setaria viridis

Patrick Z. Ellsworth[1], Patricia V. Ellsworth[1], Max Feldman[2], Darshi Banan[3], Rachel Paul[3], Ivan Baxter[2], Andrew Leakey[3], Asaph B. Cousins (acousins@wsu.edu)[1*], and Thomas P. Brutnell[2]

[1]Washington State University, Pullman, Washington; [2]Donald Danforth Plant Science Center, St.Louis, Missouri; and [3]University of Illinois, Champaign-Urbana

http://sviridis.org/ http://foxmillet.org/

Project Goals: "A Systems-Level Analysis of Drought and Density Response in the Model C4 Grass Setaria viridis". The specific aims of the proposal are to: 1) Identify QTL for the effect of drought and density on biomass and seed yield components of Setaria. 2) Conduct in–depth physiological profiles in roots and leaves of a subset of selected lines 3) Integrate datasets and develop metabolic and gene networks for Setaria 4) Develop transformation technologies for Setaria viridis 5) Functionally examine the role of candidate genes deduced by network models; and 6) Develop protocols and best practices for monitoring gene flow in transgenic Setaria. To achieve these aims we will produce one of the most extensive molecular characterizations of plant growth in the field to date, generating several million data points that will be collected from physiological and molecular genetic studies. We will develop novel informatics models and network tools that will guide future molecular characterization in S. viridis and guide breeding efforts in major feedstock targets.

We screened recombinant inbred lines (RILs) of Setaria viridis x S. italica from the 2013 density and drought field season for water use efficiency (WUE) using leaf carbon and oxygen isotopic signatures (δ13Cleaf and δ18Oleaf, respectively) and total aboveground biomass. Values of δ13C in C4 plants are related to stomatal conductance, so that screening for both δ13C and total aboveground biomass takes into account both WUE and biomass production. Alternatively, the isotopic signature of source water and rates of transpiration determine δ18Oleaf. Across the RILs and the density/drought experiments, several QTLs were associated with WUE as determined by δ18Oleaf, δ13Cleaf, and δ13Cleaf/biomass. Additionally, some of the QTLs associated with δ13Cleaf were found across both the density and drought experiments suggesting that these QTLs for δ13Cleaf are independent of growth conditions. We also found several RILs that scored high in δ13Cleaf or δ13Cleaf/biomass rankings and have common genetic components. These RILs are being tested further for differences in leaf level physiological measurements such as photosynthetic efficiency, stomatal conductance and rates of transpiration. Gas exchange measurements of leaf CO2 assimilation showed that RILs with high δ13Cleaf values had higher photosynthetic capacity than RILs with low δ13Cleaf values. Presently we are conducting in-depth physiological studies to better understand the mechanisms controlling δ13Cleaf in these C4 plants, such as stomatal conductance, respiration, and the efficiency of the CO2 concentrating mechanism. The goal is to identify further physiological traits conferring differences in WUE across the RILs and to link them to particular genetic components.

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

28. Sequence-Enabled Gene Discovery in Setaria viridis: A Model Panicoid Grass

Hui Jiang,[1] Pu Huang,[1] Jeremy Schmutz, [2] Kerrie Barry, [2] Anna Lipzen, [2] Xiaoping Li[1] and Thomas P. Brutnell[1*] (tbrutnell@danforthcenter.org)[1]

[1]Donald Danforth Plant Science Center, St. Louis, Missouri and [2]US Department of Energy, Joint-Genome Institute, Walnut Creek, California

http://sviridis.org/
http://foxmillet.org/

Project Goals: "A Systems-Level Analysis of Drought and Density Response in the Model C4 Grass Setaria viridis". The specific aims of the proposal are to: 1) Identify QTL for the effect of drought and density on biomass and seed yield components of Setaria. 2) Conduct in–depth physiological profiles in roots and leaves of a subset of selected lines 3) Integrate datasets and develop metabolic and gene networks for Setaria 4) Develop transformation technologies for Setaria viridis 5) Functionally examine the role of candidate genes deduced by network models; and 6) Develop protocols and best practices for monitoring gene flow in transgenic Setaria. To achieve these aims we will produce one of the most extensive molecular characterizations of plant growth in the field to date, generating several million data points that will be collected from physiological and molecular genetic studies. We will develop novel informatics models and network tools that will guide future molecular characterization in S. viridis and guide breeding efforts in major feedstock targets.

To accelerate the adoption of Setaria viridis (S. viridis) as a model system, genetic resources and tools are being developed that exploit high throughput sequencing technologies. To create mutant populations for forward and reverse genetics, we have conducted an NMU-mutagenesis and, to date, have generated and characterized approximately 3000 NMU mutant families, from which a TILLING population is being developed. A total of 55 mutant individuals are being sequenced to ~30 x coverage by JGI-DOE to empirically determine mutation frequency. Initial characterizations of the families' shows the majority are transition mutations and the median number of non- synonymous disruptive mutations was 48 per individual. A panicle mutant has been crossed with S. viridis accessions to fine map the gene using Bulk Segregant Analysis (BSA) followed by deep sequencing. We have also initiated the construction of six recombinant inbred populations generated by crossing diverse S. viridis accessions with A10.1. We have assembled a diverse germplasm collection of 430 S. viridis accessions with contributions from many collaborators. A subset of accessions are being sequenced at JGI-DOE to establish a panel for population genetic analysis and genome wide association studies (GWAS). We demonstrate that phenotypes of interest can be mapped to fine resolution with a thorough characterization of the standing phenotypic variation in a subset of this panel. To date, 60 lines has been characterized for phenotypic traits and have been propagated for seed distribution at the USDA GRIN (http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?430573).

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

29. Zbrowse: An interactive GWAS results browser

Greg Ziegler[1], Ryan Hartsock[2] and Ivan Baxter[1 *](ivan.baxter@ars.usda.gov) and Thomas Brutnell[2]

[1] USDA-ARS, Donald Danforth Plant Science Center, 975 North Warson Road, St. Louis, MO 63132 USA
[2] Donald Danforth Plant Science Center, 975 North Warson Road, St. Louis, MO 63132 USA

http://sviridis.org/ http://foxmillet.org/

Project Goals: "A Systems-Level Analysis of Drought and Density Response in the Model C4 Grass Setaria viridis". The specific aims of the proposal are to: 1) Identify QTL for the effect of drought and density on biomass and seed yield components of Setaria. 2) Conduct in–depth physiological profiles in roots and leaves of a subset of selected lines 3) Integrate datasets and develop metabolic and gene networks for Setaria 4) Develop transformation technologies for Setaria viridis 5) Functionally examine the role of candidate genes deduced by network models; and 6) Develop protocols and best practices for monitoring gene flow in transgenic Setaria. To achieve these aims we will produce one of the most extensive molecular characterizations of plant growth in the field to date, generating several million data points that will be collected from physiological and molecular genetic studies. We will develop novel informatics models and network tools that will guide future molecular characterization in S. viridis and guide breeding efforts in major feedstock targets.

We have developed an interactive GWAS results viewer that is an extension of the classic GWAS Manhattan Plot. Zbrowse runs on a personal computer, but is displayed in a web browser and allows for the rapid graphical comparison of GWAS experiments performed on complex traits such as multiple phenotypes measured in multiple locations. The manhattan plots are fully interactive. The browser allows zooming by dragging. Clicking a point in a genome or chromosome-wide view quickly zooms in close enough to see genes under the point. Results can be filtered to only display overlapping QTL between experiments. In addition, results with base pair ranges, such as joint linkage support intervals, can be viewed on the same plot as the GWAS results to quickly visualize overlaps. The browser allows for easy and interactive navigation between plots displaying the entire genome, down to a plot less than a mega base wide displaying gene tracks. Genes under peaks can be clicked to open a browser tab with more information about the gene and all genes under a peak can be viewed in table form in the browser or exported as a comma-separated table.

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

30. De novo assembly of a Setaria viridis reference genome

Jerry Jenkins,[1] Jeremy Schmutz,[1,2] Malia Gehan,[3] Tom Brutnell, [3] and Todd C. Mockler [3*] (tmockler@danforthcenter.org)[1]

[1]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, [2]US Department of Energy, Joint-Genome Institute, Walnut Creek, California, and 3Danforth Plant Science Center, St. Louis, Missouri

http://sviridis.org/
http://foxmillet.org/

Project Goals: Bioenergy grasses promise to provide a sustainable source of renewable fuels for the US bioenergy economy. To develop bioenergy grasses with the desirable traits needed for large scale production, it will be necessary to develop model systems that are closely related to bioenergy feedstocks, but which are more amenable to genetic analysis. One of the most promising model species is Setaria viridis. S. viridis, like all major feedstock targets, is a C4 panicoid grass. However, it is much smaller in stature, flowers within 6 weeks of planting, and can be easily transformed with genes of interest. The objectives of this project were to develop a genomic reference sequence for S. viridis that will provide insights for engineering and intelligent breeding of improved food, fuel, and fiber crops. Further, the S. viridis genome assembly will serve as a useful reference for comparative genomics in the grasses.

The assembly of the S. viridis genome was based on 118x of Illumina whole-genome shotgun sequencing data and performed using ALLPATHS. The 1,417 marker map and Setaria italica synteny were used to identify a total of 21 misjoins. Scaffolds were then oriented, ordered, and joined together based on the map and synteny, and assembled into 9 chromosomes. When the marker map and synteny were in disagreement the map was given priority. A total of 232 joins were made during this process. The final set of pseudomolecule chromosomes were numbered and oriented to match the S. italica chromosomes. A total of 50 gaps in the genome assembly were patched using ~18x coverage of PacBio reads, adding 34,444 bases to the assembly. Patching consisted of assembling the PacBio reads crossing a gap using QUIVER and then integrating the assembled sequence in the genome at the gap. Additionally, 84 homozygous SNPs and 144 homozygous INDELs were corrected using ~70x of independently collected Illumina reads to polish the final assembly. The initial release of the S. viridis genome is composed of 12,531 contigs assembled into 9 chromosomes, along with 724 unintegrated scaffolds. The main genome contig N/L50 is 1597/65.5 Kb and the scaffold N/L50 is 4/46.1 Mbp. The main genome scaffolds span a total of 394.9 Mbp.

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

31. Measurement and modeling of phenylpropanoid metabolic flux in Arabidopsis

Peng Wang[1*] (wang1155@purdue.edu), Longyun Guo,[1] Rohit Jaini,[2] Antje Klempien,[1] Rachel M. McCoy,[1] John Morgan,[1,2] Natalia Dudareva,[1] Clint Chapple[1]

[1] Department of Biochemistry, Purdue University, West Lafayette, IN; [2] School of Chemical Engineering, Purdue University, West Lafayette, IN

Project Goals: We propose to develop a kinetic model for the shikimate and phenylpropanoid pathways. Kinetic models provide insights into the distribution of flux control, thus permitting more intelligent, predictive and effective design of experiments to modulate fluxes towards pathway end products. For this work, we will compare flux measurements in wild-type Arabidopsis plants to plants that are mutant or down-regulated for genes of the lignin biosynthetic pathway, and, those that have been metabolically engineered to bypass the shikimate dependent branch or direct carbon away from lignin biosynthesis to the production of 2-phenylethanol. The outcomes of our proposed kinetic modeling are to identify what remains unknown about the regulation and control of metabolic fluxes to lignin, and to allow development of strategies and predictions of what targets are the most promising candidates for alteration of metabolic flux to lignin.

Lignin is a heterogeneous phenolic polymer that constitutes about 30% of the carbon fixed by photosynthesis in terrestrial plants. Deposited with polysaccharides in the plant secondary cell wall, lignin provides strength and hydrophobicity to plant tissues, but impedes the utilization of lignocellulosic biomass for forage, paper making, and biofuel production. Lignin is derived from the phenylpropanoid pathway, the architecture of which is well understood based upon the biochemical and genetic investigations conducted to date into the enzymes and the genes encoding them. In contrast, we lack a systematic and quantitative view of the factors that determine carbon flux into and within this branched metabolic pathway in plants. Several enzymatic steps in the lignin biosynthetic pathway have been hypothesized to be critical integrators of phenylalanine and lignin biosynthetic flux. To explore the control of carbon allocation in the lignin biosynthetic pathway, we are developing a kinetic model of the pathway in Arabidopsis and performing metabolic control analysis to test the regulatory role of several key steps. We have established an experimental system for flux analysis using excised wild-type Arabidopsis stems. We have found that excised stems continue to grow and lignify in excess of 48 hours when incubated in appropriate medium, and show a distribution of PAL and 4CL activities consistent with the pattern of lignin deposition observed. In the feeding process, PAL and 4CL activities remain constant in the stem tissue. When 13C6-ring labeled phenylalanine was supplied to excised stems, corresponding isotopologues of a number of intermediates have been detected and quantified by LC/MS/MS, and incorporation of 13C6-ring labeled monolignols into lignin was demonstrated by DFRC/GC/MS. We are analyzing the isotope abundance and metabolite concentrations in the pathway in the established system, and using these data for kinetic modeling and metabolic control analysis. Our initial modeling of the first steps of pathway shows a good fit for Phe and p-coumarate. Following the generation of kinetic model in wild-type Arabidopsis, we will take advantage of monolignol biosynthetic mutants in hand to evaluate and refine the model.

This research is supported by the Office of Biological and Environmental Research in the US Department of Energy.

32. Engineering Crassulacean Acid Metabolism (CAM) Photosynthetic Machinery into C3 Arabidopsis

Sung Don Lim[1] Won Cheol Yim,[1] Rebecca L. Albion,[1] and John C. Cushman[1*] (jcushman@unr.edu)

[1]Department of Biochemistry and Molecular Biology, University of Nevada, Reno

http://cambiodesign.org

Project Goals: The long-term goal of this project is to enhance the water-use efficiency (WUE) and adaptability to hotter, drier climates of species that normally perform C3 photosynthesis by introducing crassulacean acid metabolism (CAM). Photosynthetic performance and WUE will be enhanced in Arabidopsis and Populus by: 1) defining the genetic basis of key CAM modules in both eudicot and monocot CAM species, 2) characterizing the regulation of 'carboxylation', 'decarboxylation', and 'inverse stomatal control' gene modules of CAM using a wide variety of functional genomic approaches including loss-of-function studies in model CAM species, 3) deploying advanced genome engineering technologies to enable transfer of fully functional CAM modules into C3 species and 4) analyzing the effects of these transgenic modules on 'stomatal control', CO2 assimilation and transpiration rates, biomass yield, and WUE in Arabidopsis and Populus.

A changing global climate is predicted to lead to declines in agricultural and agroforestry productivity due to increasing water vapor-pressure deficits. One potential approach to improve the sustainability of biofuels feedstocks will be to move the water-wise photosynthetic machinery of CAM into C3 food and bioenergy crops (1). Crassulacean acid metabolism (CAM) is an adaptation that reduces water loss through stomata by shifting all or part of CO2 fixation from the daytime to the nighttime, which results in improved water-use efficiency (WUE) by reductions in evapotranspiration. Introducing the CAM pathway into C3 plants is expected to confer improved WUE in order to allow plants to withstand long episodes of drought or perhaps to expand crop production into semi-arid regions.

Introducing the CAM pathway into C3 plants such as Arabidopsis and Populus will likely require several important tasks: (i) the proper temporal and mesophyll-specific expression patterns of multiple C4 enzymes, their regulatory protein kinases/phosphatases, and related transporters in order to reconstitute the CAM photosynthetic machinery; (ii) an increased mesophyll cell size in order to store malate in the vacuoles of mesophyll cells and limit CO2 diffusion; (iii) an effective multiple gene assembly strategy to effectively express a large number of expression cassettes to build insulate genetic circuits; and (iv) strategies to overcome heterochromatin-mediated gene silencing and reduced mRNA expression of adjacent genes within cis-acting components of gene cassettes to prevent transgene silencing and ensure stable expression of CAM pathway enzymes.

To date, progress has been made in the identification CAM-specific genes that constitute the functional 'carboxylation' and 'decarboxylation' modules of CAM, the identification and selection of promoter regions that are predicted to drive the appropriate temporal (circadian- clock controlled), and mesophyll-specific expression patterns required for CAM, as well as drought-induced promoters for engineering the facultative engagement of CAM, and a GUS/LUC reporter system to validate these predicted expression patterns. Furthermore, increased leaf succulence, an anatomical trait associated with CAM that limits CO2 diffusion out of the leaf, has been engineered in Arabidopsis. Lastly, insulated genetic circuits are being created using a biologically inactive, unique nucleotide sequence (UNS)-guided Gibson isothermal assembly strategy (2, 3) using the gypsy insulator (4) to ensure the optimal expression of CAM genes cassettes by preventing transgene silencing in Arabidopsis. Future efforts will focus on moving the completed genetic circuits (e.g., 'carboxylation' and 'decarboxylation' modules) into Arabidopsis, and

testing transgenic plants for CO2 assimilation and transpiration rates, WUE, and biomass yield. Once proof-of-concept designs are functionally validated in Arabidopsis, synthetic 'carboxylation' and 'decarboxylation' modules will be moved to Populus for further functional testing.

References

1. Cushman JC, Davis SC, Yang X, Borland AM. (2015) Development and use of bioenergy feedstocks for semiarid and arid lands. J. Exp. Bot. *In press*
2. Borland AM, Hartwell J, Weston D, Schlauch KA, Tschaplinski TJ, Tuskan GA, Yang X, Cushman JC. (2014) Engineering crassulacean acid metabolism to improve water-use efficiency. Trends Plant Sci. 19: 327-338.
3. DePaoli HC, Borland AM, Tuskan GA, Cushman JC, Yang X. (2014) Synthetic biology as it relates to CAM photosynthesis: Challenges and opportunities. J. Exp. Bot. 65:3381-3393.
4. Borland AM, Wullschleger SD, Weston DJ, Tuskan GA, Hartwell J, Yang X, Cushman JC. (2015) Climateresilient agroforestry: Physiological responses to climate change and engineering of crassulacean acid metabolism (CAM) as a mitigation strategy. Plant Cell & Environ. In press.
5. Dever LV, Boxall SF, Knerova J, Hartwell J. (2015) Transgenic perturbation of the decarboxylation phase of crassulacean acid metabolism alters physiology and metabolism but only has a small effect on growth. Plant Physiol. 167: 44-59.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Genomic Science Program under Award Number DE-SC0008834. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract Number DE–AC05–00OR22725.

33. Engineering Crassulacean Acid Metabolism (CAM) to Improve Water-use Efficiency of Bioenergy Feedstocks

Anne M. Borland,[1,3] James Hartwell,[2] Susanna F. Boxall,[2] Phaitun Bupphada,[2] Jack Davies,[2] Louisa V. Dever,[2] Nirja Kadu,[2] Jade Waller,[2] David J. Weston,[3] Henrique C. De Paoli,[3] Timothy J. Tschaplinski,[3] Gerald A. Tuskan,[3] Xiaohan Yang,[3] Hao-Bo Guo,[4] Hong Guo,[4] Paul Abraham,[5] Robert L. Hettich,[5] Juli Petereit,[6] Karen A. Schlauch,[6] Rebecca Albion,[7] Travis Garcia,[7] Jungmin Ha,[7] Sung Don Lim,[7] Bernard W. M. Wone,[7] Won Cheol Yim,[7] and John C. Cushman7* (jcushman@unr.edu)

[1]School of Biology, Newcastle University, Newcastle upon Tyne; [2]Department of Plant Sciences, Institute of Integrative Biology, University of Liverpool; [3]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [4]Department of Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville; [5]Chemical Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [6]Nevada Center for Bioinformatics, University of Nevada, Reno; [7]Department of Biochemistry and Molecular Biology, University of Nevada, Reno

http://cambiodesign.org

Project Goals: The long-term goal of this project is to enhance the water-use efficiency (WUE) and adaptability to hotter, drier climates of species that normally perform C3 photosynthesis by introducing crassulacean acid metabolism (CAM). Photosynthetic performance and WUE will be enhanced in Arabidopsis and Populus by: 1) defining the genetic basis of key CAM modules in both eudicot and monocot CAM species, 2) characterizing the regulation of 'carboxylation', 'decarboxylation', and 'inverse stomatal control' gene modules of CAM using a wide variety of functional genomic approaches including loss-of-function studies in model CAM species, 3) deploying advanced genome engineering technologies to enable transfer of fully functional CAM modules into C3 species and 4) analyzing the effects of these transgenic modules on 'stomatal control', CO2 assimilation and transpiration rates, biomass yield, and WUE in Arabidopsis and Populus.

A changing global climate is predicted to lead to declines in agricultural and agroforestry productivity due to increasing water vapor-pressure deficits. One potential approach to improve the sustainability of biofuels feedstocks will be to move the water-wise photosynthetic machinery of CAM into C3 food and bioenergy crops (1). Crassulacean acid metabolism (CAM) is an adaptation that reduces water loss through stomata by shifting all or part of CO2 fixation from the daytime to the nighttime, which results in improved water-use efficiency (WUE) by reductions in evapotranspiration. Introducing the CAM pathway into C3 plants is expected to confer improved WUE in order to allow plants to withstand long episodes of drought or perhaps to expand crop production into semi-arid regions. An essential prerequisite for engineered CAM is to characterize and functionally analyze the leaf anatomy and minimal set of genes and proteins required for its efficient operation.

Introducing the CAM pathway into C3 plants such as Arabidopsis and Populus will likely require several important tasks: (i) the proper temporal and mesophyll-specific expression patterns of multiple C4 enzymes, their regulatory protein kinases/phosphatases, and related transporters in order to reconstitute the CAM photosynthetic machinery; (ii) an increased mesophyll cell size in order to store malate in the vacuoles of mesophyll cells and limit CO2 diffusion; (iii) an effective multiple gene assembly strategy to effectively express a large number of expression cassettes to build insulate genetic circuits; and (iv) strategies to overcome heterochromatin-mediated gene silencing and reduced mRNA expression of adjacent genes within cis-acting components of gene cassettes to prevent transgene silencing and ensure stable expression of CAM pathway enzymes.
To date, progress has been made in the identification CAM-specific genes that constitute the functional

'carboxylation' and 'decarboxylation' modules of CAM, the identification and selection of promoter regions that are predicted to drive the appropriate temporal (circadianclock controlled), and mesophyll-specific expression patterns required for CAM, as well as drought-induced promoters for engineering the facultative engagement of CAM, and a GUS/LUC reporter system to validate these predicted expression patterns. Furthermore, increased leaf succulence, an anatomical trait associated with CAM that limits CO2 diffusion out of the leaf, has been engineered in Arabidopsis. Lastly, insulated genetic circuits are being created using a biologically inactive, unique nucleotide sequence (UNS)-guided Gibson isothermal assembly strategy (2, 3) using the gypsy insulator (4) to ensure the optimal expression of CAM genes cassettes by preventing transgene silencing in Arabidopsis. Future efforts will focus on moving the completed genetic circuits (e.g., 'carboxylation' and 'decarboxylation' modules) into Arabidopsis, and testing transgenic plants for CO2 assimilation and transpiration rates, WUE, and biomass yield. Once proof-of-concept designs are functionally validated in Arabidopsis, synthetic 'carboxylation' and 'decarboxylation' modules will be moved to Populus for further functional testing.

References

1. Borland AM, Hartwell J, Weston D, Schlauch KA, Tschaplinski TJ, Tuskan GA, Yang X, Cushman JC. (2014) Engineering crassulacean acid metabolism to improve water-use efficiency. Trends Plant Science 19: 327-338.
2. Guye P, Li Y, Wroblewska L, Duportet X, Weiss R. (2013) Rapid, modular and reliable construction of complex mammalian gene circuits. Nucleic Acid Research 41: e156.
3. Torella JP, Boehm CR, Lienert F, Chen J-H, Wau KC, Silver PA (2014) Rapid construction of insulated genetic circuits via synthetic sequence-guided isothermal assembly. Nucleic Acid Research 42: 681-689.
4. She W, Lin W, Zhu Y, Chen Y, Jin W, Yang Y, Han N, Bian H, Zhu M, Wang J. (2010) Genetics 185: 1141-1150.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Genomic Science Program under Award Number DE-SC0008834. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract Number DE–AC05–00OR22725.

34. Metabolic Engineering of Increased 2-Phenylethanol Production in Plants

Joseph H Lynch[1], John Morgan[2*] (jamorgan@purdue.edu), Rachel M McCoy[1], Peng Wang[1], Rohit Jaini[2], Clint Chapple[1], Natalia Dudareva[1]

[1]Department of Biochemistry, 2School of Chemical Engineering, Purdue University, West Lafayette, IN 47907

Project Goals: We seek to test a metabolic engineering strategy for the overproduction of 2-phenylethanol, a potential biofuel. Our approach is to overexpress the enzymes catalyzing the multiple biosynthetic steps from phenylalanine to 2-phenylethanol in Arabidopsis thaliana. The strategy utilizes a single gene cassette to simultaneously express all proteins in near stoichiometric amounts under the control of a single promoter. The results will be incorporated into kinetic models which will be used for identifying targets of future metabolic engineering strategies for optimized biofuel production.

2-Phenylethanol is a naturally occurring organic volatile with a characteristic rose scent. Currently, 2-phenylethanol from both natural and artificial sources is utilized for flavoring and fragrance. Although the properties of 2-phenylethanol make it a potential biofuel which can be used as a substitute for petroleum-derived gasoline, there is currently no economical large scale production platform available. In the synthesis of 2-phenylethanol, phenylalanine is first deaminated and decarboxylated by a single enzyme, phenylacetaldehyde synthase, to form phenylacetaldehyde. Subsequent reduction by phenylacetaldehyde reductase forms 2- phenylethanol. We are engineering Arabidopsis as a model system for increased endogenous synthesis of this compound. We have generated gene cassettes to allow for simultaneous expression of combinations of the following genes: AAS, the Arabidopsis phenylacetaldehyde synthase; PAAS, the phenylacetaldehyde synthase from petunia that produces hydrogen peroxide as a byproduct; PAR1, the phenylacetaldehyde reductase from tomato; and KatE, an E. coli catalase which is only utilized in conjunction with PAAS. In each cassette, separate genes are linked via a synthetic 2A linker, which causes ribosome skipping, allowing expression of multiple distinct proteins from a single open reading frame without proteolytic cleavage. Gene cassettes were placed under control of either the CMV 35S promoter for constitutive overexpression, or the Arabidopsis C4H promoter to target expression specifically to lignifying cells. Arabidopsis plants from the T1 generation have been shown by quantitative real-time PCR to overexpress the genes. Analysis of these plants by GC-MS confirmed elevated levels of 2- phenylethanol. Homozygous progeny for the transgenic lines are currently being obtained by self-pollination. The homozygous lines will be used for precise determination of levels of phenylacetaldehyde, 2-phenylethanol, and glycosylated 2-phenylethanol derivatives in both leaves and stems, and comparison of the different cassettes will reveal the most effective gene combination. To analyze whether substrate availability is limiting, the lines with highest production will be crossed with Arabidopsis engineered for increased phenylalanine content, and the progeny will be analyzed to determine how this improves 2-phenylethanol production. The results will be used to refine our kinetic model currently being developed to describe flux toward lignin production.

This research is supported by the Office of Biological and Environmental Research in the US Department of Energy

35. Next Generation Protein Interactomes for Plant Systems Biology and Biomass Feedstocks Research

Shelly A. Trigg[1,3*] (swanamaker@salk.edu), Renee Garza[1], Andrew MacWilliams,[1] Joseph Nery,[1] Anna Bartlett,[1] Rosa Castanon,[1] Adeline Goubil,[1] Joseph Feeney,[1] Ronan O'Malley,[1,2] Mary Galli,[1] and Joseph R. Ecker[1,2,4]

[1]Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla CA; [2]Plant Biology Laboratory, The Salk Institute for Biological Studies, La Jolla CA; [3]Biological Sciences Department, University of California, San Diego; and [4]Howard Hughes Medical Institute

Project Goals: Elucidate interactome characteristics giving rise to desirable phenotypic traits in biofuel feedstock crops by rapidly identifying all their protein-protein interactions using the en masse next-gen yeast two-hybrid screening system, ProCREate.

In order to keep up with global energy demands, it is imperative we acquire more knowledge of biofuel feedstocks for improving their cultivation and energy yield. Knowledge of proteome- wide protein-protein interaction (PPI) networks, or interactomes, that promote robust plant growth or that are perturbed by pathogens could progress strategies for improving cultivation. However, current technologies for obtaining interactome data are not suitable for non-model plants like switchgrass or sorghum because of time, cost, and sensitivity constraints. Even the largest high quality PPI map for the model plant Arabidopsis thaliana (Arabidopsis Interactome 1 or AI-1) contains an estimated 2% of the interactome, and took upwards of 5 years and $8 million to finish1. To address this problem, we have developed a yeast two-hybrid (Y2H) system, ProCREate, that can currently generate interactome data 4x faster and 30x less expensive than the Y2H 'gold standard' assay used to generate AI-1. ProCREate enables en masse pooling and massively paralleled sequencing for the identification of interacting proteins by exploiting Cre- lox recombination. Only interacting proteins can induce reporter gene expression of Cre Recombinase and subsequent Cre-mediated recombination of plasmids containing mutant loxP sequences. The irreversible double mutant loxP linkage of each protein's corresponding coding sequence has allowed us to identify protein interactions using Illumina paired-end sequencing. Assay quality was measured using a set of ~3,300 Arabidopsis ORFs (AIrepeat) previously screened six times for estimating the sensitivity of the Y2H 'gold standard' assay. Combined data from three replicate ProCREate screens of libraries of ~1600 ORFs x libraries of ~2800 ORFs showed a significant increase in assay sensitivity and in sampling sensitivity compared to six replicate Y2H 'gold standard' assays. We retested a subset of ProCREate detected PPIs doing tradition 1x1 Y2H and confirmed 72%, suggesting replicates detected ~60% of all AIrepeat predicted PPIs. While ProCREate has the potential to detect 100% of predicted PPIs by increasing replicates and/or mating efficiency, we are moving on to screen genome size Arabidopsis libraries to estimate assay capacity. We will then use ProCREate to screen cDNA libraries made from feedstocks by shotgun cloning into our Y2H plasmids. PPI data generated will yield deeper insight into many molecular processes and pathways, enabling these interactomes to be used to guide improvement of feedstock productivity and sustainability.

References

1. Arabidopsis Interactome Mapping Consortium, Science 333, 6042 (2011).

We gratefully acknowledge the U.S. DOE Office of Biological and Environmental Research in the U.S. DOE Office of Science for funding this project; DOE-DE-SC0007078.

36. A Strategy for Genome-scale Design, Redesign, and Optimization for Ethylene Production in E. coli

Sean Lynch[1,2], Pin-Ching Maness[1*] (pinching.maness@nrel.gov), Carrie Eckert[1], Jianping Yu[1], and Ryan Gill[2]

[1]Biosciences Center, National Renewable Energy Laboratory, Golden, CO; [2]Department of Chemical and Biological Engineering, University of Colorado, Boulder

Project Goal: This project aims to apply rationale design and state-of-the-art synthetic and systems biology tools to design and optimize E. coli for sustained production of biofuels. Chassis biofuel strains, optimized for production based on predictive design and systems biology knowledge, will serve as the framework for high throughput genome re-design. Using targeted genome-scale and multiplex genome-engineering technologies, strains with improved production will be selected for, and gene-to-trait mapping will identify key factors for further optimization. Herein, we will focus on the construction of an E. coli prototype chassis strain for the production of ethylene by optimization of protein expression, growth medium composition, and metabolic pathway flux. Moreover, we are developing high throughput strategies for selection of strains with increased biofuel/precursor levels.

Abstract

Ethylene is the most highly utilized organic compound for the production of plastics and chemicals, and its catalytic polymerization to alkane fuels has been demonstrated. At present, global ethylene production involves steam cracking of a fossil-based feedstock, representing the largest CO2-emitting process in the chemical industry. Biological ethylene production has the potential to provide a sustainable alternative while mitigating CO2 emission. The expression of a single gene found in some bacteria and fungi, ethylene-forming enzyme (Efe), can catalyze ethylene formation (1, 2). Construction of the first generation chassis strain is based on E. coli MG1655 as the host and the efe gene from *Pseudomonas syringae* (Ps). EFE has been postulated to catalyze ethylene production according to the equation (3):

3. α-ketoglutarate + Arginine + $3O_2 \rightarrow 2C_2H_4$ + Succinate + $7CO_2$ + $3H_2O$ + guanidine + P5C*

However, its efficient biotechnological application requires a more in-depth understanding of the interactions between the Efe reaction and other metabolic pathways in the cell, which will be afforded by genome-scale synthetic biology. One of the substrates, α-ketoglutarate (AKG), is a key TCA cycle intermediate which sits at a node bridging carbon and nitrogen metabolism. AKG levels are therefore tightly regulated within the cell, which renders it difficult to predict which genetic changes will ultimately lead to increases in AKG substrate availability. We explored media composition (rich versus defined), addition of exogenous substrates (AKG, glutamate, and arginine), and the incorporation of predicted genetic modifications in carbon and nitrogen metabolism to improve flux to both AKG and arginine in the TCA cycle (Figure 1), the findings from which will be reported. As such, current work from our groups at the National Renewable Energy Laboratory and the University of Colorado at Boulder seeks to improve ethylene production by combining traditional metabolic engineering strategies with synthetic biology-enabled evolutionary approaches involving the high-throughput construction of genome-scale libraries. Coupled with novel screens and selections, these methods will identify strains with increased production of key intermediates and/or ethylene.

*P5C: L- $\triangle^1$-pyrroline-5-carboxylate

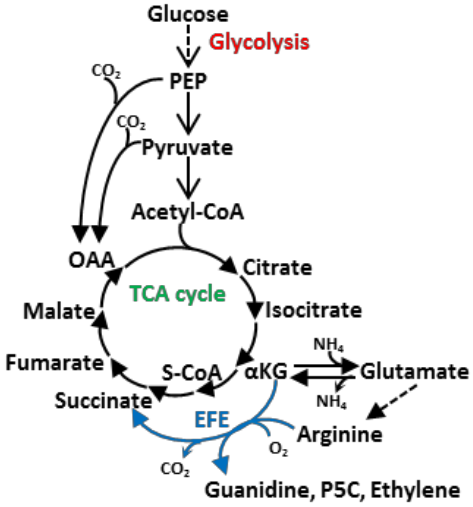


Figure 1. Putative metabolic scheme for ethylene production in E. coli. EFE: ethylene-forming enzyme. P5C: L-$\triangle^1$- pyrroline-5-carboxylate.

References

1. Fukuda, H., T. Ogawa, M. Tazaki, K. Nagahama, T. Fujii, S. Tanase, and Y. Morino. 1992. Two reactions are simultaneously catalyzed by a single enzyme, the arginine-dependent simultaneous formation of two products, ethylene and succinate, from 2-oxoglutarate by and enzyme from Pseudomonas syringae. Biochem. Biophys, Res. Comm. 188: 483-489.
2. Eckert, C., W. Xu, W. Xiong, S. Lynch, J. Ungerer, L. Tao, R. Gill, P. C. Maness, and J. Yu. 2014. Ethylene-forming enzyme and bioethylene production. Biotech. Biofuels 7:33-43.
3. Fukuda, H., T. Ogawa, M. Tazaki, K. Nagahama, T. Fujii, S. Tanase, and Y. Morino.. 1992. Molecular cloning in Escherichia coli, expression and nucleotide sequence of the gene for the ethylene-forming enzyme of Pseudomonas syringae pv. phaseolicola PK2. Biochem. Biophys. Res. Comm. 826-832.

Funding Statement
Grant title: A Platform for Genome-scale Design, Redesign, and Optimization of Bacterial Systems; Project grant number (DE-SC008812) and FWP number (ERWER44) for NREL.

37. Flux balance analysis constraint modification to locate metabolic engineering targets for ethylene production in Escherichia coli

Keesha E. Erickson[1*] (keesha.erickson@colorado.edu), Anushree Chatterjee[1,2], Ryan T. Gill[1,3]

[1]Department of Chemical and Biological Engineering, University of Colorado – Boulder; [2]BioFrontiers Institute, University of Colorado – Boulder; [3]Renewable and Sustainable Energy Institute, University of Colorado Boulder

https://sites.google.com/site/gillgroupcu/
http://www.colorado.edu/UCB/chatterjeelab/

Project Goals: This project aims to cultivate a genome-scale technology platform that allows for prognostic design and optimization of biological systems. One obstacle to successful manipulation of genome-scale systems is that the combinatorial space is intractable (i.e. >>1015). To address this challenge, we are developing computational and synthetic biology tools that allow for more effective search of the design space.

The use of in silico methods has become standard practice to correlate the structure of a biochemical network to the expression of a desired phenotype. Flux balance analysis (FBA) is one of the most prevalent techniques for modeling cellular metabolism. FBA models range from genome-scale reconstructions to simple minimal reaction sets, and have been successfully applied to obtain predictions of growth, theoretical product yields from heterologous pathways, and location of engineering targets to maximize product yield or design antibiotics. We take inspiration from high-throughput recombineering techniques, which show that combinatorial exploration can reveal optimal mutants, and apply the advantages of computational techniques to analyze these combinations. We introduce Constrictor, an in silico tool for flux balance analysis that allows gene mutations to be analyzed in a combinatorial fashion by applying simulated constraints accounting for down-regulation of gene expression. We apply this algorithm to study ethylene production in Escherichia coli through the addition of the heterologous ethylene-forming enzyme from Pseudomonas syringae. Targeting individual reactions as well as sets of reactions results in theoretical ethylene yields that are as much as 25% greater than yields calculated without constraint modification. We demonstrate additional functionality that allows Constrictor to scan a network for possible alternate metabolite products. Constrictor is an adaptable technique that can be used to generate and analyze disparate populations of in silico mutants, select gene expression levels, and troubleshoot metabolic networks.

Publications:
Erickson KE, Gill RT, Chatterjee A. 2014. CONSTRICTOR: Constraint Modification Provides Insight into Design of Biochemical Networks. PLoS ONE 9(11): e113820. doi:10.1371/journal.pone.0113820.

This work is supported by US Department of Energy grant DE-SCOOO8812.

38. The LASER Database: a comprehensive resource for metabolic engineering designs

James D. Winkler* (james.winkler@colorado.edu), Andrea L. Edwards* (andrea.edwards@colorado.edu), Ryan T. Gill

Department of Chemical and Biological Engineering, University of Colorado-Boulder

Project Goals: Metabolic engineers have accumulated a significant body of knowledge regarding strain design, pathway engineering, and media optimization. However, this information remains largely inaccessible to computational pipelines directed toward strain design. To address this disjunction, we have developed the LASER (Learning Assisted Strain EngineeRing) database that contains over 300 curated metabolic engineering studies, all represented using a defined, extensible schema. Researchers can enter additional designs using a convenient web interface to assist with database expansion. Our progress towards using these data to develop metabolic engineering design rules, statistical approaches to strain engineering, and metabolic network visualization are also described. We expect that the LASER database will serve as a resource for a broad range of metabolic engineering and synthetic biology endeavors.

https://sites.google.com/site/gillgroupcu/

39. Metabolite Profiling of the Monolignol Biosynthesis Pathway Using Reversed Phase Liquid Chromatography Coupled with Tandem Mass Spectrometry

Rohit Jaini[1*](rjaini@purdue.edu), Peng Wang[2], Clint Chapple[2], Natalia Dudareva[2], John A. Morgan[1,2]

[1]School of Chemical Engineering; [2]Department of Biochemistry, Purdue University, West Lafayette, IN-47907

Project Goals: We propose to develop a kinetic model for the shikimate and phenylpropanoid pathways. Kinetic models provide insights into the distribution of flux control, thus permitting more intelligent, predictive and effective design of experiments to modulate fluxes towards pathway end products. For this work, we will compare flux measurements in wild-type Arabidopsis plants to plants that are mutant or down-regulated for genes of the lignin biosynthetic pathway, and, those that have been metabolically engineered to bypass the shikimate dependent branch or direct carbon away from lignin biosynthesis to the production of 2-phenylethanol. The outcomes of our proposed kinetic modeling are to identify what remains unknown about the regulation and control of metabolic fluxes to lignin, and to allow development of strategies and predictions of what targets are the most promising candidates for alteration of metabolic flux to lignin.

Abstract

Monolignols constitute the fundamental units of lignin that impart strength, vascular integrity and pathogen resistance to plants. It has been observed that altering lignin synthesis improves conversion efficiency of biomass into energy, food or other industrial chemicals. As a result, a multitude of research efforts have been invested in understanding the mechanism of monolignol biosynthesis via the phenylpropanoid pathway, making the quantification of the network intermediates invaluable. We present a novel and comprehensive method for profiling the metabolites of the monolignol biosynthesis pathway based on liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) in the negative ion mode using multiple reaction monitoring (MRM).

Method improvement by altering chromatographic conditions indicated low pH and low buffer concentration were conducive for obtaining better analyte responses using standards. Arabidopsis thaliana stem tissue was used as a model system for metabolite profiling. Refining and modifying extraction protocols revealed that vortexing combined with incubation at high temperatures resulted in an enhanced extraction of hydrophobic compounds. Process efficiency, constituting matrix effects (ion suppression) and extraction recovery, was evaluated by testing the recovery of metabolites from exogenously fed stems. Twelve of the seventeen pathway intermediates considered were detected and quantified in wild-type A. thaliana stem tissue. Ferulate-5-hydroxylase knockout and overexpression lines were used to validate the analytical method by analyzing the sinapoyl derivatives in the respective strains. The analytical method would be further extended to detect and quantify CoA ester intermediates of the pathway. The CoA esters such as p-coumaryl-CoA and feruloyl-CoA constitute crucial junctions of the phenylpropanoid pathway, the abundances of which may be vital for understanding flux regulation to the monolignols.

40. Analysis of Lipid Metabolism in Saccharomyces cerevisiae: Elucidating Regulation of Triacylglyceride Synthesis

Michael Gossing[1*] (gossing@chalmers.se), Anna Olsson[1] and Jens Nielsen[1]

Systems and Synthetic Biology, Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden

The goal of this project is to improve our understanding of regulation of triacylglyceride synthesis in yeast. For this, Saccharomyces cerevisiae is engineered to accumulate increased amounts of triacylglycerides by overexpressing genes involved in their biosynthesis. Lipidome and transcriptome of the engineered strain and a reference strain will be analyzed. Integrated data analysis will reveal associations and correlations between the different components. Obtained knowledge can be transferred to Yarrowia lipolytica, a distantly related oleaginous yeast.

Lipids are a group of highly diverse molecules with a multitude of biological functions such as formation of biological membranes, storage of energy, cell signaling, and apoptosis. Triacylglycerides function as energy storage and source of membrane building blocks. Triacylglycerides are of particular interest since they can serve as a feedstock for microbial production of oleochemicals and biodiesel.

Oleaginous yeasts such as Yarrowia lipolytica, Lipomyces starkeyi or Rhodotorula glutinis are a suitable choice for microbial production, as they can accumulate significant amounts of storage lipids. Lack of genetic tools and incomplete fundamental knowledge about these organisms, however, makes optimization of production difficult. The yeast Saccharomyces cerevisiae, while not an oleaginous yeast, is a well-established model organism for studying the lipid metabolism of eukaryotes. A strong toolbox for genetic modification and extensive knowledge about this model organism makes it a good choice to study the mechanisms behind storage lipid accumulation. While Saccharomyces cerevisiae is only distantly related to the above mentioned oleaginous yeasts, the enzymes involved in biosynthesis and storage of lipids are generally well-conserved across yeast species.

To fully unlock the potential of yeast as a cell factory for products derived from triacylglycerides, an improved understanding about the regulatory mechanisms behind synthesis, storage and breakdown of triacylglycerides is required. To unravel these mechanisms, an engineered yeast strain that produces increased amounts of triacylglyceride was generated. The flux of carbon was redirected into triacylglyceride synthesis by expressing acetyl-CoA carboxylase mutant ACC1S659A S1157A, phosphatidate phosphatase

PAH1 and diacylglycerol acyltransferase DGA1 under the control of strong constitutive promoters. The engineered yeast strain and the reference strain were cultured in batch under different nutritional conditions. Strains were then used to analyze differences in gene expression and lipid composition as a result from changes in metabolic flux.

41. Regulatory Changes During Lipid Accumulation in Yarrowia lipolytica

Eduard J. Kerkhoven[1*] (eduardk@chalmers.se), Kyle Pomraning[2], Carrie D. Nicora[3], Thomas L. Fillmore[3], Samuel O. Purvine[2], Richard D. Smith[3], Thomas O. Metz[3], Scott E. Baker[2] and Jens Nielsen[1]

[1]Chalmers University of Technology, Gothenburg, Sweden; [2]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington; [3]Fundamental and Computational Science Directorate, Pacific Northwest National Laboratory, Richland, Washington

Project Goals: Our goal is to elucidate the regulation of lipid metabolism in Y. lipolytica to identify new targets to improve the TAG yield.

Oleaginous yeasts such as Yarrowia lipolytica are capable of accumulating lipids up to 70% of their biomass, predominantly in the form of triacylglycerols (TAGs), and this has fuelled interest in exploiting these fungi for the production of biodiesel. Y. lipolytica can use different carbon sources—including a genetic mutant that can grow on xylose—but TAG yields vary considerably under different growth conditions. To further optimise these yields, we are studying the metabolic fluxes and their regulation in Y. lipolytica in different growth conditions on a genomic scale.

A high lipid producing Y. lipolytica strain, with increased expression of diacylglycerol acyltransferase (DGA1), was compared during nitrogen and carbon restricted growth in chemostats at a constant dilution rate of 0.05 h-1. At steady state lipid and fatty acid compositions were determined, together with proteomics (performed under the Pan-omics Program at PNNL) and transcriptomics (RNAseq) measurements. All data were integrated using an updated genome-scale model of metabolism (GEM). Changes in internal fluxes were estimated by random sampling of the GEM solution space.

Regulation on the level of mRNA and protein abundances were primarily focused on reducing nitrogen requirements, while only limited regulation of lipid metabolism was observed. Correlation of changes in mRNA, protein and estimated flux levels were indicative of potentially transcriptionally regulated reactions.

42. Characterization of the Alginate Lyases from Vibrio sp.

Ahmet Badur[1*] (abadur2@illinois.edu), Matthew Plutz[1], Geethika Yalamanchili[1], Jan- Hendrik Hehemann[2], Sujit Sadashiv[1,3], Huimun Zhao1, Martin Polz2, and Christopher V. Rao[1]

[1]Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana- Champaign
[2]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology
[3]Department of Chemical Engineering, Konkuk University, Republic of Korea

Project Goals: This project will harvest 'biomass to biofuel' pathways from algae- associated bacteria, and develop these as reusable genetic parts. Marine algae hold great promise for biofuel production and have advantages over terrestrial biomass and freshwater algae. Despite this potential, little effort has been made to date to harness the enzymatic machinery that bacteria use to convert marine algal carbohydrates into bioenergy substrates. Our project capitalizes on this unexplored opportunity via three distinct activities: bioprospecting for novel algal polysaccharide-degrading genes, functional screening for enzymes with desired biochemical properties, and repackaging pathways in reusable genetic modules.

Brown seaweeds are an attractive source of feedstocks for biofuel production, since they have advantages over terrestrial feedstocks. Brown seaweeds have higher growth rates than terrestrial plants, and they lack crystalline cellulose and lignin. Additionally, brown seaweeds do not impinge on arable land, thus negating the conflict between food and fuel. One of the primary components of brown seaweeds is alginate. Alginate is a copolymer consisting of 1,4 linked epimers α- L-guluronate (G) and β-D-mannuronate (M). The local structure of alginate can take one of three forms: short stretches of polyguluronate (polyG), short stretches of polymannuronate (polyM), or alternating sequences of guluronate and mannuronate. The enzymes that can degrade the linkages within alginate are called alginate lyases. Alginate lyases are classified based on their specific dyad G-G (EC 4.2.2.11), M-M (EC 4.2.2.3), and M-G/G-M bonds that they cleave. Additionally, alginate lyases are classified based on whether they have exolytic or endolytic cleavage.

We are investigating the mechanism of alginate metabolism within marine Vibrio sp. To this end, we next cloned, purified, and characterized the alginate lyases within Vibrio splendidus 12B01 and 13B01 and Vibrio breoganii 1C10. We found that these enzymes are most active between pH 7.5 and 8.5, 20°C and 25°C, and 250 and 1000 mM NaCl. We then determined the enzyme kinetics for each enzyme. We found that each enzyme had a Vmax between 0.090 and 1.7 μM s-1, KM between 22 and 300 μM laminarin, and a turnover number between 0.19 and 4.9 s-1. We also determined the dyad specificity of each lyase; we found G-M, G-G, M-M, and M-G specificity. Between the three organisms, we have characterized 18 alginate lyases, which allows for a broad sampling of how Vibrio sp. degrade alginate. We will next seek to elucidate synergies between lyases. Since a single organism contains multiple lyases, we want to find how alginate lyases can be used simultaneously to degrade alginate to completion and with faster kinetics.

This project is a part of the Biosystems Design Program supported by the Office of Biological and Environmental Research in the DOE Office of Science

43. Characterization of the Laminarinases from Vibrio breoganii 1C10

Ahmet Badur[1*] (abadur2@illinois.edu), Matthew Plutz[1], Geethika Yalamanchili[1], Jan- Hendrik Hehemann[2], Martin Polz[2], and Christopher V. Rao[1]

[1]Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana- Champaign
[2]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology

Project Goals: This project will harvest 'biomass to biofuel' pathways from algae- associated bacteria, and develop these as reusable genetic parts. Marine algae hold great promise for biofuel production and have advantages over terrestrial biomass and freshwater algae. Despite this potential, little effort has been made to date to harness the enzymatic machinery that bacteria use to convert marine algal carbohydrates into bioenergy substrates. Our project capitalizes on this unexplored opportunity via three distinct activities: bioprospecting for novel algal polysaccharide-degrading genes, functional screening for enzymes with desired biochemical properties, and repackaging pathways in reusable genetic modules.

Brown seaweeds are an attractive source of feedstocks for biofuel production, since they have advantages over terrestrial feedstocks. Brown seaweeds have higher growth rates than terrestrial plants, and they lack crystalline cellulose and lignin. Additionally, brown seaweeds do not impinge on arable land, thus negating the conflict between food and fuel. One of the primary components of brown seaweeds is laminarin. Laminarin is a polysaccharide consisting of β-1,3 and β-1,6 linked glucose. The enzymes that can degrade these linkages are called glycoside hydrolases (GHs). More specifically, the β- 1,3 linkage is degraded by enzymes belonging to seven GH families: GH3, GH5, GH16, GH17, GH55, GH64, and GH81. β-1,6 degrading GHs are remain unknown.

We are investigating the mechanism of laminarin metabolism within marine Vibrio sp. Firstly, we have found that laminarinases are induced by growth on laminarin. We next sought to clone, purify, and characterize the four laminarinases within Vibrio breoganii 1C10. We found that these enzymes are most active between pH 6.5 and 8.0, 20°C and 25°C, and 50 and 400 mM NaCl. We then determined the enzyme kinetics for each enzyme. We found that each enzyme had a Vmax between 0.148 and 0.92 μM s-1, KM between 3.4 and 6.0 mM laminarin, and a turnover number between 0.69 and 6.1 s-1.

These results now allow for metabolic engineering of microorganisms that degrade laminarin as their sole carbon source.

This project is a part of the Biosystems Design Program supported by the Office of Biological and Environmental Research in the DOE Office of Science

44. Defining Determinants and Dynamics of Cellulose Microfibril Biosynthesis, Assembly and Degradation

Jocelyn Rose,[1*] (jr286@cornell.edu) Thomas Brutnell,[2] Ying Rong,[2] Indrajit Kumar,[2] Kazuhiro Kikuchi,[2] Seth DeBolt,[3] Detlef Smilgies,[1] Larry Walker,[1]

[1]Cornell University, Ithaca, New York; [2]Donald Danforth Plant Science Center, St. Louis, Missouri; [3]University of Kentucky

Project Goals: The goals of this multidisciplinary project are to: (1) Establish platforms through reverse and forward genetics to identify and manipulate candidate genes that influence cellulose microfibril synthesis and structure; (2) Characterize the effects of altered candidate gene expression on cellulose microfibril synthesis and structure, and develop a mechanistic model for microfibril crystallization; (3) Determine the consequences of altering microfibril architecture on digestibility and integrate this information with nano-scale observations of enzymatic hydrolysis.

The central paradigm for converting plant biomass into soluble sugars for subsequent conversion to transportation fuels involves the enzymatic depolymerization of lignocellulosic plant cell walls by microbial enzymes. Despite decades of intensive research, this is still a relatively inefficient process, due largely to the recalcitrance and enormous complexity of the substrate. A major obstacle is still insufficient understanding of the detailed structure and biosynthesis of major wall components, including cellulose. For example, although cellulose is generally depicted as rigid, insoluble, uniformly crystalline microfibrils that are resistant to enzymatic degradation, the in vivo structures of plant cellulose microfibrils are surprisingly complex.

Crystallinity is frequently disrupted, for example by dislocations and areas containing chain ends, resulting in “amorphous” disordered regions. Importantly, microfibril structure and the relative proportions of crystalline and non-crystalline disordered surface regions vary substantially and yet the molecular mechanisms by which plants regulate microfibril crystallinity, and other aspects of microfibril architecture, are still entirely unknown. This obviously has a profound effect on susceptibility to enzymatic hydrolysis and so this is a critical area of research in order to characterize and optimize cellulosic biomass degradation.

The entire field of cell wall assembly, as distinct from polysaccharide biosynthesis, and the degree to which they are coupled, are relatively unexplored, despite the great potential for major advances in addressing the hurdle of biomass recalcitrance. Our overarching hypothesis is that identification of the molecular machinery that determine microfibril polymerization, deposition and structure will allow the design of more effective degradative systems, and the generation of cellulosic materials with enhanced and predictable bioconversion characteristics.

We believe that the most effective way to address this long standing and highly complex question is to adopt a broad ‘systems approach’. Accordingly, we have assembled a multi-disciplinary collaborative team with collective expertise in plant biology and molecular genetics, polymer structure and chemistry, enzyme biochemistry and biochemical engineering. Our team is using a spectrum of cutting edge technologies, including plant functional genomics, chemical genetics, live cell imaging, advanced microscopy, high energy X-ray spectroscopy and nanotechnology, to study the molecular determinants of cellulose microfibril structure.

Specifically we are coupling with an analytical pipeline to characterize the effects of altering microfibril architecture on bioconversion potential, with the goal of generating predictive models to help guide the

identification, development and implementation of new feedstocks. We are using Arabidopsis thaliana and Brachypodium distachyon as model dicotyledon and grass species, respectively.

We have established an EMS-mutagenized B. distachyon Bd21-3 population of approximately 5000 M2 families. By integrating TILLING methods with Illumina sequencing of target gene amplicons, we have developed an informatics protocol for efficiently identifying single base pair mutations. To date, we have screened for mutations in six genes of interest. A total of 67 SNPs were detected and 55 SNPs were confirmed through Sanger sequencing in 6 target regions. Our results to date indicate that the B. distachyon mutant population and the TILLING by sequencing protocol have great potential for functional studies and should greatly facilitate reverse genetic approaches for gene discovery in this model grass system.

This project is supported by the Office of Biological and Environmental Research in the DOE Office of Science.

45. Developing a robust genetic and omic tool suite for the oleaginous yeast, *Yarrowia lipolytica*

Kyle R.Pomraning[1], Erin L. Bredeweg[1], Ziyu Dai[2], Mary S. Lipton[3], Eduard Kerkhoven[4], Jens Nielsen[4], Galya Orr[1], Thomas O. Metz,[3] and Scott E. Baker (scott.baker@pnnl.gov) [*1]

[1]Environmental Molecular Sciences Laboratory, [2]Energy and Environmental Directorate, [3]Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, Washington and [4]Chalmers University of Technology, Gothenburg, Sweden

Project Goal: Our goal is to enable rapid engineering of the oleaginous yeast, *Yarrowia lipolytica*, for increased rates of biofuel production.

Yarrowia lipolytica is a dimorphic oleaginous ascomycete yeast that produces high levels of lipids under nitrogen limiting conditions. We are developing a suite of genetic tools that allows for rapid and precise target gene deletion, overexpression and tagging. In order to maximize targeted transformation efficiency, we deleted a Y. lipolytica gene (ku70) involved in non-homologous end- joining DNA repair. To validate this system, we targeted four transcription factors for replacement with ura3. Split marker fragments for each were constructed and transformed into FKP424 (ura3, ku70::hph). Confirmation of gene replacement was confirmed by PCR amplifying across the gene of interest and looking for a size shift indicative of replacement with ura3. 5/6 transformants assayed had deletions which were further confirmed phenotypically. We have also collaborated with the PNNL Pan-omics program to identify and validate new Y. lipolytica genes for inclusion for metabolic modeling studies. Six enzyme codes were found not previously identified in Y. lipolytica annotations from KEGG or by annotation of the reference protein sequences with Blast2GO. Identification of candidate genes whose manipulation may lead to increased lipid accumulation in Y. lipolytica is accelerated by integration of- omic data with genetic engineering tools.

We appreciate the funding for this research provided by the DOE BER Genome Science Program through a Biosystems Design project. Proteomic analysis was provided by the PNNL Pan-omics Program.

46. Determination of lipase gene products relevant to lipid mobilization in Yarrowia lipolytica and characterization of their respective knockout strains

Andrew M. Silverman[1,*](asilver@mit.edu), and Gregory Stephanopoulos[1]

[1]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA

Project Goals: The purpose of this investigation is to improve lipid production in *Y. lipolytica* by knocking out lipase genes and preventing lipid mobilization. As *Y. lipolytica* possesses many genes annotated to have lipase domains or activity, it was important to first determine which lipases are critical to lipid catabolism. This research identified three lipase genes that may be responsible for limiting lipid accumulation in Yarrowia.

Y. lipolytica is an oleaginous yeast, with wild-type strains able to accumulate lipids to account for over 20% of its dry cell weight. Over 90% of these lipids are stored in the form of triacylglycerols (TAGs); biodiesel can be made from TAGs through transesterification with short chain alcohols to form fatty acid alkyl esters. Y. lipolytica is capable of consuming many different carbon sources, with acetate being one particularly attractive example as a major component of many agro-industrial or municipal waste streams.

Previously, our lab developed a strain of Yarrowia with increased expression of the native ACC1 and DGA1 genes in order to greatly improve flux towards lipid synthesis. In a bioreactor, this strain typically achieves a maximum lipid content (grams lipids per gram dry cell weight) between 55-65% when grown in minimal medium with acetate as a carbon source, with lipids being produced throughout the fermentation until an abrupt stationary phase is reached. MicroGC measurements for carbon dioxide in the bioreactor headspace during this stationary phase indicate that these cells are still metabolically active at that time, leading us to consider the possibility that the static lipid content during stationary phase could be the result of simultaneous lipid synthesis and degradation (which occur in separate cellular compartments) rather than a more physical limitation of the cell in being unable to store more lipids.

In the first step of lipid degradation in Yarrowia, a lipase or esterase removes fatty acids from TAGs or other glycerolipid molecules. Free fatty acids are then degraded through β-oxidation by the POX (1-5) enzymes into acetyl-CoA units, which can be oxidized in the TCA cycle to make ATP. We carried out a combination of literature review with the results of BLAST searches to find gene products in Yarrowia that are homologous to intracellular lipases from other yeasts. This search allowed us to identify 23 genes that could potentially be responsible for lipid mobilization. This number was further reduced by determining which of these genes may be the most responsible for lipid mobilization and therefore the best candidates for knockouts.

We took RNA samples of the ACC1- and DGA1-overexpressing strain growing in an acetate bioreactor at 2 days and 10 days to represent times during which the cells are in a lipid- producing growth phase and in stationary phase, respectively. We then performed what we will call a Differential Expression Analysis (DEA) on all 23 genes, comparing their expression levels at 10 days (relative to the constitutive actin gene) to their expression levels at 2 days (also relative to actin) to find genes that were most highly upregulated only when the cells have a high (~50-55%) lipid content. We have identified three genes, LIP1 (fold change: 12.1), LIP4 (fold change: 18.7), and LIP10 (fold change: 107.0), that showed large increases in expression during the lipid-static phase. In a ΔKu70 background engineered for us by collaborators at Pacific Northwest Nation Laboratory (PNNL), we have been attempting to create strains with a single lipase removed and a strain with all three of these lipases removed, and we will overexpress ACC1 and DGA1 in all of them to compare the contribution of each lipase to the process of lipid mobilization and achieve a lipid content of Yarrowia higher than what was originally possible.

This project was supported by the U.S. Department of Energy Office of Science.

47. Improving de novo lipids production through engineering the availability of cytosolic NADPH in Yarrowia lipolytica

Kangjian Qiao[1] (kjqiao@mit.edu), Tomas Wasylenko[1] and Gregory Stephanopoulos[1]

[1]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA

Project Goals: We focused on achieving a fundamental understanding of the metabolism of the oleaginous yeast Yarrowia lipolytica and developing tools to characterize and engineer it towards cost-effective lipids production. More specifically, we aimed to improve its fermentation characteristics towards the development of a cost-effective process which converts renewable resources to lipids for biodiesel production. The conversion yield and volumetric productivity on various carbon sources are the key metrics for optimization.

Fatty acids are an important product that can be used for the production of biofuels such as biodiesel. Living organisms share the enzymatic machinery to biosynthesize fatty acids. De novo biosynthesis of fatty acid is an energy-intensive process, requiring two NADPH and one ATP to install one two-carbon brick-acetyl-CoA onto fatty acid chain. Oxidative pentose phosphate pathway was found to provide majority of the intracellular NADPH supporting lipogensis in most of the organisms, in particular, oleaginous yeast Yarrowia lipolytica in light of recent performed flux analysis of lipid hyper-producing strain ACCDGA. Engineering NADPH availability could potentially improve the yield and productivity of lipids. Toward this end, six different strategies have been tested to enrich the cytosolic NADPH in Y. lipolytica and three of them were demonstrated to be effective. In the 2 liter bioreactor experiments, the engineered strain exhibited an improved lipid productivity and yield comparing to the control strain ACCDGA. Our investigation here has sucessfully demonstrated that increasing the engineering NADPH availability is an effective strategy towards cost-effective productions of lipids from glucose using oleaginous yeast.

48. Reprogramming acetyl-CoA metabolism for efficient production of lipid biofuels in Yarrowia lipolytica

Peng Xu* (pengxu@mit.edu), Kangjian Qiao and Gregory N. Stephanopoulos

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

Project Goals: We focused on achieving a fundamental understanding of the metabolic pathways of the oleaginous yeast Yarrowia lipolytica and developing tools to characterize and engineer it. More specifically, we aimed to improve its fermentation characteristics towards the development of a cost-effective process which converts renewable resources to lipids for biodiesel production. The conversion yield and volumetric productivity on various carbon sources are the key metrics for optimization.

Yarrowia lipolytica, an oleaginous yeast, can naturally accumulate large quantity of neutral lipids using a variety of carbon sources. Previous metabolic engineering efforts working on the acyl- CoA related pathways (Tai M et al, Metabolic engineering, 2013) have resulted in efficient triacylglyceride producers by increasing the carbon flux towards malonyl-CoA and sequestrating fatty acyl-CoAs in neutral lipids. This was achieved through the overexpression of native acetyl- CoA carboxylase (ACC) and diacylglycerol acyltransferase (DGA1) genes which may subject to post-transcriptional and post-translational regulations that complicate the metabolic process and limit the pathway efficiency. For example, the S. cerevisiae ACC has more than 20 phosphorylation sites that are subject to allosteric or kinase-related regulations (Li et al, Biotechnology and Bioengineering, 2014). To circumvent these limitations, we engineered alternative heterologous pathways that rewire and optimize the supply of acyl-CoA precursors to unleash the metabolic potential of oleaginous yeast. Specifically a prokaryotic four subunit acetyl-CoA carboxylase along with a plant-derived diacylglycerol acyltransferase have been functionally expressed and led to efficient synthesis of lipid biofuels.

Another major obstacle for efficient production of lipids in Y. lipolytica is pertinent to the unique nitrogen starvation conditions which distinctly separate the lipid accumulation phase from the cell growth phase, which is industrially undesirable due to the relatively low productivity and yield. We have engineered alternative cytosolic acetyl-CoA pathways that are less sensitive to intrinsic nitrogen sensing mechanism and ROS (reactive oxygen species) -scavenging pathways to alleviate the ATP citrate lyase (ACL) internal regulation. In this way, we could partially decouple nitrogen starvation and lipogenesis and move the lipid production from stationary phase to exponential phase. As a result, we were able to achieve very high level of oil content and lipid productivity by reprogramming the acetyl-CoA metabolism in Y. lipolytica. Since acetyl-CoA, malonyl-CoA and other acyl-CoAs are direct precursors for production of oleochemicals including fatty alcohols, fatty alkanes and fatty alkyl esters, the strategies reported in this work should be applicable to develop an yeast biorefinery platform that potentially upgrades low value carbons to high value commodity chemicals.

49. The Novel Cellulolytic Strategy of Cellulomonas gilvus

Caryn S. Wadler[1*], Meagan C. Burnet[2], Stephen J. Callister[2], and Garret Suen[1]

[1]University of Wisconsin-Madison, Madison; [2]DOE Pacific Northwest National Laboratory, Richland

Project Goals: The use of cellulolytic bacteria for generating cellulosic ethanol has recently been the focus of extensive research. This research has primarily focused on using engineered strains of cellulolytic bacteria, but there is growing interest in understanding the fundamental processes of cellulose degradation for identifying novel enzymes or for finding organisms capable of more efficient cellulose fermentation. One family of Actinobacteria, the Cellulomonadaceae, includes the only reported bacteria capable of both degrading cellulose aerobically and anaerobically and producing ethanol as a byproduct of this metabolism. This could thus represent a novel strategy for cellulolytic activity, as the known strategies for aerobic and anaerobic cellulose degradation are functionally distinct with any given bacterium utilizing only one approach. The goal of this project is to understand the strategy through which the cellulomonad Cellulomonas gilvus accomplishes cellulose degradation.

The ability to degrade cellulose is found in a handful of organisms including bacteria. Within bacteria, two mechanisms for cellulose deconstruction exist, including the cell-attached and cell-free models. Both strategies are known to be associated with either anaerobic or aerobic bacteria, respectively, and thus represent mutually exclusive approaches for this difficult physiological process. An exception to this observation is Actinobacteria within the genus Cellulomonas, which are known to degrade cellulose under both aerobic and anaerobic conditions. However, very little is known regarding the process through which this occurs. In this study, we investigated the strategy that cellulomonads use for cellulose degradation, using Cellulomonas gilvus as a model organism. We employed a combination of RNA-seq and proteomics to determine that C. gilvus expresses the same sets of genes and proteins to degrade cellulose under both aerobic and anaerobic conditions. This suggests that it uses a conserved mechanism regardless of condition. To better understand this mechanism, we performed a series of experiments examining the motility, chemotaxis, cellulase localization, and fiber adherence capabilities of C. gilvus. Our results suggest that C. gilvus degrades cellulose using a novel hybrid of the canonical aerobic and anaerobic strategies and provides further insights into the metabolism of this organism.

This work was supported by a DOE Early Career Research Program Award DE–SC0008104 to GS. PNNL is a multiprogram national laboratory operated by Battelle for the DOE under Contract DE-AC05-76RLO 1830.

50. Expanding the Breeders Toolbox: The Development of Biotechnological Tools for Switchgrass Improvement and Transgene Containment.

Sangwoong Yoon[1], Ray Collier[1], Richard Bryan Hernandez[2], Eduardo Blumwald[2], Roger Thilmony[1], John Vogel[3], Christian Tobias[*1] (christian.tobias@ars.usda.gov)

[1]USDA-ARS-WRRC, Crop Improvement and Genetics Research Unit, 800 Buchanan Street, Albany, CA 94710, USA
[2]Dept. of Plant Sciences, University of California, One Shields Ave, Davis, CA 95616, USA.
[3]DOE Joint Genome Institute, 2800 Mitchell Dr., Walnut Creek, CA 94598, USA

There is a great need for the development of alternative and sustainable fuel sources, and switchgrass is poised to become one of the first dedicated biomass crops due to its inherently high yield, C4 metabolism, and perennial nature. Because it is difficult to fix beneficial alleles within outcrossing switchgrass populations, we have initiated a biotechnology-based approach for inbred production via doubled haploids through a process referred to as centromere-mediated genome elimination (Ravi et al., 2010). For this purpose we are using TALENs to target cleavage and repair of several loci encoding CenH3 histones involved in centromere function with the goal of generating whole or partial loss of the functional alleles. Six TALEN pairs were designed to knock out endogenous swichgrass genes PviCENH3-1 and PviCENH3-2. Another 6 pairs of TALEN were designed to knock out the CENH3 and COMT genes of the model grass Brachypodium distachyon. Activity of each TALEN pair was first tested using a yeast-based cleavage assay and then used to create Agrobacterium transformation vectors for stable plant transformation. To express both left and right TALENs in planta, coding regions were separated by a T2A translational skipping sequence and expressed under control of a single maize ubiquitin promoter.

Transgenic T0 lines were regenerated from embryogenic calli and a T7 endonuclease assay was conducted to screen juvenile and flag leaves of these plants for mutations of the target loci. Evidence of somatic mutation at the target loci of these T0 lines will be presented.

Since biotechnologically improved switchgrass will ultimately be grown in open fields, they must be designed to minimize the potential impact of the transgenes on non-target ecosystems. Thus we are implementing a strategy designed to minimize transgene flow from the genetically engineered crop by ablating the transgenic pollen, thus reducing the transmission of transgenes in the environment. We are using the model perennial grass Brachypodium sylvaticum to evaluate the utility of novel transformation constructs to block pollen-mediated transgene flow. We have generated transgenic plants that express barnase under the control of four rice-derived pollen-specific promoters (PS1, PS2, PS3 and OsGEX2). Multiple independent transgenic lines for each construct have been produced and are being evaluated by pollen staining and genetic segregation analyses. Alexander's staining revealed that, relative to wildtype plants, >50% of the pollen collected from the hemizygous T0 containment lines was dead or severely deformed. Analysis of the selfed T1 progeny showed that transgene heritability was 1:1, consistent with the expected segregation frequency for a single locus male lethal transgene, supporting the conclusion that successful ablation of transgenic pollen was achieved in these Brachypodium sylvaticum transgenic plants. Future work is focused on transferring these functionally validated transgene containment constructs into switchgrass.

51. An Innovative Cloning Platform for Pathway Engineering

Henrique C. DePaoli, Gerald A. Tuskan and Xiaohan Yang* (yangx@ornl.gov)

BioSciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

http://cambiodesign.org/

Project Goals: Crassulacean acid metabolism (CAM) is a specialized mode of photosynthesis that features a temporal CO2 pump with nocturnal CO2 uptake, facilitates increased water-use efficiency (WUE), and enables CAM plants to inhabit water-limited environments such as semi-arid deserts or seasonally dry environments. CAM provides an excellent opportunity for engineering both enhanced WUE and photosynthetic performance into bioenergy crops. This project has two main goals: 1) to identify the CAM-associated genes using systems biology approaches and 2) to engineer CAM gene modules into C3 species using synthetic biology approaches. The success of the project will allow biomass production on semi-arid, abandoned, marginal or degraded agricultural lands.

To establish synthetic biology capability for transferring CAM gene modules into C3 species, we developed an innovative method for assembling DNA fragments in vitro. Several cloning strategies are available to support genetic engineering (DePaoli et al., 2014). However, the current availability of cloning methods imposes several limitations to seamless cloning of large and multi-gene constructs. Our technology, called "TNT-cloning", assembles DNA parts in a simple, fast, efficient and flexible manner. Our system combines DNA fragments to be cloned in one single universal library, leaves no undesirable sequences behind, allows "one-pot" reaction with up to 3 fragments to be joined at once and automatically maintains the open reading frame (ORF) in-frame between genes of interest. By combining all cloning elements into one single universal library, the method allows a pre-determined assembly without the need of linkers/adaptors, resulting in a "scar-free" product. In addition to the cloning system, a new buffer, called "TNT-buffer", was developed to allow quick and simultaneous digestion and ligation of DNA fragments, enhancing the efficiency several times compared with current commercial products. The vectors represent a binary platform, making the final gene construct reusable as well as immediately ready for plant transformation. Additionally, the set of plasmids support secondary and tertiary assembling in a loop format, with an exponential reduction of steps for assembling a large number of DNA parts. We demonstrated that the system is wholly functional by cloning, assembling and testing several fragments ranging from 30 bp to 12 kb.

Because this technique is compatible with isothermal (Gibson) assembly, virtually any fragment can be used as an element in the library and circularized without the need to carry the binary- backbone, expanding the technology's use to other systems and making it of special interest for constructing new plasmids/circular molecules. This novel cloning platform will accelerate the creation of multiple gene constructs necessary for CAM engineering and will greatly support the construction and tuning of different pathways in a wide range of organisms.

Note: The TNT cloning system is protected under the invention disclosure 201403357, DOE S- 124,978 and is patent pending. The TNT-buffer formulation is protected under the invention disclosure 201403356, DOE S-124,977 and is patent pending.

References

1. DePaoli, H. C., Borland, A. M., Tuskan, G. A., Cushman, J. C., & Yang, X. (2014). Synthetic biology as it relates to CAM photosynthesis: challenges and opportunities. Journal of Experimental Botany, 65:3381-3393.

This material is based upon work supported by the Department of Energy, Office of Science, Genomic Science Program (under award number DE-SC0008834). Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the US Department of Energy (under contract number DE-AC05-00OR22725).

52. Isotopically nonstationary 13C flux analysis of isobutyraldehyde production in Synechococcus elongatus

Lara J. Jazmin[1], Adeola O. Adebiyi[1], Yao Xu[2], Carl H. Johnson[2], Jamey D. Young[1,3*]
(j.d.young@vanderbilt.edu)

[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 strains of cyanobacteria that are optimized for producing advanced biofuels. This will be done by applying isotopically nonstationary ^{13}C metabolic flux analysis (INST-MFA) to quantify photoautotrophic metabolism in engineered cyanobacterial strains and then redirecting carbon flux toward biofuel production using rational pathway manipulations identified by ^{13}C MFA.

Recent studies have demonstrated the feasibility of converting energy from sunlight and carbon from CO^2 directly into biofuels using photosynthetic microorganisms. Despite the advances made in cyanobacterial biofuels production, the productivities achieved are currently too low for industrial feasibility and few tools are available that specifically address the challenges of determining and redirecting metabolic flux in photosynthetic microbes. Our group is developing novel approaches that use 13C metabolic flux analysis (MFA) to quantitatively assess in vivo metabolic phenotypes of photoautotrophs. Previously, we mapped carbon fluxes in the cyanobacterium Synechocystis sp. PCC 6803 by applying isotopically nonstationary MFA (INST-MFA), which involves model-based regression of transient 13C-labeling patterns of intracellular metabolites. The flux analysis revealed unanticipated photosynthetic inefficiencies tied to oxidative metabolic pathways, despite minimal photorespiration.

We have now applied a similar modeling approach to mapping photoautotrophic metabolism in another strain of cyanobacteria, Synechococcoccus elongatus PCC 7942. Metabolism was quantified in the wild-type strain of S. elongatus, as well as a strain engineered to produce isobutyraldehyde (IBA, a direct precursor of isobutanol). GC-MS profiling of isotopically labeled intracellular metabolites, growth, and IBA production were used to create flux maps of central carbon metabolism in the WT and IBA-producing strains. The flux analysis results quantitatively describe increased flux through phosphoenolpyruvate carboxylase (PEPC) rather than pyruvate kinase (PK) in both studied strains, with an increased overall PEPC flux in the IBA-producing strain. This pointed to an alternative carbon route to pyruvate formation, a major precursor to IBA (phosphoenolpyurvate → oxaloacetate →malate →pyruvate →→IBA).

Based on these results, we generated three cyanobacterial single gene overexpression strains in the IBA-producing background for reactions involved in the pyruvate formation bypass: IBA/$PEPC^{ox}$, IBA/MDH^{ox}, and IBA/ME^{ox}. We found that the IBA/ME^{ox} strain showed a significant improvement in IBA production, while maintaining a growth rate comparable to its IBA-producing parental strain. We also generated an overexpressed pyruvate kinase in the IBA-producing background strain, but found that IBA production did not increase and the growth rate actually decreased.

This work is funded by the U.S. Department of Energy (DOE) Award DE-SC008118.

53. Engineering a *Bacillus subtilis* Strain Capable of Utilizing Marine Macroalgae for Biofuels Production

Eva Garcia-Ruiz1,[2*] (egarciar@illinois.edu), Erik Andersen[1,2], Sujit S. Jagtap[1,3] and Huimin Zhao[1,2]

[1]Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Illinois; [2]Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Illinois; and 3Department of Chemical Engineering, Konkuk University, Seoul, Republic of Korea

http://scs.illinois.edu/~zhaogrp/

Project Goals: This project is focused on unveiling pathways from algae-associated bacteria and refactoring them to use algal polysaccharide as a feedstock for biofuel production. Algal polysaccharides are considered a promising carbon source and are emerging as an important feedstock for the production of biofuels. Despite this potential, little effort has been made to date to harness the enzymatic machinery that bacteria use to convert marine algal carbohydrates into bioenergy substrates. This project harness the unexplored bacterial polysaccharide-degrading pathways to 1) bioprospect novel algal polysaccharide-degrading genes, 2) characterize enzymes with desired biochemical properties, and repackage pathways in reusable genetic modules. This project will yield a set of functional modules for the producing biofuels from marine macroalgae.

Marine macroalgae is emerging as an important feedstock for biofuels production. A number of marine microbes are able to degrade and catabolize efficiently macroalgal polysaccharides by specialized enzymatic pathways which convert these carbohydrates into bioenergy substrates. However, there are few studies to unveil and exploit this enzymatic machinery. Vibrio splendidus is a marine bacterium capable of degrading and catabolizing alginate (a linear copolymer of two uronid acids: β-D-mannuronate and α-L-guluronate) by specialized enzymatic pathways. In this work, we harnessed the alginate-degrading machinery of V. splededus to engineer a model gram-positive bacterium Bacillus subtilis capable of utilizing alginate to produce ethanol. The alginate-degrading pathway in V. splededus is clustered in two separated fragments of DNA that contain a set of genes for alginate transport and metabolism. This cluster of ~49 kb was assembled using the DNA assembler method and was heterologously expressed in B. subtilis that was previously engineered to produce ethanol. The growth of the engineered B. subtilis was limited when using alginate as the sole carbon source. It is well known that heterologous pathway expression is challenging and it can be hampered by differences in regulatory elements. To circumvent this drawback the biosynthetic pathway was refactored using a collection of heterologous constitutive promoters, ribosomal binding sites, and terminators that are functional in Bacillus sp. Additionally, it was further optimized by addition of novel alginate lyases and removal of redundant enzymes. This microbial platform can be further engineered to create recombinant B. subtilis capable of using macroalgae polysaccharides to produce a variety of biofuels and chemicals.

References:

1. S. S. Jagtap, J.-H. Hehemann, M. F. Polz, J.K. Lee, and H. Zhao. “Comparative Biochemical Characterization of Three Exolytic Oligoalginate Lyases from Vibrio splendidus Revealed Complementary Substrate Scope, Temperature and pH Adaptations.” Applied and Environmental Microbiology, 80, 4207-4214 (2014).
2. H. Badur, S. S. Jagtap, G. Yalamanchali, J.-K. Lee, H. Zhao, and C.V. Rao. “Characterization of the Alginate Lyases from Vibrio splendidus 12B01.” Applied and Environmental Microbiology, doi:10.1128/AEM.03460-14.

This study was supported by the U.S. Department of Energy DE---SC0008743.

54. Responses of and interactions between nitrifying bacteria to environmental changes: a systems level approach

Sayavedra-Soto Luis (sayavedl@science.oregonstate.edu), Brett Mellbye, Frank Chaplen, Jeff Chang, Peter J. Bottomley, Daniel J. Arp.

Oregon State University, Corvallis OR

URL: http://www.science.oregonstate.edu/bpp/Labs/arpd/

Project Goals: The main focus of the project is to create predictive models of ammonia---oxidizing bacteria (AOB) and nitrite---oxidizing bacteria (NOB) that incorporate metabolism, the regulatory interactions that influence metabolism, and the signaling network for interaction with the environment. These models are derived from analysis of AOB and NOB cells grown in chemostats after the determination of their transcritptome, phsyiological responses, and changes in biochemistry. The models are providing tools and predictions for understanding the response of nitrifying organisms to changing environmental conditions, and likely contributions to climate change.

Nitrification is the microbially driven process that oxidizes ammoniacal---N (NH4+/NH3) into nitrate---N (NO3---) via the partially oxidized intermediate, nitrite (NO2). Phylogenetically specific types of beta---proteobacteria (AOB) and thaumarchaea (AOA) oxidize NH3 to NO2, and, sequentially several phylogenetically diverse bacterial genera oxidize NO2-to NO3 (NOB) (Fig. 1). Nitrifying microorganisms can also carry out denitrification under micro--- aerobic or anaerobic conditions producing NO and N2O, gases that can impact climate change. The project is characterizing the interactions between the two model nitrifying bacteria AOB Nitrosomonas europaea and the NOB Nitrobacter winogradskyi.

We have carried out transcriptome studies that showed ~30% of the genes of the AOB, N. europaea, are differentially expressed in steady-state chemostat co-culture with the NOB, N. winogradskyi, versus the transcriptome in single chemostat cultures under identical NH4+-limited conditions (1). The yield of AOB was higher in co-culture with the NOB, providing evidence that in NH4+-limited co-culture the AOB benefit more than did the NOB (1). In the case of N. winogradskyi, there was less evidence of transcriptome differences between co-culture and single culture (~11%). When N. winogradskyi was cultured singly however, in the presence of NO2 and excess

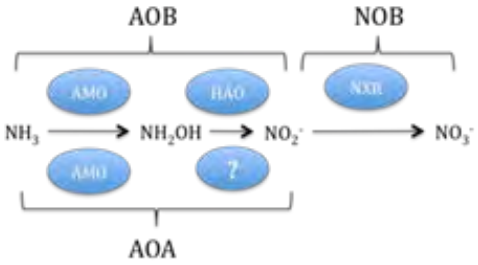


Fig. 1. The biological process of nitrification. Ammonia-oxidizing bacteria (AOB) Oxidize ammonia (NH_3) to nitrite (O_2) using ammonia monooxygenase (AMO) and hydroxylamine oxidoreductase (HAO). Nitrite oxidizing Bacteria (NOB) Oxidize NO_2 To nitrate (NO_3) using nitrite oxidoreductase (NXR). Ammonia-oxidizing Archaea Oxidize NH_3 To NO_2-Using AMO by A mechanism that is not as yet established.

We have made progress in detecting evidence of bacterial cell---cell signaling or quorum sensing (QS) in the culture media of nitrifying bacteria, an understudied area. N. europaea and N. winogradskyi might interact via acyl-homoserine lactone (acyl-HSL) dependent QS. N. europaea showed a cell density-dependent recovery from starvation likely through acyl HSL auto inducer signals (3, 4). Using a broad-range acyl-HSL bioassay we detected evidence of acyl-HSL in N.europaea and N. winogradskyi grown singly and in co-culture in different amounts depending on the cell density of the cultures (Mellbye, manuscript in preparation).

We have established a stoichiometric model for energy metabolism in N. europaea and N. winogradskyi singly and in co-culture based upon the annotated N. europaea genome, and model building through the SEED. The results are in close agreement with our experimental data and with the literature. For example the N. europaea model predicted that growth under hypoxic conditions could be supported by the addition of pyruvate and nitrite, and that parameters that impact NOx production, include the ratio of CO2/HCO3- uptake rates, the O2 uptake rate, the proton (H+)-secretion rate, the NO2 secretion rate, and availability of potential exogenous electron donors for NO2 reduction (Chaplen, manuscript in preparation). We are attempting to refine the constraints models to better understand the factors impacting reductant allocation sinks (including NOx gases) in single and co-cultures.

References:

1. Pérez J, Buchanan A, Mellbye B, Ferrell R, Chang J, Chaplen F, Bottomley P, Arp D, Sayavedra-Soto L. 2015. Interactions of Nitrosomonas europaea and Nitrobacter winogradskyi grown in co-culture. Archives of Microbiology:1-11.
2. Sayavedra-Soto L, Ferrell R, Dobie M, Mellbye B, Chaplen FWR, Buchanan A, Chang J, Bottomley PJ, Arp D. 2014. Nitrobacter winogradskyi transcriptomic response to low and high ammonium concentrations. . FEMS Microbiol. Lett. DOI: http://dx.doi.org/10.1093/femsle/fnu040:First published online: 4 December 2014.
3. Batchelor SE, Cooper M, Chhabra SR, Glover LA, Stewart GS, Williams P, Prosser JI. 1997. Cell density-regulated recovery of starved biofilm populations of ammonia-oxidizing bacteria. Appl Environ Microbiol 63:2281-2286.
4. Burton EO, Read HW, Pellitteri MC, Hickey WJ. 2005. Identification of acyl homoserine lactone signal molecules produced by Nitrosomonas europaea strain Schmidt. Appl. Environ. Microbiol. 71:4906---4909.

This research was supported by the Office of Science (BER) U.S. Department of Energy under grant DE-SC0006943.

55. Multi-'Omic' Analyses of the Dynamics, Mechanisms, and Pathways for Carbon Turnover in Grassland Soil

Cristina N. Butterfield[1*] (butterfc@berkeley.edu), Zhou Li[2], Peter Andeer[3], Susan Spaulding[1], David Burstein[1], Susannah G. Tringe[3,4], Brian C. Thomas[1], Robert Hettich[2], Trent Northen[3], Chongle Pan[2], and Jillian F. Banfield[1]

[1]University of California, Berkeley; [2]Oak Ridge National Laboratory, Oak Ridge, Tennessee; [3]Lawrence Berkeley National Laboratory, Berkeley, CA; [4]DOE Joint Genome Institute, Walnut Creak, CA

http://ggkbase.berkeley.edu

Project Goals: Climate change will alter terrestrial ecosystems. However, the strength and the direction of change will be shaped by feedbacks, most of which will be difficult to predict. Of primary importance in this regard is how the distribution of carbon between the atmosphere and the subsurface will change in response to altered rainfall, temperature and vegetation patterns. Metagenomics, proteomics, transcriptomics, and metabolomics will be used to compare the membership and functioning of soil communities at three different depths below the grass root zone. Overall, this program will result in critical advances in our understanding of soil microbial community carbon cycling to address the question of how the grassland ecosystem will respond to future climate change.

Abstract:
In this project, we will investigate how different rainfall impacts the dynamics of carbon stored in grassland soil, and the potential consequences for release of carbon and other nutrients from soils to streams. However, little is known of the microorganisms that play vital roles in the processing of this largely vegetation-derived soil carbon and how the metabolic activities that occur at different soil depths ultimately impact DOC discharged into streams. We are studying the period around the first Fall rainfall event, when soil- associated carbon fixed during Spring growth is rapidly metabolized, focusing on climate manipulations that differ 1) in the amount of Spring rainfall (above-ground carbon stocks) and 2) on the period of time following the first Fall rain events (soil microbial communities) and 3) soil depth.

Genomic information was obtained for ten soil samples from two sites at 10- 20 cm and 30- 40 cm depth, two of which were collected before the first rainfall and the rest after. Metagenomic data were assembled and draft genomes, including dozens of partial to near complete genomes, were binned and reconstructed based on time series coverage analysis and tetranucleotide frequency using ggKbase and emergent self-organizing maps. While present in most samples, many microbes were more abundant (estimated by coverage) at certain depths. Chloroflexi, Actinobacteria, and Verrucomicrobia were more abundant in the 10- 20 cm sample than 30- 40 cm while the Archaea domain, Methylomirabilis, GAL15, and NC10 were more abundant at 30- 40 cm than the 10- 20 cm depth.

Phylogenetic analysis showed that the soil is dominated by Archaea and several phyla of Bacteria. Most of the organisms clustered with closely related species, creating a dandelion-like structure at the branches' termini. In many cases there were no reference sequences of considerable similarity, representing new phyla and many novel classes and orders within the Gemmatimonadetes, Verrucomicrobia, Deltaproteobacteria, Acidobacteria (most novelty here), and Chloroflexi phyla.

Polar metabolites were extracted from selected pre- and post-rain soil samples using an aqueous extraction method and detected on an Agilent 6550 iFunnel Q-TOF LC/MS system following separation on a zwitterionic stationary phase column. Approximately 140 putative compounds were tracked across soil depth profiles and preliminary analysis of the metabolite composition indicated that carbon and

nitrogen sources (simple sugars and quaternary ammonium compounds) clustered with depth, decreasing from 10 cm to 40 cm.

Total protein samples were extracted from soils using a modified MoBio NoviPure protocol and cleaned up and digested using the Filter-Aided Sample Preparation (FASP) method. The proteome digest samples were analyzed using 2-dimensional liquid chromatography coupled with high-resolution tandem mass spectrometry on a LTQ- Orbitrap Elite mass spectrometer. The mass spectral data were searched using the Sipros algorithm on the Titan supercomputer against a combined soil metagenome database constructed above. On average approximately two thousand proteins or protein groups were identified per soil sample with a peptide-level false discovery rate at 1% and a protein-level false discovery rate at ~2%. Proteomics analysis revealed that corresponding sugar transporters are present and that there are abundant proteins in all samples are those involved in methanotrophy and ammonification.

The results from these "omics" methods are quickly coming together to form a narrative about the dynamic microbial community and its processes below the grass root zone with respect to carbon flow during the Fall rainfall event. We anticipate these below ground terrestrial system feedbacks will impact the grassland ecosystem and the global carbon cycle.

This research is supported by the Office of Science, Office of Biological and Environmental Research, of the U. S. Department of Energy Grant DOE-SC10010566.

56. Correlated Molecular Imaging Applied to Chemically Communicating Microbial Communities of Relevance to the Rhizosphere

Sneha Polisetti[1], Nameera Baig[1], Jennifer Morrell-Falvey[2], Mitchel Doktycz[2], Sage Dunham[3], Nydia Morales-Soto[4], Joshua Shrout[4], Jonathan Sweedler[3], and Paul Bohn[1]

[1] Department of Chemistry and Biochemistry and Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN 46556
[2] BioSciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831
[3] Department of Chemistry and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801
[4] Department of Civil and Environmental Engineering and Earth Sciences and Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556

Project Goal: Our goal is to develop a multi-modal analysis program by integrating confocal Raman microscopy (CRM) and mass spectrometry imaging (MSI – both by SIMS and MALDI) and apply them to a system that closely mimics the rhizosphere. Specifically we are developing a three-component system comprised of two microorganisms and a plant root. We seek to understand the biological interactions, processes, and communication events that occur at the intra and inter species level between the microorganisms as well as the interactions occurring between the plant root and the individual microorganisms by following the spatial and temporal characteristics of the various biomolecules within the system.

http://www.bohnresearchgroup.com

Correlating image information across orthogonal platforms, such as CRM and MSI can provide insights that are superior to those obtained from either modality alone and uniquely positions our team to: (1) develop selective, non-perturbative probes to measure the spatial and temporal concentration profiles of nutrients, metabolites, signaling molecules, extracellular matrices and other biomolecules; (2) develop instrumentation to enable simultaneous observation of different intracellular processes with high spatial and quantitative resolution; and (3) investigate the spatial relationships, physical connections, and chemical dynamics that are responsible for the flow of information and materials at the interface between organisms and their environment.

Correlated CRM-MSI imaging has already uncovered new principles by which low-level signaling molecules, specifically the quinolones, govern the spatiotemporal organization of microbial communities of Pseudomonas aeruginosa. These new chemical imaging tools fundamentally change the nature of the questions that can be posed regarding microbial communities, permitting us to move beyond a taxonomic approach – cataloguing observed compounds and how they relate to gene expression – to an approach that addresses the temporal and structural features of community organization, especially in biologically relevant multi- organismal co-culture systems. Furthermore, correlated CRM-MSI imaging produces information across multiple dimensions, since molecular identities acquired from MSI data are combined with high spatial resolution chemical class information available from CRM. Multiplexing of these two platforms has now reached a stage of maturity where correlated CRM-MSI can be applied to multi-organismal microbe-plant communities under more complex and realistic conditions. We are currently extending the chemical imaging capabilities in two important ways – by enhancing the technological capabilities of correlated imaging and by applying the resulting capabilities to develop critical understanding of spatiotemporal organization in co-cultured models of the rhizosphere.

We have recently extended our correlated CRM-MSI work to the challenging problem of bacterial

biofilms. First, planktonic P. aeruginosa was compared to biofilms using CRM and metal-assisted LDI. Planktonic cells produced CRM spectra dominated by DNA and RNA scattering, while biofilms expressed significant quantities of rhamnolipids, polysaccharides and proteins. Corresponding LDI analysis provided MS/MS identification and spatial mapping for six rhamnolipids, including both Rha-C10 and Rha-C10-C10, which are critical surfactants for biomolecule distribution during biofilm development. Comparing a quorum-sensing deficient mutant to the wild-type strain showed dramatic differences, with the mutant producing little or no rhamnolipid and other extracellular polymeric substances (EPS) under conditions designed to elicit biofilm growth.1 The spatial registry problem was solved for these relatively featureless bacterial biofilms, by printing a nanoparticle fiducial array on the dried biofilm prior to analysis, viz. Figure 1, providing a general scheme for co-localizing spectral features across multiple platforms implementing image acquisitions under disparate conditions.

The authors acknowledge funding from the Department of Energy Office of Science (BER) through grant DE-SC0006642 and the Genomic Science Program, Plant-Microbe Interfaces Project (ORNL).

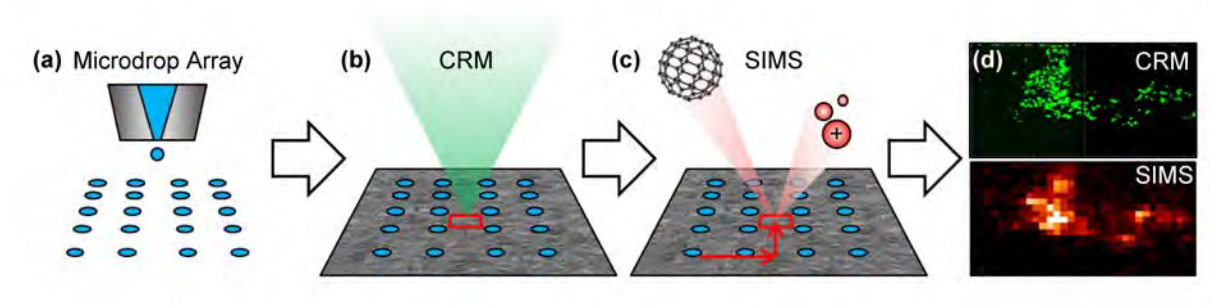


Figure 1. CRM-SIMS correlated imaging workflow. (a) A microdroplet array is applied to the dried biofilm. (b) CRM is performed to locate ROIs and array coordinates are recorded. (c) The sample is transferred to the SIMS instrument and the array is used to navigate back to the ROIs. (d) The CRM and SIMS data are correlated, using the array for alignment.

1. Lanni, E.J.; Masyuko, R.N.; Driscoll, C.M.; Aerts, J.T.; Shrout, J.D.; Bohn, P.W.; Sweedler, J.V., Correlated Imaging with C60-SIMS and Confocal Raman Microscopy: Visualization of Cell-Scale Molecular Distributions in a Bacterial Biofilm. Analyt. Chem. 2014, 85, 10885-10891.

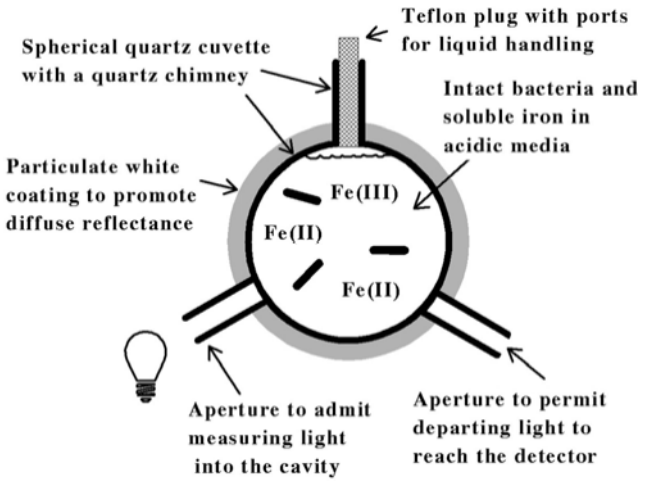
57. A Spectroscopic Device to Monitor Respiratory Electron Transfer in Suspensions of Live Organisms

Robert Blake II,[1] *(rblake@xula.edu), TingFeng Li, [1] Bhupal Ban,[1] and Richard Painter[1]

[1]Xavier University of Louisiana, New Orleans, LA

URL: www.xula.edu

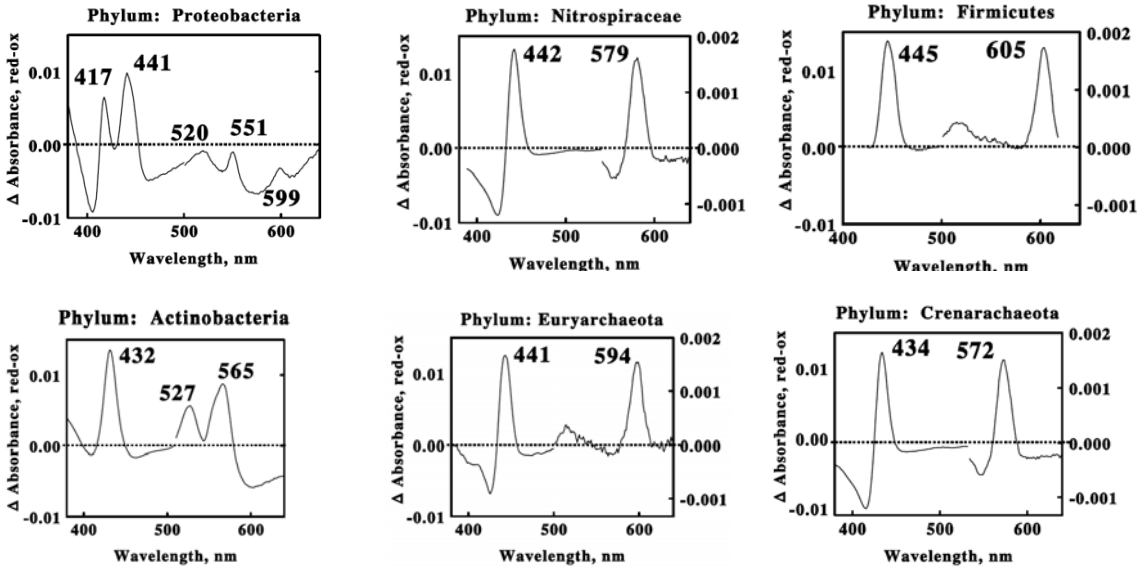
Project Goals: The practical goal is to develop an integrating cavity absorption meter (ICAM) to directly observe respiratory electron transfer reactions in live bacteria as they exchange electrons with soluble or insoluble substrates under physiological conditions. The premise is that accurate UV-visible spectroscopy of electron transfer reactions among colored biomolecules can be conducted in highly turbid suspensions if the live bacteria are irradiated in an isotropic homogeneous field of incident measuring light. We exploited this ICAM to test the hypothesis that acidophilic bacteria in different phyla express different types of electron transfer proteins to respire on extracellular iron. We have also studied the dynamic behavior of the electron transport system at the microbe-mineral interface. The outcome of this project provides a new means to examine the extents and rates of biochemical events in situ without disrupting the complexity of the cellular environment.



An experimental beta unit of an integrating cavity absorption meter was obtained from On Line Instrument Systems (Bogart, GA), in which the cuvette is a reflecting cavity completely filled with the absorbing sus- pension. A simplified model of this ICAM designed to permit accurate absorbance measurements in media that scatter light is shown on the right. The observation cell is comprised of a spherical quartz cuvette fused with a quartz tube to permit access. The chamber is surrounded by a tightly packed white powder that serves to maximize diffuse reflectance of light on the exterior walls of the spherical flask. The apertures in the reflecting sphere through which the measuring light enters and the transmitted/ scattered light exits to the photomultiplier tube are positioned at a 90o angle such that the light must undergo many reflections and cell transversals before it is quantified using the detector.

Initial studies focused on respiratory electron transfer reactions in live bacteria that respire aero- bically on extracellular iron at pH 1.5. Difference spectra such as those shown in Fig. 1 were obtained as soon as the suspension of cells was mixed with soluble ferrous iron. The resting absorbance spectrum of the bacterium observed under air-oxidized conditions was always regenerated from that of the Fe(II)-reduced bacterium initially observed in the presence of Fe(II).

Figure 1. Iron-reduced minus oxidized difference spectra of representative intact microorgan- isms from six different phyla. Spectra were obtained using the ICAM under physiological conditions at pH 1.5. Organisms: Proteobacteria, *Acidithiobacillus ferrooxidans*; Nitrospi- raceae, *Leptospirillum ferriphilum*; Firmicutes, *Sulfobacillus thermosulfidooxidans*; *Actino- bacteria, Ferrimicrobium acidiphilum*; Euryarchaeota, Acidiplasma aeolicum; and Crenarach- aeota, *Acidianus brierleyi*.



In each case, the velocity of product ferric iron formation at any time point was directly proportional to the concentration of the principal reduced biomolecule(s) observed in that particular organism. Further, the integrals over time of the concentrations of the reduced biomolecules were directly proportional to the total concentrations of ferrous iron in each reaction mixture. These kinetic data obtained using whole cells were consistent with the hypothesis that the prominent reduced electron transfer biomolecule(s) observed with each organism were obligatory intermediates in the iron respiratory chains of the respective organisms. The ability to respire aerobically on extracellular iron is currently distributed among 19 genera in 6 different phyla. The different biomolecules observed in Fig. 1 were consistent with the central hypothesis that phylogenetically distinct microorganisms express different types of electron transport biomolecules to achieve respiration on extracellular iron.

There is no better means to establish physiological relevance in a metabolic function than to directly observe it as it occurs in the intact bacterium. The ability to conduct direct spectro- photometric studies under non-invasive physiological conditions represents a new and powerful approach to examine the extents and rates of biological events in situ without disrupting the complexity of the live cellular environment.

This research was supported by the Office of Biological and Environmental Research in the US Department of Energy Office of Science.

58. Correlated Molecular Imaging Applied to Chemically Communicating Microbial Communities of Relevance to the Rhizosphere

Sneha Polisetti[1], Nameera Baig[1], Jennifer Morrell-Falvey[2], Mitchel Doktycz[2], Sage Dunham[3], Nydia Morales-Soto[4], Joshua Shrout[4], Jonathan Sweedler[3], and Paul Bohn[1]

[1] Department of Chemistry and Biochemistry and Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN 46556
[2] BioSciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831
[3] Department of Chemistry and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801
[4] Department of Civil and Environmental Engineering and Earth Sciences and Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556

Project Goal: Our goal is to develop a multi-modal analysis program by integrating confocal Raman microscopy (CRM) and mass spectrometry imaging (MSI – both by SIMS and MALDI) and apply them to a system that closely mimics the rhizosphere. Specifically we are developing a three-component system comprised of two microorganisms and a plant root. We seek to understand the biological interactions, processes, and communication events that occur at the intra and inter species level between the microorganisms as well as the interactions occurring between the plant root and the individual microorganisms by following the spatial and temporal characteristics of the various biomolecules within the system.

http://www.bohnresearchgroup.com

Correlating image information across orthogonal platforms, such as CRM and MSI can provide insights that are superior to those obtained from either modality alone and uniquely positions our team to: (1) develop selective, non-perturbative probes to measure the spatial and temporal concentration profiles of nutrients, metabolites, signaling molecules, extracellular matrices and other biomolecules; (2) develop instrumentation to enable simultaneous observation of different intracellular processes with high spatial and quantitative resolution; and (3) investigate the spatial relationships, physical connections, and chemical dynamics that are responsible for the flow of information and materials at the interface between organisms and their environment.

Correlated CRM-MSI imaging has already uncovered new principles by which low-level signaling molecules, specifically the quinolones, govern the spatiotemporal organization of microbial communities of Pseudomonas aeruginosa. These new chemical imaging tools fundamentally change the nature of the questions that can be posed regarding microbial communities, permitting us to move beyond a taxonomic approach – cataloguing observed compounds and how they relate to gene expression – to an approach that addresses the temporal and structural features of community organization, especially in biologically relevant multi- organismal co-culture systems. Furthermore, correlated CRM-MSI imaging produces information across multiple dimensions, since molecular identities acquired from MSI data are combined with high spatial resolution chemical class information available from CRM. Multiplexing of these two platforms has now reached a stage of maturity where correlated CRM-MSI can be applied to multi-organismal microbe-plant communities under more complex and realistic conditions. We are currently extending the chemical imaging capabilities in two important ways – by enhancing the technological capabilities of correlated imaging and by applying the resulting capabilities to develop critical understanding of spatiotemporal organization in co-cultured models of the rhizosphere.

We have recently extended our correlated CRM-MSI work to the challenging problem of bacterial

biofilms. First, planktonic P. aeruginosa was compared to biofilms using CRM and metal-assisted LDI. Planktonic cells produced CRM spectra dominated by DNA and RNA scattering, while biofilms expressed significant quantities of rhamnolipids, polysaccharides and proteins. Corresponding LDI analysis provided MS/MS identification and spatial mapping for six rhamnolipids, including both Rha-C10 and Rha-C10-C10, which are critical surfactants for biomolecule distribution during biofilm development. Comparing a quorum-sensing deficient mutant to the wild-type strain showed dramatic differences, with the mutant producing little or no rhamnolipid and other extracellular polymeric substances (EPS) under conditions designed to elicit biofilm growth.1 The spatial registry problem was solved for these relatively featureless bacterial biofilms, by printing a nanoparticle fiducial array on the dried biofilm prior to analysis, viz. Figure 1, providing a general scheme for co-localizing spectral features across multiple platforms implementing image acquisitions under disparate conditions.

The authors acknowledge funding from the Department of Energy Office of Science (BER) through grant DE-SC0006642 and the Genomic Science Program, Plant-Microbe Interfaces Project (ORNL).

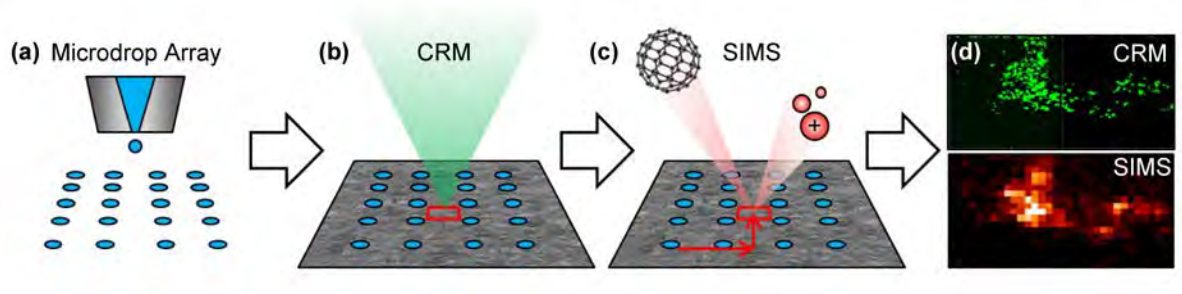


Figure 1. CRM-SIMS correlated imaging workflow. (a) A microdroplet array is applied to the dried biofilm. (b) CRM is performed to locate ROIs and array coordinates are recorded. (c) The sample is transferred to the SIMS instrument and the array is used to navigate back to the ROIs. (d) The CRM and SIMS data are correlated, using the array for alignment.

1. Lanni, E.J.; Masyuko, R.N.; Driscoll, C.M.; Aerts, J.T.; Shrout, J.D.; Bohn, P.W.; Sweedler, J.V., Correlated Imaging with C60-SIMS and Confocal Raman Microscopy: Visualization of Cell-Scale Molecular Distributions in a Bacterial Biofilm. Analyt. Chem. 2014, 85, 10885-10891.

59. How do plant-microbial interactions regulate litter decomposition?

Mary Firestone[1,2*] (mkfstone@berkeley.edu), Shengjing Shi[1], Thea Whitman[1], Erin Nuccio1,[3], Donald Herman[1], Jason Shi[4], Zhili He,[4] Jennifer Pett-Ridge[3], Eoin Brodie[1,2], Jizhong Zhou[4], and Trent Northen[5]

[1] University of California, Berkeley; [2] Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California; [3] Nuclear and Chemical Sciences Division, Lawrence Livermore National Laboratory, Livermore, California; [4] University of Oklahoma, Norman, Oklahoma; [5] Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California.

Project Goals: Our project (Mapping soil carbon from cradle to grave: drafting a molecular blueprint for C transformation from roots to stabilized soil organic C) works towards a fundamental understanding of C cycling in soil as mediated by soil microorganisms and their interactions with plants. How do the interactions between roots and soil microorganisms affect transformations of root derived C, decomposition and loss as CO2, as well as C sorption and stabilization in soil? We seek to gain a mechanistic understanding of the conversion of root-derived C to stabilized soil C, clarify the impacts of microbial activities on soil C sequestration, and substantially expand our understanding of molecular regulation of terrestrial C cycling.

Plants transfer atmospheric CO2 to belowground soil C pools, while microbes mediate C transformation and mineralization in the soil. However, the molecular mechanisms underlying soil-plant-microbial interactions are poorly understood. We have examined the effects of live Avena fatua roots (a common annual grass) on decomposition of 13C-labeled root litter in a California grassland soil over two growing seasons. The presence of live roots consistently suppressed rates of litter decomposition; however this effect disappeared with plant senescence. Presence of live roots significantly altered the abundance, composition and functional potential of microbial communities (assessed by qPCR, MiSeq 16S and ITS sequencing, and GeoChip 4, respectively). Two possible mechanisms – preferential substrate utilization and drying stress – are proposed as factors explaining the observed decrease in soil organic C decomposition in the presence of live roots. A simple model based on our recently collected data indicates that the decomposition of litter is primarily controlled by microbial functional potential. We are currently examining the molecular mechanisms underlying altered decomposition rates in the rhizosphere using metatranscriptomic analyses over the life of a root in the presence and absence of litter.

Meanwhile, we have developed a functional gene oriented pipeline for rapid processing and annotation of large metagenomic/metatranscriptomic reads with key functional gene information relevant to geochemical and ecological processes. The pipeline is being applied to investigate the functional gene composition and structure of soil microbial communities and will continue to be validated and benchmarked throughout our project. We are also profiling the spatial dynamics of enzymatic activity surrounding a plant root through the development of "Nimzyme" imaging technology. This approach uses nanostructure-initiator mass spectrometry (NIMS) technology to image in situ enzyme activity in intact systems. This type of spatial mapping of enzyme activities combined with transcriptome-based, genome-based, and proteome-based analyses of rhizosphere soil promises to reveal the bases of root-induced changes in microbial decomposition processes.

Funding for this research was provided by awards SA-DOE-29318 and DE-SC0010570 through the Genomic Sciences program of the Office of Biological and Environmental Research, U.S. Department of Energy to the University of California, Berkeley. Part of this work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344 and FWP SCW 1421 and at Lawrence Berkeley National Laboratory under the auspices of the University of California - contract DE- AC02-05CH11231.

60. Mapping soil carbon from cradle to grave: Identifying metabolic processes underlying patterns of rhizosphere microbial succession and organic matter priming

Kateryna Zhalnina[1,2*](kzhalnina@lbl.gov), Heejung Cho[1], Ulas Karaoz[2], Stefan Jenkins[2], Richard A. White III[3], Mary S. Lipton[3], Jizhong Zhou[4], Jennifer Pett-Ridge[5], Trent R. Northen[2], Mary K. Firestone[1,2], and Eoin L. Brodie[1,2]

[1]University of California, Berkeley, California; [2]Lawrence Berkeley National Laboratory, Berkeley, California; [3]Pacific Northwest National Laboratory, Richland, Washington; [4]University of Oklahoma, Norman, Oklahoma; [5]Lawrence Livermore National Laboratory, Livermore

Project Goals: Our project (Mapping soil carbon from cradle to grave: drafting a molecular blueprint for C transformation from roots to stabilized soil organic C) works towards a fundamental understanding of C cycling in soil as mediated by soil microorganisms and their interactions with plants. How do the interactions between roots and soil microorganisms affect transformations of root derived C, decomposition and loss as CO2, as well as C sorption and stabilization in soil? We seek to gain a mechanistic understanding of the conversion of root-derived C to stabilized soil C, clarify the impacts of microbial activities on soil C sequestration, and substantially expand our understanding of molecular regulation of terrestrial C cycling.

Interactions between plants, soil microorganisms and soil minerals underlie both the stabilization and decomposition of soil organic carbon (SOC). Plants increase the concentration of labile C in soil through exudation, providing readily assimilated substrates for rhizosphere microbiota and stimulation or repression of SOC mineralization may result (i.e. positive or negative priming). The soil microbiome displays strong patterns of succession in response to living plant roots and although both positive and negative rhizosphere priming have been widely observed, the metabolic processes underlying these successional patterns and SOC priming are not well defined. Using combined modeling and experimentation we are working to define key traits of soil microorganisms relevant to their fitness in the rhizosphere and transformation of carbon. From a library of approximately 300 bacterial isolates we sequenced 38 heterotrophic bacteria representative of the dominant organisms identified in metagenomes of this Mediterranean grassland soil. Analysis of genome properties and content allowed us to predict important physiological features of the isolates, such as minimum generation times (MGTs), optimum temperature and substrate utilization capabilities. MGTs inferred directly from genomic sequences agreed well with in vitro observations. All analyzed genomes revealed a repertoire of features relevant to life in the rhizosphere, such as high gene copy number of sugar and organic acid transporters and polymer degrading enzymes (glycoside hydrolases and polyphenoloxidases). Isolates from the Alphaproteobacteria and Actinobacteria had the highest occurrence of these features in their genomes. To test genomic predictions we have performed a number of functional analyses using enzymatic, exometabolomic and proteomic assays.

We have developed and curated draft genome-scale metabolic models of select isolates and have used exometabolomic data to perform gap filling. These metabolic models will be used to evaluate via simulation the metabolic response of soil heterotrophic bacteria to root exudates and their impact on rhizosphere succession and SOC mineralization. Explicit features of the genome-scale models (electron donors and acceptors, C use efficiency, metabolic feedbacks and trade-offs) together with experimental data (gene expression, protein secretion, metabolite uptake and release) will be used in future work to parameterize genome-informed trait-based models of the soil microbial community aimed at predicting succession and C transformation in the rhizosphere.

Funding for this research was provided under contract FOA DE-PS02-09ER09-25 through Genomic

Sciences program of the Office of Biological and Environmental Research, U.S. Department of Energy to the University of California, Berkeley. Part of this work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344 and at Lawrence Berkeley National Laboratory under the auspices of the University of California - contract DE-AC02-05CH11231.

61. Tracing the Influence of Mineralogy, Microbiology, and Exudate Chemistry on the Stabilization of Root-Derived Carbon

Rachel Neurath[1*](rneurath@berkeley.edu), Marco Keiluweit[2], Jennifer Pett-Ridge[4], Peter Nico[3], and Mary Firestone[1]

[1] University of California, Berkeley; [2] School of Earth Sciences, Stanford University; [3] Earth Sciences Division, Lawrence Berkeley National Laboratory; [4] Chemical Sciences Division, Lawrence Livermore National Laboratory

Project Goals: Our project (Mapping soil carbon from cradle to grave: drafting a molecular blueprint for C transformation from roots to stabilized soil organic C) focuses on a fundamental understanding of C cycling in soil as mediated by soil microorganisms and their interactions with plants. Our work investigates how the interactions between soil minerals and microorganisms affect C sorption and stabilization in soil, particularly in the rhizosphere. Through our research we seek to provide a mechanistic understanding of the conversion of root-derived C to stabilized soil C, clarify the impacts of microbial activities on soil C sequestration, and substantially expand our understanding of molecular regulation of terrestrial C cycling.

Mineral surfaces provide sites for carbon stabilization in soils, protecting soil organic matter (SOM) from microbial degradation for up to thousands of years (Mikutta et al., 2006). While plants provide the primary supply of carbon to soils, minerals are central controllers of the ultimate fate of that carbon. Free plant carbon in soil may quickly be metabolized by microbial communities and released back to the atmosphere (CO2, CH4), while carbon that associates with mineral surfaces has the potential for long-term storage. However, we know that not all mineral-associated carbon is protected from microbial degradation for long time periods. Our previous work tracking plant carbon into soil under elevated atmospheric CO2 showed that while soil minerals initially stabilized new plant-derived carbon, naturally occurring plant and soil dynamics resulted in loss of carbon from soil minerals and a subsequent pulse of CO2 to the atmosphere. Thus, soil minerals hold a conditional potential for long-term carbon stabilization.

In more recent experiments, we followed the fate of 13C-labeled plant-derived carbon in soil microcosms incubated with three minerals representing a spectrum of structure and reactivity: quartz, kaolinite, and ferrihydrite-coated quartz. These minerals, which were isolated in mesh bags to exclude plant roots but not microorganisms, were extracted and measured for total C and 13C atom% after 1, 2, and 2.5 months incubation. At 2.5 months, the quartz had the least mineral-bound C (0.40 mg-g-1) and ferrihydrite the most (0.78 mg-g-1). Ferrihydrite and kaolinite also accumulated more plant-derived carbon (3.0 and 3.1%, respectively) based on 13C isotopic tracing. These results, which we coupled with Fourier Transform Infrared Spectroscopy (FTIR) and 13C-Nuclear Magnetic Resonance Spectroscopy (NMR) analysis of the chemical character of mineral-associated SOM, suggest that more reactive minerals have a higher potential to associate with simpler compounds, which otherwise may have been easily degraded by microbes. Through extraction of cells from the mineral surfaces and preliminary quantification of DNA (PicoGreen), we see that on all minerals, a significant portion of the associated carbon is from living or recently living microbial bodies, emphasizing the importance of microbial biomass not only in the breakdown of SOM, but as a component of SOM.

Despite the importance of plant roots as a conduit for carbon substrates that may eventually become mineral stabilized, certain root exudates may actually destabilize mineral associated SOM, liberating carbon for microbial degradation. We conducted a study with 13C-labeled individual exudate solutions pumped through artificial roots into soil rhizotrons. With the addition of the common root exudate oxalic

acid, 75-95% oxalic acid C was released as CO2. Accounting for carbon use efficiency (CUE) and expected biological oxygen demand (BOD), 50% of CO2 released was not, in fact, from oxalic acid metabolism, but rather from priming of carbon already in the soil. Analysis of mineral-organic complexes and soil pore water (NanoSIMS, STXM/NEXAFS, and LDPI-MS) indicate this carbon was 'old carbon' released from mineral surfaces. Thus, despite the importance of plants in supplying carbon to soils, 'priming' may release carbon from mineral-organic complexes, such that it is microbially available. Our research shows that both soil mineralogy and the chemical character of plant-derived compounds are important controls of mineral stabilization and destabilization of soil carbon.

References

1. Mikutta, R., M. Kleber, M.S. Torn, R. Jahn. Stabilization of soil organic matter: Association with soil minerals or chemical recalcitrance? 2006. Biogeochemistry, 77:1, p. 25-56.
2. Keiluweit, M., J.J. Bougoure, J. Pett-Ridge, P.K. Weber, P.S. Nico, M. Kleber. Root exudates counteract mineral control on soil carbon turnover. 2015. Nature Climate Change, in press.

Funding

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 SA-DOE- 29318 and DE-SC0010570 at UC Berkeley and FWP SCW1421 at LLNL. Funding was also provided by an LLNL LDRD "Microbes and Minerals: Imaging C Stabilization" and a Lawrence Scholar Fellowship to M. Keiluweit. Work at LLNL was performed under the auspices of the U.S. Department of Energy under Contract DE-AC52-07NA27344.

62. Key microbial decomposition pathways in soils from bioenergy agroecosystems.

Kirsten Hofmockel[1*](khof@iastate.edu), Ryan Williams[1], Adina Chuang Howe[1]

[1]Iowa State University, Ames, IA

Project Goals: Because of the amazing diversity of soil, it has been identified as a "Grand Challenge" for exploring microbial diversity. Here we investigate the diverse metabolic potential of soil microbial communities involved in decomposition by identifying key functions and organisms involved in cellulose decomposition. Our goal is to develop effective strategies to analyze soil metagenomes allowing exploration of the biological basis for decomposition. This work is being done in soils cultivated for bioenergy feedstocks, where we aim to optimize plant-microbe- soil interactions for the development of sustainable biofuel production systems.

The Comparison of Biofuel Systems (COBS) research site in central Iowa, USA is a replicated field experiment managed to investigate the potential for alternative biomass production systems to produce fossil---fuel replacement and concurrently support biogeochemical benefits. Corn continues to be the most readily abundant and available crop for biofuel production in the United States1. However the current corn---soy management that dominates the Corn Belt is no longer sustainable in the context of climate feedbacks, demand for clean water, and the requirement for cellulosic biofuel. One promising solution is to incorporate diversified biofuel cropping systems that provide multiple ecological benefits coincident with provisioning harvestable cellulosic biomass. The potential benefits of diversified agroecosystems are derived from the known relationship between biodiversity and the productivity and sustainability of Earth's ecosystems2,3,4,5. Diversified biofuel cropping systems may augment microbial diversity and the associated biogeochemical benefits, including enhanced decomposition of plant detritus, increased soil aggregation and reduced loss of soil organic matter. Yet the biological basis for how plant and microbial diversity interact to influence soil carbon (C) storage remains an important frontier.

Identifying the microbial metabolic pathways that most strongly influence ecosystem C cycling requires a deeper understanding of the relationships between microbial community functions in a field context. However, characterizing these community--scale functions requires adoption of a scale relevant to microbial habitats in soil. Soil aggregate fractions can be used to enhance the ability to detect metabolic relationships, because aggregates approach the scale of ecological drivers that structure microbial communities. The present study addresses the composition and functional potential of C cycling microbial communities using soil aggregate fractions from diversified, reconstructed, fertilized prairies that are harvested for bioenergy feedstock. We posit that the most abundant sequences may not capture the central C cycling steps regulating decomposition and soil organic matter formation. Because C---cycling occurs in soil as the concerted action of multiple, interacting microbial groups, we developed a novel co---occurrence workflow that uses Bayesian hierarchical methods to identify community-scale relationships that may drive C---cycling. In addition to producing multivariate networks of co-occurring microbial taxa and functions, this methodology allows for greater consideration of uncertainty between co---occurrence relationships and provides a straightforward framework for testing our pathways, we concentrated this analysis on protein---coding genes identified with the Carbohydrate Active Enzyme (CAZy) database6. With a focus on glycoside hydrolases (GH's), we saw many families of unclassified enzymes positively co-occurring with families of well-known enzymes. For example, we detected a strong and positive co-occurrence between the abundance of the xyloglucanase family, GH12, and a poorly characterized β-xylosidase in family GH54, that may be important for breaking down complex, structural plant biomass. Co-occurrence between these functions illustrates complementary functions between multiple bacterial phyla and the fungal phylum, Ascomycota. Together these results suggest that the presence of keystone organisms or functions may be central to initiating decomposition in soils

and producing a diversity of C-rich substrates that can maintain diversity and coexistence across multiple domains of microbial life.

In addition to co-occurrence strength, position in the network can be used to generate hypotheses about metabolic dependencies. Our analysis suggests that genes in GH52 and GH2 may be pivotal to decomposition in our soils. A GH52 (xylosidase) present only in γ-Proteobacteria was central in the network, supporting a high dependency of other enzymes. Notably, a variety of proteins encoding for GH2, β---acting enzymes (galactosidase, mannosidase, glucoronidase, glucosaminidase) were strongly correlated with the γ-Proteobacteria associated xylosidase. While some functions like those involving GH52 were phylogenetically limited, many other co---occurring functions involved a larger diversity of microorganisms. These analyses have provided key phylogenetic and functional targets to monitor in both field and lab based studies of microbial decomposition in feedstock production systems.

References:

1. US DOE (2011) U.S. Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry, Oak Ridge, TN, Oak Ridge National Laboratory. Available at: http://www1.eere.energy.gov/bioenergy/pdfs/billion_ton_update.pdf
2. Loreau, M., Naeem, S. & Inchausti, P. Biodiversity and Ecosystem Functioning: Synthesis and perspectives (Oxford Univ. Press, 2002).
3. Hooper, D. U. et al. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecol. Monogr. 75, 3–35 (2005).
4. Tilman, D. Ecological consequences of biodiversity: a search for general principles. Ecology 80, 1455–1474 (1999).
5. Wardle, D. A., Bardgett, R. D., Callaway, R. M. & Van der Putten, W. H. Terrestrial ecosystem responses to species gains and losses. Science 332, 1273–1277 (2011).
6. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Research 42: D490-D495 (2013).

This research was supported by the Genomic Science Program, U.S. Department of Energy, Office of Science, Biological and Environmental Research under Contract No. DESC0010775.

63. Identification of genomic elements required for uranium resistance by Caulobacter crescentus

Mimi C. Yung,[1*] (yung6@llnl.gov), Dan M. Park,[1] Cindi A. Hoover,[2] Matthew J. Blow,[2] and Yongqin Jiao[1]

[1] Biosciences and Biotechnology Division, Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA
[2] DOE Joint Genome Institute, Walnut Creek, CA

(Work at LLNL conducted under Contract DE-AC52-07NA27344. LLNL-ABS-667012.)

Project Goals: Uranium (U), particularly in its water soluble form U(VI), has been shown to be a significant environmental toxin and a major contaminant at DOE legacy sites. One potential method to remediate U contamination is to use microbes that are highly tolerant of U and are able to decrease its mobility in the environment. Caulobacter crescentus is an aerobic, aquatic bacterium that is highly tolerant of U and has the potential to be used for U bioremediation especially in oxic zones. The overall aim of our studies is to understand the mechanisms by which Caulobacter crescentus tolerates and immobilizes U in the environment using a combination of genetic, biochemical, and omic approaches.

Caulobacter crescentus is a ubiquitous, aerobic bacterium known to survive in nutrient limited environments and to tolerate high levels of U(VI). The detoxification mechanisms by which C. crescentus resists U are poorly understood. Here, we show that C. crescentus is able to facilitate U (VI) biomineralization through the formation of U-Pi precipitates via its native alkaline phosphatase activity. The U-Pi precipitates, deposited on the cell surface in the form of meta-autunite structures, have a lower U/Pi ratio compared to U precipitates formed abiotically. The enzyme that is responsible for the phosphatase activity and thus the biomineralization process is identified as PhoY, a periplasmic, alkaline phosphatase with broad substrate specificity. PhoY is shown to confer a survival and growth advantage to C. crescentus under U (VI) stress. These results highlight U(VI) biomineralization as a resistance mechanism in aerobic microbes with potential applications for bioremediation in the environment.

In order to gain a further understanding of how C. crescentus resist U, we employed a transposon mutagenesis screening approach (Tn-seq) to identify essential genomic elements that are required for U resistance. In our method, highly saturated transposon (Tn) mutagenesis was first performed to generate a library of 106 mutants. The library was then grown on solid agar plates containing U, Cd, or no metal control. Mutants surviving each exposure were subsequently sequenced via high-throughput Illumina sequencing at JGI to identify Tn insertion sites. Genes that accumulated fewer Tn insertions under U stress compared to the Cd or no stress controls were identified as genes specific for U tolerance. Using this method, we identified 15 genes potentially involved in U tolerance, which were subsequently tested through mutational analysis. Genes identified to be involved in U tolerance include those for TolC-like transporters RsaFa and RsaFb, previously identified to be involved in S-layer protein transport, and stress factors CztR and CztA. These key genes provide important insight into the various resistance pathways employed by C. crescentus for survival under U stress.

64. Uranium exposure elicits a multifaceted stress response in Caulobacter crescentus

Dan M. Park*, Mimi Yung, Celena Carriilo, Pejman Naraghi-Arani, Yongqin Jiao

Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA 94550

*park36@llnl.gov

(Work at LLNL conducted under Contract DE-AC52-07NA27344. LLNL-ABS-648404.)

Project Goals. The oxidized form of uranium (U(VI)) predominates in oxic environments and poses a major environmental threat due to its high toxicity and mobility.

Microorganisms play an important role in governing U speciation and transport in the environment and represent a promising U bioremediation platform. Although mechanisms of microbial-mediated U(VI) immobilization have been well-studied, information regarding the specific mechanisms involved in U(VI) toxicity remains limited. A major aim of our project is to understand the mechanisms by which the obligate aerobe Caulobacter crescentus mitigates U toxicity in the environment. To this end, we have applied a functional genomics approach to identify genomic elements critical for uranium tolerance.

To gain insight into the mechanism of U(VI) toxicity, we characterized the stress response pathways critical for uranium tolerance in the obligate aerobe Caulobacter crescentus that is of interest for bioremediation because of its ability to mineralize U(VI) under aerobic conditions. .RNA-seq and proteomics were used to identify upregulated stress response pathways during exposure to toxic concentrations of U, with the contribution of such pathways to U tolerance assessed by gene deletions. Among the differentially expressed genes include several global stress regulators and regulon members, including those responding to ROS, DNA damage, heat shock and extracytoplasmic stresses. In particular we found that a recA deletion mutant was hypersensitive to uranium exposure. Intriguingly this phenotype was independent of SOS induction since a recA mutant lacking the ability to induce the SOS response did not exhibit increased U sensitivity, suggesting that the homologous recombination function of RecA is critical for U tolerance. Additionally, the loss of ClpAP and DegP protease function also significantly reduced growth rate during U exposure, as did a σ32 partial loss of function allele. Consistent with an oxidative stress response, U shifted the GSSG/GSH redox couple towards the oxidized state in a concentration dependent manner and the loss of key antioxidant enzymes, including superoxide dismutase (sodB), catalase/peroxidase (katG ahpF double mutant) and glutathione reductase (CCNA_02387) reduced U tolerance. Finally, our data revealed a link between uranium toxicity and cell cycle progression, since nearly the entirety of the Caulobacter cell cycle machinery was downregulated and protein levels of the two master cell cycle regulators, CtrA and DnaA, were depleted upon exposure to U. Together, our data suggest that uranium exerts a multifaceted toxicity by damaging DNA, disrupting cytoplasmic and extracytoplasmic protein homeostasis and by causing oxidative stress.

References

1. Hu P, Brodie EL, Suzuki Y, McAdams HH, Andersen GL. 2005. Whole-genome transcriptional analysis of heavy metal stresses in Caulobacter crescentus. Journal of bacteriology 187:8437-8449.

65. New Insights into the Factors Defining the Composition of Methane-oxidizing Communities and the Role of Oxygen

Ludmila Chistoserdova[1*] (milachis@uw.edu), Maria E. Hernandez[1], Sascha Krause[1], David A.C. Beck[1,3], and Mary E. Lidstrom[1,2]

[1] Departments of 1Chemical Engineering and [2] Microbiology and [3]eScience Institute, University of Washington, Seattle, USA

Project Goals: This project focuses on developing model microbial communities active in methane utilization. Two approaches are utilized: a top-down approach that employs natural communities from environmental samples, and a bottom-up approach that employs axenic cultures of model bacteria. Through manipulations of these communities, we are striving to understand the molecular mechanisms that form a basis for specific species interactions in microbial oxidation of methane. Transcriptomics and metabolomics are among the approaches that are used to identify candidate factors involved in observed co-occurrence patterns. Hypotheses validation will be approached via strain and community manipulations, such as site-specific mutagenesis or introduction of synthetic functional modules into the specific partner strains. Data integration will allow for metabolic and ecological modeling of such communities.

In the current phase of the project, we are addressing the nature of the proposed relationships among different physiological groups of microbes involved in methane utilization and the mechanisms underlying their specificity. In order to obtain mechanistic details into specific species interaction as part of microbial methane oxidation, we are now using both top-down and bottom-up community manipulation approaches. In the first, we manipulate natural communities using methane as the only carbon source and monitor complex community deconvolution toward dramatically simplified, low-species communities. In the second, we use pure cultures of bacteria, all originating from Lake Washington sediment, all with sequenced genomes, to build and manipulate simple synthetic communities. Data from the top-down approach demonstrate that, under methane pressure, the complex natural communities simplify rapidly, with a Methylococcaceae species becoming one dominant species, and a Methylophilaceae species becoming the second dominant species. However, different species persist under different conditions, one major factor being oxygen availability. At high oxygen tensions, the major players are, respectively, species of the genera Methylosarcina and Methylophilus. At low oxygen tensions, when communities experience periods of hypoxia, the major players are, respectively, species of the genera Methylobacter and Methylotenera. These data suggest that oxygen availability is a major factor determining specific partnerships in methane oxidation. Data from the bottom-up approach demonstrate that two randomly selected species of Methylococcaceae and Methylophilaceae isolated from the same environment do not readily form stable communities and that additional partners may be required, such as Flavobacteriaceae. At the same time, two-species communities were observed when both partners originated from naturally selected pairs of Methylococcaceae and Methylophilaceae. In these cases as well, oxygen appeared to play a major role in community composition and behavior. Data from both types of experiments strongly suggest that speciation within Methylococcaceae and Methylophilaceae may be responsible for niche adaptation, driven by factors such as oxygen availability. These results suggest that specific ecotypes of, respectively, Methylococcaceae and Methylophilaceae are tailored to form pairs that possess competitive advantage in specific conditions, such as the placement within the methane/oxygen counter gradient.

Publications

McTaggart TL, Beck DAC, Setboonsarng U, Shapiro N, Woyke T, Lidstrom ME, Kalyuzhnaya MG, Chistoserdova L. Genomics of methylotrophy in Gram-positive methylamine-utilizing species.

Microorganisms, Submitted.

Hernandez ME, Beck DAC, Lidstrom ME, Chistoserdova L. Oxygen availability is a major factor in determining the composition of microbial communities involved in methane oxidation. PeerJ 2015, In Press.

Beck DAC, McTaggart TL, Setboonsarng U, Vorobev A, Kalyuzhnaya MG, Goodwin L, Shapiro N, Woyke T, Lidstrom ME, Chistoserdova L. Multiphyletic origins of methylotrophy in Alphaproteobacteria, exemplified by comparative genomics of Lake Washington isolates. Environmental Microbiology 2015, In Press.

Oshkin I, Beck DAC, Lamb AE, Tchesnokova V, Benuska G, McTaggart TL, Kalyuzhnaya MG, Dedysh S, Lidstrom ME, Chistoserdova L. Methane fed microcosms show differential community dynamics and pinpoint specific taxa involved in communal response. ISME Journal 2014, Online.

Beck DAC, McTaggart TL, Setboonsarng U, Vorobev A, Kalyuzhnaya MG, Ivanova N, Goodwin L, Woyke T, Lidstrom ME, Chistoserdova L. The expanded diversity of Methylophilaceae from Lake Washington through cultivation and genomic sequencing of novel ecotypes. PLoS One 2014, 9:e102458.

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-0010556.

66. Denitrification in the Absence of Bacterial Nitrite Reductases

Steven A. Higgins[1], Jenny Onley[1], Sukhwan Yoon[1], Luis Orellana-R[2], Konstantinos Konstantinidis[2], Joanne Chee-Sanford[3], Robert A. Sanford[4], and Frank E. Löffler[1,5*] (frank.loeffler@utk.edu)

[1]University of Tennessee, Knoxville, TN; [2]Georgia Institute of Technology, Atlanta, Georgia; [3]US Department of Agriculture-Agricultural Research Service, Urbana, IL; [4]University of Illinois, Urbana, IL; [5]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

Project Goals: The goals of this project are to fill existing knowledge gaps in our understanding of N-flux and associated C-turnover in soils and sediments. Novel information about the diversity, distribution, abundance and expression of genes contributing to N-transformation is required to link desirable (i.e., N-retention) and undesirable (i.e., N2O emission) activities with measurable microbial parameters. Correlating molecular- and organismal-level information with environmental factors that control N- and C-turnover can predict the impact of land management practices on greenhouse gas (N2O, CO2) emissions. Such integrated approaches generate novel information at multiple scales of resolution and contribute to system-level understanding of key nutrient cycles in soils. In the present work, we design tools to assess the contribution of fungi to N2O production in soils, and demonstrate that bacteria without the nitrite reductase genes nirS or nirK contribute to denitrification.

Abstract: The overuse of synthetic, nitrogen-based fertilizers stimulates microbial N- turnover, effectively accelerating the global N-cycle. Bacterial denitrification is generally considered the major pathway for nitrate turnover under anoxic conditions although recent findings suggest that respiratory ammonification can predominate under certain geochemical conditions (1). Monitoring nirS or nirK genes, which encode the nitrite reductases, is often employed to assess denitrification potential. Interestingly, Anaeromyxobacter dehalogenans, a Deltaproteobacterium lacking nirS or nirK, reduced nitrite to dinitrogen gas in medium amended with ferric iron and acetate as electron donor. Detailed studies revealed that ferrous iron, the product of ferric iron reduction, chemically reduced nitrite to N2O (i.e., chemodenitrification). Anaeromyxobacter dehalogenans strains possess a functional atypical nosZ gene conferring the ability to reduce N2O to dinitrogen (2), and the cultures reduced N2O to dinitrogen. Despite the absence of key denitrification genes, A. dehalogenans contributed to the conversion of nitrite to dinitrogen in the presence of ferric iron. Ferric iron minerals are common to soils and sediments suggesting that coupled abiotic-biotic processes (i.e., chemodenitrification followed by enzymatic N2O reduction) contribute to N-cycling and affect N2O flux. These findings also demonstrate that the role of a microorganism in N- cycling cannot be predicted based on gene content alone, and ferric iron reducers lacking nitrite reductase genes can contribute to denitrification.

Fungal activity is another source of N2O in soils; however, the involvement of fungi in soil geochemical transformations is poorly understood. Filamentous Ascomycetes and Basidiomycetes are key contributors to the degradation of plant-derived organic matter (e.g., lignin, cellulose), and affect soil carbon flux. Further, members of these taxa have been implicated in nitrate/nitrite reduction to N2O (3). Though relevant contributions of fungi to C- and N-turnover in soils have been recognized, molecular tools to selectively target fungal taxa and their functional genes involved in geochemical cycling are lacking. To address these shortcomings, more than 200 fungal isolates capable of nitrate or nitrite reduction were obtained from two distinct soil sites in Illinois. Degenerate PCR primers amplifying an approximately 650-bp fragment of the fungal p450nor gene, which is responsible for N2O production, were designed. The primer set amplified p450nor genes from many denitrifying fungal isolates, and amplicons were also obtained from soil DNA extracts. Sequence analysis confirmed specific amplification of p450nor gene fragments. This PCR tool is now being applied to assess the dynamics of fungal p450nor genes over spatial and temporal scales at the two study sites. To assess the diversity and dynamics of soil fungi from

the two different soil types, automated ribosomal intergenic spacer analysis (ARISA) was conducted on samples collected over defined spatial and temporal scales (e.g., with depth and across seasons). The average number of ARISA fragment sizes representing unique operational taxonomic units (OTUs) was higher in the well- drained sandy soil (n=95) than in the clay-containing silt loam (n=76). Overall fungal communities were significantly distinct across soil depths at any time of year, but assemblages shifted seasonally with depth. qPCR assays targeting the ribosomal intergenic spacer regions of 15 fungal morphotypes were designed and are being applied in concert with the p450nor primers to assess fungal population dynamics in terms of abundance and denitrification potential. This study provides new tools to assess soil denitrification and associated C-turnover in the absence of nirS or nirK in soil ecosystems.

References:

1. Yoon, S., C. Cruz-García, R. Sanford, K. M. Ritalahti, and F. E. Löffler. 2014. Denitrification versus ammonification: environmental controls of two competing respiratory nitrate/nitrite reduction pathways in Shewanella loihica. ISME J. doi: 10.1038/ismej.2014.201.
2. Robert A. Sanford, Darlene D. Wagner, Qingzhong Wu, Joanne C. Chee-Sanford, Sara H. Thomas, Claribel Cruz-García, Gina Rodríguez, Arturo Massol-Deyá, Kishore K. Krishnani, Kirsti M. Ritalahti, Silke Nissen, Konstantinos T. Konstantinidis, and Frank
3. E. Löffler. Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils. Proc. Natl. Acad. Sci. U. S. A. 109: 19709-19714.
4. Shoun H, Fushinobu S, Jiang L, Kim S-W, Wakagi T. 2012. Fungal denitrification and nitric oxide reductase cytochrome P450nor. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 367:1186-94.

Funding Statement: This research was supported by the US Department of Energy, Office of Biological and Environmental Research, Genomic Science Program, Award DE-SC0006662.

67. Detecting and quantifying genes in soil short-read metagenomes: implications for the Nitrogen cycle.

Luis H Orellana (lhorellana@gatech.edu),[1*] Luis Rodriguez-R,[2] Joanne Chee-Sanford,[3] Robert Sanford,[4] Frank Löffler,[5,6,7] and Konstantinos Konstantinidis[1,2]

[1]School of Civil and Environmental Engineering, [2] and School of Biology, Georgia Institute of Technology, Atlanta, Georgia; [3]US Department of Agriculture—Agricultural Research Service, Urbana, Illinois; [4]Department of Geology, University of Illinois, Urbana, Illinois; [5]Department of Microbiology, University of Tennessee, Knoxville, Tennessee; [6]Bioscience Division, Oak Ridge National Laboratory, Oak Ridge; and Department of Civil and [7]Environmental Engineering, University of Tennessee, Knoxville, Tennessee.

Project Goals:
The goals of this project are to fill existing knowledge gaps in our understanding of N- flux and associated C-turnover in soils and sediments. Novel information about the diversity, distribution, abundance and expression of genes contributing to N- transformation is required to link desirable (i.e., N-retention) and undesirable (i.e., N-loss such as that associated with N2O emissions) activities with measurable microbial parameters. Linking molecular- and organismal-level information with environmental factors that control N- and C-turnover can predict the impact of land management practices on greenhouse gas (N2O, CO2) emissions. Such integrated approaches generate novel information at multiple scales of resolution and contribute to system-level understanding of key nutrient cycles in soils. In the present work, we developed and applied a bioinformatic strategy that calculates the best thresholds for detecting gene fragments related to specific genes in short-read metagenomes and use it to quantify the abundance of N-cycle genes in important soil ecosystems for bioenergy crop production in Midwest US across the year.

Abstract:
Metagenomics studies have shed light onto the natural diversity and abundance for the microbial genetic potential participating in key environmental processes. However, accurate thresholds that can discriminate between true and false positive results from commonly used similarity search methods are rarely discussed nor evaluated. To overcome these difficulties, we developed a methodology aimed to identify position- specific most-discriminant thresholds from manually curated reference datasets. Our method applies Receiver-Operator Curve (ROC)-analysis to metagenomic datasets generated in silico to simulate current short-read sequencing technologies. From the evaluation of true and false positive matches from similarity searches using simulated datasets, models with the best position-specific most-discriminating bitscore threshold for nitrogen cycle genes were constructed and evaluated. The combination of a similarity search algorithms and our strategy, showed an improved false discovery rate (FDR; from ~4 to 14 times) when compared to traditional arbitrary fixed e-value strategies. This strategy also showed to have better sensitivity (average increase of ~24%) compare to hidden markov model searches, although at the expense of more computational resources. In order to apply our tool to real metagenomic samples, we analyzed short-read metagenomes obtained from two agricultural sites with contrasting soil textures (sandy versus silty-loam) during four seasons in 2012 at two depths: surface (0-5cm) and deep (20-30 cm). Most nitrogen cycling genes (e.g., nosZ, amoA and nirK among others) varied in abundance over the course of the year. For instance, a remarkably high abundance of metagenomic reads related to atypical nosZ (reduction of N2O to N2) sequences were observed over the year, accounting for approximately 90% of the total nosZ reads found in both soil layers. Approximately 12% of the nosZ reads were taxonomically assigned to the Anaeromyxobacter genus, indicating their potential relevance for N2O reduction. In addition, six amoA (ammonia oxidation) genes, each encoded by distinct archaeal and bacterial populations, became abundant in the deep sandy samples when seasonal nitrogen

fertilization was applied. This study provides a bioinformatic tool for reliable detection of target short gene fragments in metagenomes and advances our understanding of the abundance and diversity of the nitrogen cycle genetic potential found in different ecosystems. Our publicly available pipeline "ROCker" is fully automated, freely available, and can be used to investigate any other genes or processes of interest (www.enve-omics.gatech.edu).

References:

1. Robert A. Sanford, Darlene D. Wagner, Qingzhong Wu, Joanne C. Chee-Sanford, Sara H. Thomas, Claribel Cruz-García, Gina Rodríguez, Arturo Massol-Deyá, Kishore K. Krishnani, Kirsti M. Ritalahti, Silke Nissen, Konstantinos T. Konstantinidis, and Frank E. Löffler. 2012. Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils. Proc. Natl. Acad. Sci. U. S. A. 109: 19709–14.
2. Orellana, L. H., L. M. Rodriguez-R, S. Higgins, J. C. Chee-Sanford, R. A. Sanford, K. M. Ritalahti, F. E. Löffler, and K. T. Konstantinidis. 2014. Detecting nitrous oxide reductase (nosZ) genes in soil metagenomes: method development and implications for the nitrogen cycle. mBio 5: e01193-14.

Funding Statement: This research was supported by the US Department of Energy, Office of Biological and Environmental Research, Genomic Science Program, Award DE-SC0006662.

68. Genome-scale Metabolic Modeling of Microbial Communities at Sites Undergoing Natural and Stimulated Attenuation of Uranium

Jiao Zhao1, Christopher M Gowen*1 (chris.gowen@utoronto.ca), Kelvin Li2, Monika Bihan2, Radhakrishnan Mahadevan1, and Barbara A. Methé2

Department of Chemical Engineering and Applied Chemistry, University of Toronto, Ontario, Canada. J. Craig Venter Institute, Rockville, Maryland, USA.

Project Goals: The interactions between microorganisms in subsurface microbial communities are highly connected and complex, playing important roles in the cycling of carbon, nitrogen, and metals. As part of a larger research project using 'metaomics' approaches to study these processes (see also "The use of 'Metaomic' and Modeling Approaches at Subsurface Sites to Study Carbon, Nutrient and Metal Cycling in Microbial Communities") the objectives of this work are to: 1) use data from 'metaomic' approaches to reconstruct metabolic networks of microbial communities undergoing natural attenuation of uranium, 2) provide insights into the potential mechanisms that lead to numerical dominance of specific bacterial taxa under different environmental conditions, 3) study the interaction and collaboration among members of the microbial community and 4) integrate gene expression data characterizing metabolic activity into genome-scale models to describe microbial activity in sub-optimal states.

Reductive immobilization of soluble U(VI) to the insoluble U(IV) is an efficient remediation strategy of subsurface groundwater contaminated with uranium. While U(VI) can be reduced to U(IV) by stimulation of indigenous bacteria with amendments of organic carbon, there are zones of natural attenuation of U(VI) at the Rifle Integrated Field Research Challenge (RIFRC) site, where U(VI) can be both reduced and remobilized in the absence of biostimulation. Elucidation of the potential mechanisms of microbial community structure and metabolism under these conditions of natural attenuation can facilitate more efficient remediation design and management strategies.

In order to meet the goals of this ongoing project, our earlier efforts investigated the microbial community structure at a site previously identified as undergoing natural attenuation of uranium using metagenomic data. Due to a similarity in genome content between subsurface microbial community members assigned to the same taxonomic class at our study site, we pursued a pan-genome-scale approach to subsequently analyze metabolic potential at the class-level. A statistical analysis of the functional profiles from the JB site indicated that within the numerically dominant taxonomic classes (α- , β-, γ- and δ-proteobacteria, and Actinobacteria) there is an abundance of enzymes related to CO2 fixation (e.g. Rubisco in the α- β-, and γ- proteobacteria, and PEPCase in the Actinobacteria). In contrast, in acetate amended sites within the numerically dominant δ- proteobacteria, there is a high abundance of enzymes related to N2 fixation (e.g. Nitrogenase). Collectively, these results reveal different community structures and metabolic functions mediating C and N cycling under contrasting environmental conditions.

Using metagenomic and reference genome datasets, pan-genome-scale metabolic networks were reconstructed for α- , β-, γ- and δ-proteobacteria, and Actinobacteria, respectively. The models were optimized and gaps filled to ensure that they are capable of growth in geochemical conditions similar to that of the RIFRC site. These class-level models were then integrated into a Dynamic Multi-species Metabolic Modeling (DMMM) framework for investigating the interaction and collaboration among community members. The model analysis indicates that β-proteobacteria like Thiobacillus denitrificans may dominate the community at the JB site due to its ability to use inorganic electron donors for energy and fix CO2 as its major carbon source. Through electron transport with cytochrome bc1 complexes and NADH-Q oxidoreductase, a tight coupling between Fe(II) oxidation and NO3 reduction can be established to support use of CO2 as the main source of carbon. Similarly, reduced inorganic sulfur

compounds may be oxidized to sulfate by ferricytochrome c with reduction of NO3 as a terminal electron acceptor and fixation of CO2 as the major carbon source.

Interaction and collaboration of microbial community members were quantitatively estimated through the DMMM approach. While competitive interactions mainly occur in the community for electron donors and acceptors, and carbon sources, the simulations indicate that there are potential syntrophic interactions between β-proteobacteria (e.g. T. denitrificans) and Actinobacteria (e.g. Streptomyces). Actinobacteria may, for example, use the products of sulfur oxidation (e.g., sulfate) from β-proteobacteria as the final electron acceptor for CO2 fixation under anoxic conditions. Although these class-level models were able to capture potential syntrophic interactions between classes, further complexity exists at the species level, leading to important metabolic differences between class members. To facilitate species-level modeling efforts, we have currently focused on developing and validating models for both Anaeromyxobacter dehalogenans and T. denitrificans. Both models were developed by combining the Model SEED web resource with a novel algorithm for resolving bioenergetic inconsistencies. Furthermore, both models were curated to ensure consistency with available experimental data for growth on a variety of carbon sources and electron acceptors. We were then able to investigate the microbial community dynamics with species-level resolution using genome scale models of Geobacter sulfurreducens, G. metallireducens, A. dehalogenans, Rhodoferrax ferrireducens, and T. denitrificans, integrated using the DMMM approach.

Subsurface microorganisms may grow within either optimal or sub-optimal states depending on ever changing environmental conditions. Hence, application of a flux balance analysis (FBA) that seeks to maximize or minimize an objective function may not be always appropriate for describing and predicting microbial activities in the subsurface. Therefore, we are currently developing methods to integrate metatranscriptomic (gene expression) data into the genome-scale models, thereby elucidating functional mechanisms of the community indicative of, and relevant to, sub- optimal states. These combined models will be incorporated into the DMMM framework to improve the predictive capability of the genome-scale models.

This project is supported by the Office of Biological and Environmental Research in the DOE Office of Science DESC0006837.

69. The use of 'Metaomic' and Modeling Approaches at Subsurface Sites to Study Carbon, Nutrient and Metal Cycling in Microbial Communities

Kelvin Li1, Monika Bihan1, Amy Boaro2, Neha Gupta1, Christopher Gowen3, Jiao Zhao3, Kenneth H. Williams4, Mary Lipton2, Krishna Mahadevan3, Barbara A. Methé1* (bmethe@jcvi.org)

1J. Craig Venter Institute, Rockville MD, 2Pacific Northwest National Laboratory, Richland WA, 3University of Toronto, Canada, 4Lawrence Berkeley National Laboratory, Berkeley CA

Project Goals: Diverse microbial communities exist in subsurface environments that possess significant metabolic potential to effect global carbon, nitrogen and metal cycles including the transformation of radionuclides. Objectives of this ongoing project are: 1) to apply systems-level biology through application of 'metaomics' approaches (collective analyses of whole microbial community DNA, RNA and protein) to the study of microbial environmental processes and their relationship to C, N and metals including the influence of microbial communities on uranium contaminant mobility in subsurface settings undergoing natural attenuation, 2) improve methodologies for data generation using metaomics technologies and analysis and interpretation of that data and 3) use the data generated from these studies towards microbial community-scale metabolic modeling (see also "Genome-scale Metabolic Modeling of Microbial Communities at Sites Undergoing Natural and Stimulated Attenuation of Uranium").

To meet the goals of this ongoing project, two subsurface sites from the Department of Energy (DOE) Rifle Integrated Field Research Challenge Site (RIFRC) are being interrogated using a suite of metaomic approaches. The first site consists of sediments from the Winchester 2007 gallery, 'JB" well locations and was chosen due to the occurrence of natural attenuation of uranium (uranium reduction in the absence of biostimulation or other remedial interventions). The second and more recent sites of study within this project have been collected from Colorado River Floodplain (CORFP) sediments representing recent sediment depositions. Overbank deposits in the floodplain have become enriched in C, Fe and S minerals. Aggradation processes have led subsequent burial of these enriched sediments creating "hotspots" of biogeochemical activity which serve as analogs to the buried naturally reduced sediments at the JB sites.

Previous applications of metagenomic (DNA) and metatranscriptomic (RNA) inestigations from the 'JB' sediments have been generated using the Illumina HiSeq and MiSeq platforms Taxonomic profiles generated from both assemblies and high quality alignments (>60bp quality trimmed read length at >80% composite identity) to the NCBI NT database revealed that for both the metagenomic and metatranscriptomic data sets, the most abundant species based on best matches (~28% DNA, ~14% RNA) are to relatives of the facultative anaerobic chemolithotroph, Thiobacillus denitrificans, capable of coupling the oxidation of inorganic sulfur compounds to the reduction of oxidized nitrogen compounds and oxidization and potential remobilization of U(IV) to soluble U(VI) in the presence of nitrate. Evidence for the presence of metal reducing bacterial relatives (although not necessarily demonstrated to reduce uranium) such as Rhodoferax ferrireducens, were also determined.

The relatively high abundances of T. denitrificans and R. ferrireducens in the JB data served as motivation to develop a single combined genome-scale model based on a constraint-based metabolic reconstruction of the two organisms. An automatically generated model (through Model SEED) was available for T. denitrificans which required additional manual curation to facilitate its use in this study. Both individual models have been evaluated for their ability to produce flux through their respective biomass reactions after which the stoichiometry matrices underlying each organism's model were then combined. Among the findings from this model are that as the proportion of the biomass objective function shifts in dominance from 100% R. ferrireducens to 100% T. denitrificans, the flux through the combined biomass reaction increases towards 50%, suggesting a possible synergistic instead of

competitive relationship between the two organisms. Other outcomes include the finding of an increase of maximum flux towards T. denitrificans which may suggest a slightly more flexible set of capabilities for obtaining carbon and energy from a minimal soil media.

The CORFP samples represent an important opportunity to contrast microbial community diversity and function especially the coupling of carbon, nutrient and metal cycles with the results obtained from the JB sites using metatranscriptomic and metaproteomic approaches. To that end, biological and technical replicates of the metatranscriptome from both Illumina (short- read, ~100bp) and PacBio (long-read, ~3kb) sequencing platforms along with metaproteomic data from LC MS-MS approaches are currently being analyzed. Both data types support the finding that a significant proportion of the microbial community has been assigned to the beta- proteobacteria, including T. denitrificans, similar to the taxonomic profiles at the JB sites.

Investigations of important biological functions are underway using both data types which include support for the activity of proceses related to energy metabolism via glycolytic pathways and oxidative phosphorylation. For example, highly expressed genes from metatranscriptomic data were determined by using a statistical analysis in which the abundances of a collection of housekeeping genes was used to compute a null distribution to which the relative abundance of transcripts were compared to identify those genes whose expression was significantly greater (p- value <0.05) than the housekeeping genes. Significant expression was determnined for genes such as NADH dehydrogenase (oxidative phosphorylation) and glyceraldehyde phosphate (glycolysis). Using metaproteomic approaches; 620 proteins (FDR=8) were identified from a total of 3144 spectra. The most abundant proteins include ATP synthase subunits (n=58), and K(+)-insensitive proton pumps (n=11) (oxidative phosphorylation) as well as enolase (n=12) (glycolysis), and glyceraldehyde-3-phosphate dehydrogenase (n=11). The metatranscriptomic data also revealed significant expression of genes related to functions including heavy metal efflux pumps and other transporters, and PAS domain containing genes important to sensing environmental conditions. While metaproteomic data supported the presence of an active sulfur cycle including proteins from T. denitrificans, related to sulfur oxidation as well as proteins associated with sulfate reducers from the family Desulfobacteraceae (n=32). Currently, qRT- PCR approaches are also underway to validate metatranscriptomic results. Collectively, these investigations provide an opportunity to identify both the specific and potentially unique contributions of each metaomic data type (metranscriptome and metaproteome), including their individual sensitivities and specificities along with the nature and extent of the added value from data integration.

This project is supported by the Office of Biological and Environmental Research in the DOE Office of Science DESC0006837.

70. Soil Metabolomics to Decipher the Metabolic Foodwebs in Biological Soil Crusts

Tami Swenson1, Andrea Lubbe1* (alubbe@lbl.gov), Richard Baran1, Ulas Karaoz1, Rebecca Lau1, Suzanne Kosina1, Benjamin P. Bowen1, Ferran Garcia-Pichel1,2, Eoin L. Brodie1, Trent R. Northen1,3 (trnorthen@lbl.gov)

1Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720
2School of Life Sciences, Arizona State University, 427 E Tyler Mall, Tempe, AZ 85287
3DOE Joint Genome Institute, 2800 Mitchell Dr., Walnut Creek, CA 94598

https://portal.nersc.gov/webofmicrobes/

Project Goals: The goal of this project is to understand the small molecule composition of biological soil crust organic matter and the specificity of soil microbes towards specific soil metabolites.

Biological soil crusts (biocrusts) are communities of organisms inhabiting the upper layer of soil in arid environments. The crust itself is essentially microbial exopolysaccharide (EPS) linked sand particles and is critical to soil stabilization. Biocrusts persist in a dessicated dormant state for extended periods with rare pulsed activity events following precipitation. Microcoleus vaginatus, a non-diazotrophic filamentous cyanobacterium, is the key primary producer in bacterially dominated biocrusts in the Colorado Plateau and is an early pioneer in colonizing arid environments. Over decades, biocrusts proceed through developmental stages with increasing complexity of constituent microorganisms and macroscopic properties. Since Microcoleus vaginatus does not fix nitrogen, metabolic interactions with other biocrust microorganisms in the Microcoleus vaginatus-associated 'cyanosphere' presumably play a key role in the cycling of soil organic matter and in determining biocrust community dynamics.
Mass spectrometry-based soil extraction methods were developed in order to characterize the metabolite composition of biocrust. Chloroform fumigation of soil prior to extraction with water allowed the detection of the broadest range of intracellular and extracellular metabolites (amino acids, carboxylic acids, nucleotides, sugars, sugar alcohols, and fatty acids). To determine substrate preferences of key soil bacteria, exometabolomics analysis was performed using liquid chromatography coupled to mass spectrometry. Sixteen bacterial isolates were incubated in minimal media containing biocrust extracts and metabolite profiles were compared to uncultured fresh media to identify uptake and release of metabolites by specific microbes and these data are made available through our new exometabolomics data repository, the Web of Microbes (webofmicrobes.org).
Ultimately we anticipate that linking exometabolite cycling to specific microbes will prove to be invaluable datasets for both functional genomics and understanding soil carbon cycling.

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 Contract to Lawrence Berkeley National Lab No. DE-AC02-05CH11231.

71. Understanding microbial carbon cycling in soils using novel metabolomics approaches

Tami Swenson1, Richard Baran1, Ulas Karaoz1, Rebecca Lau1, Ferran Garcia-Pichel1,2, Eoin L. Brodie1, Trent R. Northen1,3* (trnorthen@lbl.gov)

1Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720
2School of Life Sciences, Arizona State University, 427 E Tyler Mall, Tempe, AZ 85287
3DOE Joint Genome Institute, 2800 Mitchell Dr., Walnut Creek, CA 94598

http://www.northenlab.org/research/biological-soil-crusts-biocrusts/

Project Goals: The Department of Energy has made major investments in soil sequencing efforts that have the potential to revolutionize predictive models of soil nutrient cycling. Yet we lack vital data to link sequence data to metabolic transformations in soils. This program aims to help bridge this gap by pioneering new soil metabolomics approaches that link microbial community structure to soil organic matter dynamics.

The organic matter that is cycled by soil microbes is a complex mixture of metabolites derived primarily from microbes, plant exudates and decomposing plant litter. It has a major impact on community structure since microbes may be both stimulated or inhibited by specific soil metabolite concentrations. Therefore, understanding soil metabolite composition and utilization by microbes would provide critical insights into how microbial communities are shaped and how, in turn, they shape soil metabolite composition.

Using tractable desert biological soil crust microbial communities (biocrusts) we take advantage of the cascades of microbial activities that follow wetting of dry soil to correlate soil metabolite composition and turnover to active microbes. These biocrusts are communities of organisms inhabiting the upper layer of soil in arid environments. They persist in a desiccated dormant state for extended periods of time and experience pulsed periods of activity facilitated by infrequent rainfall. Microcoleus vaginatus, a non-diazotrophic filamentous cyanobacterium, is an early pioneer in colonizing arid environments and key primary producer in early successional biocrusts. Metabolic interactions among biocrust microorganisms are poorly understood yet presumably play a key role in determining the community dynamics and cycling of carbon and nitrogen. This talk will describe our initial studies aimed at systematically deconstructing the metabolic foodwebs within these biocrusts. We have developed liquid chromatography tandem mass spectrometry soil metabolomic methods which we have used to characterize hundreds of metabolites from soils including many novel compounds. Exometabolite profiling of spent vs. fresh media is being used to link the uptake or release of these metabolites to specific bacterial isolates from biocrust. From this we find that Microcoleus vaginatus releases a broad range of metabolites. Many of these metabolites are found to be uptaken up by heterotrophs but there were surprisingly few metabolites uptaken by all soil bacteria. This points to a competition for a small set of central metabolites and specialization of individual heterotrophs towards a diverse pool of available organic nutrients. Overall, this initial study indicates that understanding the substrate specialization of biological soil crust bacteria can help link community structure to nutrient cycling.

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.

72. Agricultural land-usage completely reshapes nitrifier diversity

Anthony Bertagnolli and David Stahl

Civil and Environmental Engineering, University of Washington, Seattle WA 98195- 2700

Project Goals: This abstract represents ongoing work that is part of the DOE-BER funded project entitled "Archaeal Influence on Nitrogen Transformations in Soil and Root Systems of Bioenergy Crop Species". The most general objectives are to resolve the relative contributions of ammonia oxidizing archaea and ammonia oxidizing bacteria to nitrification and production of atmospherically active gases, evalulate biotic and abiotic variables that determine their distribution and activity, and identify control methods through amendment with selective inhibitors or land management practices within the context of bioenergy cultivars.

The transition of marginalized lands into agricultural usage significantly changes terrestrial systems, altering the biological, physical and chemical properties of soil ecosystems. However, it remains uncertain as to whether these changes have broader consequences on the underlying genetic diversity of microbial communities that facilitate specific biogeochemical cycles. Aerobic ammonia and nitrite oxidizing organisms are responsible for nitrification in terrestrial environments. This process has significant consequences on the distribution and form of nitrogen in soils, as well as the release of fugitive nitrogen oxide gases. Thus, controlling the environmental impacts of nitrogen fertilizer application is essential to achieving sustainable agriculture methods. To assess whether long-term management practices associated with cropping of potential bioenergy feedstocks (switchgrass cultivars) were associated with different nitrifying populations, amplicon sequencing (i.e., tag sequencing) was applied to biomarkers associated with aerobic ammonia and nitrite oxidation (ammonia monooxygenase-amoA, nitrite- oxidoreductase-nxrB, respectively) from long-term managed and native soils in eastern Washington, USA. These data indicated that aerobic ammonia-oxidizing archaea and bacteria (AOB and AOA, respectively) associated with soils under managed conditions were significantly different from native soils. At both broad and fine phylogenetic resolution (i.e., genus and sub-cluster), specific groups of AOA and AOB were identified that contributed either exclusively with managed or native lands, perhaps indicating different life strategies that maybe intimately related to the net-input of carbon and other nutrients associated with switchgrass cultivation. Nitrite oxidizing Bacteria (NOB) also displayed significant differences associated with management. However, nxrB genes observed in this study were highly diverse, displaying low nucleotide identity to previously observed sequences from environmental surveys, indicating tremendous unrecognized breadth of phylogenetic diversity within this group. Taken together, these results suggest that management practices associated with propagating soils of bioenergy cultivars select for specific taxonomic groups of aerobic nitrifiers.

73. Nitrous oxide emissions in fields of bioenergy crops are correlated with the form of nitrogen addition and ammonia-oxidizing microbial populations

Manmeet W. Pannu1, Kelley A. Meinhardt1, Anthony Bertagnolli1, Sally L. Brown2, Steve C. Fransen3, David A. Stahl1 and Stuart E. Strand1* (sstrand@uw.edu)

Civil and Environmental Engineering1 & College of Forestry2, University of Washington, Seattle, WA 98195, Department of Crop and Soil Science3, Washington State University at Prosser, WA 99350

Biofuels derived from biomass crops such as switchgrass (Panicum virgatum L.) are proposed as possible alternatives to fossil fuels. Nitrogen (N) fertilizer, required for high biomass yields, has known environmental impacts - including post-application emission of the greenhouse gas nitrous oxide (N2O) and nitrate leaching. We examined the impact of N nitrogen provided to switchgrass as inorganic fertilizer, or supplied via intercropping with nitrogen-fixing alfalfa, on N2O emissions, crop yields, and the abundance of ammonia oxidizing archaea (AOA) and ammonia oxidizing bacteria (AOB). Replicated field plots were established in 2010 at the Washington State University experiment station in Prosser, WA. Switchgrass (cv. Blackwell, an upland ecotype) N treatments included no N (control), 224 kg/yr of inorganic N (agronomic rate), and N derived from intercropped alfalfa (Medicago sativa L.). Plots were sampled during the 2012 and 2013 growing seasons. Nitrous oxide fluxes were measured at monthly intervals using static chambers from April through October in both 2012 and 2013. Soil samples were simultaneously collected and assayed for AOA and AOB abundance via quantification of the ammonia monooxygenase (amoA) gene. The N2O fluxes were elevated (maximum flux=12.5 g N2O-N/ha/d) immediately after fertilization and irrigation compared to the unfertilized control plots (average=0.75 g N2O-N/ha/d). Intercropping with alfalfa resulted in a significantly smaller N2O flux than the fertilized treatments (maximum flux ~4 g N2O-N /ha/d). Nitrous oxide emissions were higher immediately following irrigation compared to dry soils. The average 2- year biomass yield from were higher in fertilized plots (24.5±1.2 Mg/ha/yr) and intercropped plots (16.8 ±1.1 Mg/ha/yr) relative to the control plots (10.5±1.1 Mg/ha/yr). The AOA abundance was greater than the AOB abundance in soils from all treatments. AOA were more abundant in the intercropped treatments than the control and fertilized treatments. The AOB abundance was positively correlated with N2O emissions, but AOA abundance did not correlate with flux of this greenhouse gas. We conclude that intercropping switchgrass with alfalfa does enhance yields (relative to control plots) with concomitant reduction in N2O emissions (relative to plots receiving inorganic nitrogen) from AOB.

74. Advances in mass spectrometry imaging instrumentation and sample treatment protocols for in situ chemical characterization of biological samples

Troy J. Comi1 (comi2@illinois.edu), Sage J.B. Dunham1, Bin Li1, Nameera F. Baig2, Nydia Morales-Soto3, Eric J. Lanni1, Joshua D. Shrout3, Paul W. Bohn2, Jonathan V. Sweedler1

1Department of Chemistry and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801; 2Department of Chemistry and Biochemistry and Department of Chemical and Biomoleculer Engineering, University of Notre Dame, Notre Dame, IN 46556; 3Department of Civil and Environmental Engineering and Earth Sciences and Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556

http://www.scs.illinois.edu/sweedler

Project Goals: We are developing new instrumentation and complementary sample preparation methods which allow visualization of chemical composition and physical structure across multiple size scales. To this end, we have constructed a hybrid mass spectrometer capable of imaging in three distinct, complementary modes: matrix assisted laser desorption/ionization mass spectrometry imaging (MALDI MSI), secondary ion mass spectrometry (SIMS) imaging, and secondary electron microscopy (SEM). Here we explore mutually compatible sample treatments including sublimation of conventional organic MALDI matrix, and spray deposition of both matrices and metal nanoparticles, in order to enhance molecular coverage, ionization efficiency, and image contrast. We apply this new instrument and the developed treatments to visualize multi-scale chemical heterogeneity in bacterial biofilms to further understand the role of chemical signals in bacterial biofilm formation.

Microprobe mass spectrometry imaging (MSI) utilizes spatially-resolved ionization to interrogate a sample surface at regularly spaced positions. Each point generates a mass spectrum, containing information on the chemical species present at that location. Selected ion intensities at each position can be mapped to pixel color to display spatial information of the analyte abundance. Our lab has developed a mass spectrometer equipped with both a UV laser and a buckminsterfullerene (C60) ion source for MALDI- and SIMS-mode imaging, respectively.1 Interfacing these two ionization techniques with a commercial quadrupole/time-of-flight (QToF) mass spectrometer allows us to obtain in-situ molecular identification and the combined benefits of both MALDI and SIMS ionization. MALDI offers superior molecular coverage for high mass analytes such as lipids and peptides, whereas cluster SIMS offers superior spatial resolution, but tends to fragment molecules larger than ca. m/z 2000. We have further extended the imaging capabilities of our hybrid mass spectrometer by incorporating a secondary electron detector, which enables inline electron microscopy in conjunction with the MSI functionality. With a 5 μm spatial resolution, this microscope rapidly provides architectural and topographical information to overlay with chemical features from MSI.

Improving spatial resolution with MSI necessitates reducing microprobe size and therefore the quantity of sample available for analysis. Decreasing the limits of detection is imperative for extracting useful chemical information at higher spatial resolutions. One method for improving SIMS sensitivity is to apply a matrix or nanoparticles that more efficiently dissipates projectile energy and increases ionization yield. However, application of the matrix solution leads to analyte migration as well as extraction. The importance of these two effects needs to be balanced to achieve the best spatial resolution. We have observed increased signal intensity when the sample is coated with the MALDI matrix, 2,5-dihydroxybenzoic acid (DHB), or nanoparticles. The matrix application method also greatly affects analyte extraction and imaging spatial resolution, and we are currently adapting two matrix application methods, sublimation and robotic spray coating, for MALDI and SIMS analysis of bacterial biofilms.

We have optimized methods for DHB sublimation onto biofilms, for both SIMS and MALDI imaging. We first test the methods on rat spinal cord, a well-characterized bioanalytical standard for MSI, and then transition to biofilm analysis. Our results show that DHB sublimation is a powerful method for the analysis of lipid distributions on biofilm surfaces. Analyte delocalization is limited due to the dry coating and the small crystal size. A number of lipid-associated ions are detectable, with many potassium or sodium adducts. For MALDI, the optimal DHB coating is limited in the range of 0.1 to 0.3 mg/cm2, while for C60-SIMS the optimal coating is between 0.1 and 0.2 mg/cm2. Sublimation offers a promising avenue for matrix enhancements, however substituting nanoparticles would eliminate interference from DHB while still improving SIMS sensitivity. To uniformly deposit nanoparticles, we have developed a robotic matrix application system. Early trials with nanoparticle coatings have demonstrated speed, uniformity and high transfer efficiency. Two further modifications under development are electrospray and heated nebulizing gas.

We are investigating additional modifications to improve the biofilm-imaging capabilities. A concern with any analysis is that sample preparation will perturb the system, causing loss of information from the pristine sample. The current room-temperature vacuum analysis chamber requires complete desiccation of biofilms which may cause some analyte degradation or redistribution. A future goal is to begin SIMS imaging of hydrated biofilms by maintaining the vacuum chamber at cryogenic temperatures. MSI on our custom system presently utilizes stepper motors to move the sample in a regular array. Along with the SED upgrade, we can now control the ion beam location by electrical deflection. Instead of moving the sample, we can raster the ion beam to generate higher resolution images. We are developing methods to couple MS acquisition with the ion beam raster to achieve a lateral resolution of approximately 2 microns. Achieving sub-micron resolution, which is necessary for sampling sub-cellular components and cell-cell signaling molecule exchange, will require increasing our beam energy to 40 kV through a commercial source upgrade.

References

1. Lanni, E.J., Dunham, S.J.B., Nemes, P., Rubakhin, S.S. & Sweedler, J.V. Biomolecular imaging with a C60-SIMS/MALDI dual ion source hybrid mass spectrometer: Instrumentation, matrix enhancement, and single cell analysis. Journal of the American Society for Mass Spectrometry 25, 1897-1907 (2014).

This work was funded through DE SC-0006642: "In Situ Correlated Molecular Imaging of Chemically Communicating Microbial Communities"—DOE Office of Biological and Environmental Research

75. Multimodal chemical and architectural characterization of Pseudomonas aeruginosa microbial communities

Sage J.B. Dunham1 (dunham3@illinois.edu), Nameera F. Baig2, Nydia Morales-Soto3, Eric J. Lanni1, Joshua D. Shrout3, Paul W. Bohn2, Jonathan V. Sweedler1

1Department of Chemistry and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801;
2Department of Chemistry and Biochemistry and Department of Chemical and Biomoleculer Engineering, University of Notre Dame, Notre Dame, IN 46556;
3Department of Civil and Environmental Engineering and Earth Sciences and Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556

http://www.scs.illinois.edu/sweedler/

Project Goals: We are performing untargeted in situ chemical imaging of microbial communities. Here we combine complementary information collected from mass spectrometry imaging (MSI), confocal Raman microscopy (CRM), and electron microscopy (EM) to study the spatio-temporal dynamics underlying Pseudomonas aeruginosa bacterial biofilm formation and development. Each technique offers a unique advantage: MSI uses an ion's mass-to-charge (m/z) ratio to provide in-situ chemical identification and chemically specific ion imaging, CRM is a nondestructive light spectroscopy technique that differentiates compound classes by their inelastic scattering profile, and EM measures secondary electron emission profiles to provide an architectural map of the sample surface. Application of this combined approach highlights inter-film chemical heterogeneity and developmental differences both during biofilm development and resulting from different nutritional sources.

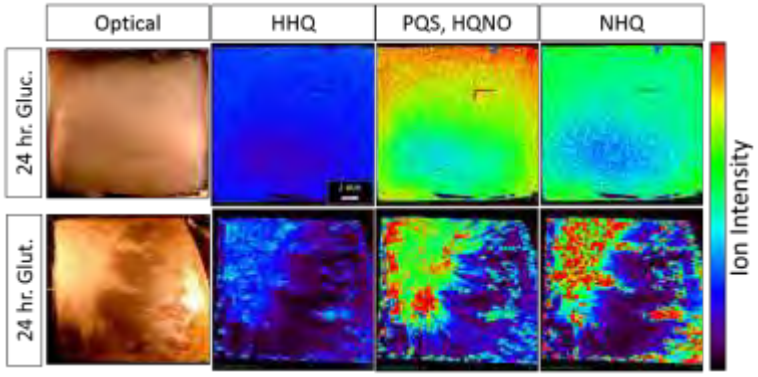
Many microorganisms, such as the gram-negative bacterium Pseudomonas aeruginosa, exist in highly cooperative communities that exhibit collective gene expression and coordinated behavior resembling that of a multicellular organism. The coordination of this ensemble is accomplished by way of a chemical communication system, termed quorum sensing (QS), in which the activation of genes □and therefore downstream protein synthesis and phenotypical differentiation □is controlled as a func population density; each cell secretes a baseline level of a chemical messenger and when the population density reaches a threshold level this messenger activates or deactivates the transcription of certain genes, leading to phenotypical expression and behavior exhibition. For example, the transition of a P. aeruginosa cell from the planktonic phenotype into that of a sessile biofilm, a process that results in expression changes in over 10 % of the genome, is largely dictated by the local population density and the concentration of specific QS chemical messengers.

Here we present the combination of two untargeted chemical imaging capabilities, namely SIMS imaging and CRM, in conjunction with the architectural characterization provided by EM, for in situ identification and chemical mapping of the QS molecules and their downstream products across the surface of P. aeruginosa biofilms. We primarily focus on the class of biomolecules known as quinolones, which are pivotal for gene control and biofilm development. The prominent quinolones studied here are the two QS molecules 2-heptyl-3- hydroxy-4(1H)-quinolone, or Pseudomonas quinolone signal (PQS), and its precursor 2-heptyl- 4(1H)-quinolone (HHQ), as well as a host of structurally similar molecules, a category that includes 2-nonyl-4(1H)-quinolone (NHQ), 2-Heptyl-4-hydroxyquinoline N-oxide (HQNO), 2- undecyl-4-hydroxyquinoline (C11:db-UHQ), C9:db-NHQ, C9-PQS, and C9:db-PQS.

In an early study, we developed a correlative methodology that utilizes an array of nanoparticle (NP) microdroplets as fiducial markers for accurate navigation and easy sample transfer between analytical

systems.1 For this analysis, a precisely spaced microdroplet grid is first applied to the dried biofilm surface with a chemical inkjet printer. The sample is then independently subjected to both CRM and cluster secondary ion mass spectrometry (SIMS), while recording the precise grid coordinates of each chemical signature. Finally, the regions of interest identified from the correlated CRM/SIMS work are analyzed using EM, which facilitates the assignment of chemical information to specific spatial constructs. Using this methodology we discovered highly localized chemical pockets containing HHQ and PQS, in addition to several other quinolones, specifically distributed over exposed cells on the film surface. The results of this experiment, as well as those of numerous follow-up studies, show that architecturally similar areas of the film can have vastly different chemical compositions, possibly suggesting phenotypical differentiation in purportedly clonal cellular consortiums.

Our studies have also revealed dramatic changes in chemical composition both during film development and under different sources of nutrition. We prepared separate cultures of P. aeruginosa on either glucose or glutamate as the nutritional source and monitored their development at 2 different stages: in the planktonic state (overnight culture), and after 24 hours of film growth. A combination of SIMS and CRM reveals that planktonic organisms grown on both glucose and glutamate produce low concentrations of several quinolones including HHQ and 2-nonyl-4(1H)-quinolone (NHQ). By 24 hours of film growth dramatic differences between the nutritional conditions begin to emerge. Highly localized quinolone regions are found to be protruding from the glutamate film, while the glucose film shows a much more diffuse distribution. These results highlight the power of our multimodal analytical approach for biofilm study and are of great interest to the microbiology community.



SIMS imaging of 24 hour *P. aeruginosa* films grown on glucose and glutamate. PQS and HQNO are isomers.

References

1. Lanni, E.J. et al. Correlated imaging with C60-SIMS and confocal Raman microscopy: Visualization of cell-scale molecular distributions in bacterial biofilms. Analytical Chemistry 86, 10885-10891 (2014).

This work was funded through DE SC-0006642: "In Situ Correlated Molecular Imaging of Chemically Communicating Microbial Communities"—DOE Office of Biological and Environmental Research.

76. Nutrient Stoichiometry Drives Carbon Turnover and Microbial Community Composition in Mineral and Organic Soils Under Rice Cultivation

Wyatt H. Hartman1, Rongzhong Ye2, William R. Horwath2, Susannah G. Tringe1* (sgtringe@lbl.gov)

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

Project Goals: We aim to characterize the microbial communities in wetland environments and evaluate their impact on carbon cycling, in order to develop strategies for optimizing carbon sequestration. Here we sought to determine the impact of soil carbon, nitrogen and phosphorus on carbon turnover and microbial community composition in an experimental rice field system.

Inundation of soils combined with emergent vegetation growth may enable high rates of soil carbon sequestration, and both rice cultivation and wetland restoration have been proposed as strategies to enhance carbon storage and reverse subsidence in degraded agricultural soils of the Sacramento-San Joaquin Delta. To isolate effects of wide variation in soils of the Delta on belowground C turnover, we studied a series of rice field trials as a model system with controlled vegetation and hydrology. Although soil C cycling may depend on the availability of carbon (C), nitrogen (N) and phosphorus (P), the role of P is particularly poorly understood due to a historical emphasis on soil C:N ratios informed by terrestrial nutrient limitation. To determine the effects of N and P availability on soil C turnover, we compared soil respiration over the course of a growing season in four adjacent rice fields with 5%, 10%, 20% and 25% soil C, each with control and N addition treatment plots (80 kg N/ha urea). Although soil P was not manipulated in parallel, prior work has shown soil P concentrations decline markedly with increasing soil C content. Soil CO2 and CH4 fluxes were monitored using static chambers at biweekly intervals during the growing season, and soils were collected at the end of the growing season for biogeochemical analysis and DNA extraction. Seasonal CO2 fluxes (per m2) were highest in 10% C soils (N:P =16:1), while soil N addition increased CO2 flux and soil C turnover (seasonal CO2 flux per unit soil C) in lower C fields (5% and 10% C), but not in higher C fields (20% and 25%). These patterns may be more clearly interpreted in light of shifts in soil N:P stoichiometry, which increased with soil C pools. Soil carbon turnover was greatest in mineral soils and inversely related to soil N:P ratios, suggesting progressive P limitation might limit both soil metabolism and its response to N at higher levels of soil C. Microbial community composition, based on 16S rRNA sequencing, was strongly influenced by soil C and pH along the gradient, but not by N additions as commonly observed in upland soils. Like soil carbon turnover, bacterial communities were also closely linked with soil N:P and inorganic P, and these relationships were significant even after accounting for covariance with soil C and pH. Functional gene content inferred from microbial phylogeny suggests substantial shifts in the potential to utilize carbon and phosphorus substrates between low and high C soils. Our results show that soil P availability and stoichiometry may affect microbial communities and their mediation of soil C turnover, even where primary producers appear limited by N.

The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02-05CH11231. This project was funded by the DOE Early Career Research Program, grant number KP/CH57/1 to Susannah Green Tringe, and was also supported by DOE JGI Community Sequencing Program.

77. The Complex Role of Restoration in Wetland Greenhouse Gas Production

Susanna Theroux1*(stheroux@lbl.gov), Wyatt Hartman1, Susannah Tringe1

1DOE Joint Genome Institute, Walnut Creek, California

Project Goals: This project aims to identify the geochemical and biological controls on greenhouse gas (GHG) production in coastal wetlands. Current projections of wetland response to global climate change are poorly constrained due to a limited understanding of the microbial mediators of carbon cycling in wetland soils. Using large-scale environmental genome sequencing, we aim to identify the microbial sentinels of GHG production and consumption in historic and restored wetlands of the San Francisco Bay-Delta. An enhanced comprehension of microbial metabolic arsenals will allow us to recommend management practices for future wetland restoration projects to maximize carbon sequestration and minimize microbial production of greenhouse gases.

Wetland environments play a critical role in the global carbon cycle, storing up to 35% of all terrestrial carbon and producing up to 75% of all non-anthropogenic methane. However, estimates of wetland GHG budgets are difficult to constrain given large variations in salinity, tidal regimes, soil saturation, organic carbon content, restoration status, and other geochemical and biological variables. Belowground microbial communities are the main drivers of greenhouse gases (GHG) cycling and their response to climate change will dictate whether a wetland serves as a net carbon sink or source. Wetland restoration has been proposed as a potential long-term carbon sequestration strategy, however wetland site selection and management practices are critical for ensuring restored wetlands sequester more GHG than they emit.

In an effort to better understand the underlying factors that shape the balance of carbon flux in wetland soils, we targeted the microbial communities along a salinity gradient ranging from freshwater tidal marshes to hypersaline ponds in the San Francisco Bay-Delta region. Seawater delivers sulfate to wetland ecosystems, providing a natural barrier to methane production, and traditional dogma predicts a decrease in methane production as salt content increases. Using 16S rRNA gene sequencing and shotgun metagenomics, coupled with greenhouse gas monitoring, we sampled sixteen sites capturing a range of salinities and restoration statuses. Soils were characterized for a suite of geochemical parameters including soil moisture, percent organic carbon, pH, and micronutrient, cation, and metal concentrations. The results of our microbial diversity survey showed that sampling location, plant type, and salinity were the primary drivers of both methane production and belowground microbial community composition. As expected, freshwater wetland soils produced more methane than brackish and saline sites, also reflected in the increased relative abundance of methanogen populations. Notably, restored freshwater and brackish wetlands produced orders of magnitude more methane than their historic counterparts. And conversely, unrestored hypersaline ponds produced methane at rates rivaling those of brackish restored wetlands. Both 16S and metagenome sequencing allowed for the identification of hallmark methanogen species whose relative abundance trended consistently with methane production. Our study links belowground microbial communities with their aboveground greenhouse gas production and provides a benchmark for predicting wetland soil microbial response to the inevitable influence of climate change.

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.This project was funded by the DOE Early Career Research Program, grant number KP/CH57/1 to Susannah Green Tringe, and was also supported by DOE JGI Community Sequencing Program.

78. The Thaw: Investigating Microbial Contribution to Permafrost Carbon Liberation through Functional Meta-omics

CM Singleton1* (caitlin_singleton@uq.edu.au) and BJ Woodcroft1, JA Boyd1, PM Crill2, CK McCalley3, SR Saleska3, VI Rich3, and GW Tyson1

1University of Queensland, Brisbane, Australia; 2Stockholm University, Stockholm, Sweden; 3University of Arizona, Tucson

http://www.ecogenomic.org/melting-permafrost

Project Goals: Microbial communities are key drivers of the carbon cycle, and through climate change induced thaw, permafrost associated carbon is increasingly exposed to microbial degradation. The consequent release of carbon dioxide and methane may exacerbate further thaw through a positive feedback mechanism. This project endeavors to determine the fate of cryosequestered and newly deposited carbon, by investigating a succession of microbial communities along a chronosequence of gradating permafrost thaw in Stordalen mire, subarctic Sweden. This study utilizes a range of meta-omic techniques including metagenome enabled community profiling and population genome recovery as well as metatranscriptomics to gauge in situ functional activity. Visual linkage of function to microorganisms using quantum dot labelling and confocal microscopy is also being explored. Importantly, this microbial ecology component is part of a much broader collaborative project featuring interdisciplinary expertise in metaproteomics, biogeochemistry and systems modelling.

Permafrost regions lock away around 50% of global soil carbon (1). However, the permafrost carbon sink is becoming a source for diverse microbial life that may fix carbon as biomass, or release it into the atmosphere as carbon dioxide or methane. Methane producing archaea (methanogens) reduce acetate (acetoclasts), or carbon dioxide and hydrogen (hydrogenotrophs), to produce energy for growth. The byproduct of this process is methane, which has 25 times the radiative forcing of carbon dioxide per molecule, making it a potent greenhouse gas (2). High methanogen abundances in thawing and thawed sites at Stordalen mire (3) contribute to increased methane generation (1), however aerobic methane-oxidizers (methanotrophs) also play an important role in controlling the flux of this methane. Observations of fluctuations in methanotroph diversity and metabolic efficiency following a thaw progression are currently lacking, yet they are central to understanding this rapidly changing environment. Thorough characterization of the microbial constituents of the thaw affected carbon cycle is facilitated by a comprehensive meta-omics approach.

An expansive data set (over 200 metagenomes and 24 metatranscriptomes spanning four depths, four months and five years) is revealing interactions between community members despite the enormous complexity of each environment in the thaw progression. GraftM (4), a new metagenome analysis tool, was developed specifically for such datasets, and efficiently and rapidly searches data for marker genes that can be taxonomically classified. GraftM identified and phylogenetically placed the methanogenesis gene mcrA, and methanotrophy genes pmoA and mmoX, in the metagenomes and transcriptomes. Initial analysis has determined distinct successional changes within communities as thaw progresses. Within the methanotrophs, there is a definite shift from Alphaproteobacteria to Gammaproteobacteria populations at both the genome and transcriptome level. Populations have also been identified outside of expected niches with aerobic methanotrophs appearing in deeper anoxic peat levels, alluding to the heterogeneity of the permafrost microenvironments. The recovery of over 100 population genomes is expanding insight into metabolic pathways, and when linked with metatranscriptomic information, can show up-regulation of some genes over others and may shed light on the propensity or tolerance of clades to particular environments.

Future efforts will be directed towards visualizing key players driving important metabolic processes such as methanogenesis, methane oxidation and cellulose degradation. Microbial data will be integrated with an extensive chronicle of geochemical data that encompasses carbon isotope, αC, dissolved organic matter, and methane concentration measurements. Ultimately, amalgamation of biotic and abiotic data will allow for the training of models and enable predictions on the impact of thaw to the global carbon system.

References:

1. McCalley, CK. et al. (2014). Methane dynamics regulated by microbial community response to permafrost thaw. Nature Communications 514: 478-481.
2. Schuur, EA. et al. (2008). Vulnerability of permafrost carbon to climate change: implications for the global carbon cycle. BioScience 58: 701-714.
3. Mondav, R. and Woodcroft, B. et al. (2014). Discovery of a novel methanogen prevalent in thawing permafrost. Nature Communications 5: 3212.
4. Boyd J. and Woodcroft B. (2015). GraftM. https://github.com/geronimp/graftM/

This work is supported by the U.S. Department of Energy under funding opportunity announcement number DE-FOA-0000866. Pathways to carbon liberation: a systems approach to understanding carbon transformations and losses from thawing permafrost.

79. Using GraftM to generate phylogenetically informed gene profiles in thawing permafrost metagenomes.

JA Boyd1* (j.boyd3@uq.edu.au), CM Singleton1, BJ Woodcroft1, and GW Tyson1

1 Australian Centre for Ecogenomics (ACE), University of Queensland, Australia; 2University of Arizona, Tucson http://ecogenomic.org/melting-permafrost

Project Goals: Microbes play a pivotal role in mediating the release and capture of carbon in the atmosphere and as a result are heavily involved in the regulation of the global carbon cycle. As Arctic and Antarctic frozen soil (permafrost) thaws in response to climate change, the extent of greenhouse gas (GHG) release from these large carbon sinks hinges on the microbial communities that inhabit them. This project focuses on establishing how microbes influence carbon release from these environments by identifying links between community dynamics, established by culture independent shallow metagenomics, and biogeochemical measurements. A model study site in Stordalen Mire, Sweden is used where a natural thaw gradient provides an ideal location for understanding how thaw progression impacts microbial communities. Here, a rapid and accurate community profiling tool (GraftM) is developed and applied to metagenomes extracted from samples taken at Stordalen Mire to recover the community composition of key GHG producing microbes.

Permafrost environments, once a collective carbon sink, are rapidly collapsing into GHG emitting wetlands as a result of climate change induced thaw. The potential for this GHG release to contribute to a positive feedback mechanism with global warming is worrying in light of the models that predict ~1700 Pg of carbon is cryosequestered within permafrost globally (1). The key producers of methane (methanogens), a GHG that is around 25 times more potent than carbon dioxide (2), are archaea that inhabit the expanding water saturated environments resulting from thaw. Previous work (3) has suggested that the abundance and population structure of methanogens is highly influential on the rate of methane release. However much of our current understanding of methanogen communities within these environments is limited to traditional microbiological approaches. These techniques, such as cultures or amplicon sequencing, have limitations such as culture bottlenecking and amplification bias which can negatively impact results. A view of methanogen community structure free from the limitations of traditional techniques would reveal how these populations respond to thaw, and clarify the implications of these responses for net CH4 release.

Here, culture independent shallow metagenomics is used to bypass traditional limitations in >200 environmental samples collected across the thaw gradient over multiple years and several depths at Stordalen Mire, Sweden. This massive dataset provides the unprecedented opportunity to explore and understand microbial communities in both spatial and temporal dimensions, however, rapidly recovering accurate community profiles from large metagenomic and metatranscriptomic data sets remains a non-trivial task. Moreover, the recent improvements in sequencing technology have pushed the yield of reads from sequencing into the billions, thus far outpacing the current bioinformatic tools available to analyze this data.

In this project, the methanogen population structure and abundance is elucidated by developing a novel metagenome/transcriptome analysis tool ("GraftM"), for use on the expansive metagenomic dataset at our disposal. GraftM uses Hidden Markov Models (HMMs), statistical models which use pattern recognition to rapidly identify fragments of key marker genes (such as 16S rRNA) within meta~omic data sets, and constructs community composition based on the placement of these fragments into phylogenetic trees. This approach sets GraftM apart from traditional BLAST based methods of identification and taxonomic assignment where the use of direct sequence-to-database comparisons are both slow and in the case of

complex environmental communities, potentially inaccurate. GraftM allows the exploration of microbial communities to remain unhindered by incomplete databases that often impede high resolution taxonomic assignment, and the size datasets being searched. These improvements make GraftM well suited to the task of analyzing the many complex metagenomes derived from Stordalen Mire in this project.

Here, the utility of GraftM as a profiling tool is demonstrated using in silico datasets where the community abundance was estimated consistently and accurately. Furthermore, GraftM is shown to perform more rapidly than the only published bioinformatic tool that utilizes tree insertion methods to infer phylogeny from marker gene fragments, PhyloSift. Using GraftM, the methanogen community composition was successfully recovered from metagenomes extracted from Stordalen Mire using three separate marker genes, two specific to methanogens (mer and mcrA) and one universal (16S rRNA). Highly significant correlations between the estimated abundance of methanogens as predicted by each of the three marker genes across metagenomes suggest that GraftM consistently assigns taxonomy using different marker gene sets. The population of methanogens was shown to diversify as thaw progresses, with a metabolically diverse community being found in the final stage of thaw hinting at an increase in potential methane production as permafrost thaws. Subsequent linking of key members within the methanogen community with biogeochemical data was found to be coherent with previously conducted amplicon studies on the site (3).

Future directions include the development of a marker gene database for use with GraftM, the integration of checks for copy number, and normalization of gene length. GraftM holds promise as a taxonomic gene profiler capable of rapid identification and classification of conserved marker and key functional genes in meta~omic datasets. This concept could further be expanded into a whole-pathway annotation tool for genome bins.

References

1. Tarnocai C. et al. Soil organic carbon pools in the northern circumpolar permafrost region (2009).
2. Torbjörn J. et al. Decadal vegetation changes in a northern peatland, greenhouse gas fluxes and net radiative forcing. Global Change Biology (2006).
3. Mondav R. and Woodcroft B. et al. Discovery of a novel methanogen prevalent in thawing permafrost. Nature communications (2014).

This work is supported by the U.S. Department of Energy under funding opportunity announcement *number DE-FOA-0000866. Pathways to carbon liberation: a systems approach to understanding carbon transformations and losses from thawing permafrost.*

80. Denitrifying and diazotrophic community responses to artificial warming in permafrost and tallgrass prairie soils

C. Ryan Penton1,2*(crpenton@asu.edu), James R. Cole1, Liyou Wu3, Yiqi Luo4, E.A.G. Schuur5, Jizhong Zhou3,6,7, James M. Tiedje1

1 Center for Microbial Ecology, Michigan State University, East Lansing, MI
2 College of Letters and Sciences, Arizona State University, Polytechnic Campus, Mesa, AZ
3 Institute for Environmental Genomics and Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK
4 Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK
5 Center for Ecosystem Science and Society, Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ
6 State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China
7 Earth Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

http://ieg.ou.edu/

Project Goals: The overall goal of this project is to advance system-level predictive understanding of the feedbacks of belowground microbial communities to multiple climate change factors and their impacts on soil C cycling processes. Towards this goal, we are pursuing the following objectives: (i) To determine the responses of microbial community structure, functions and activities to climate warming, altered precipitation, soil moisture regime and/or clipping in the tundra and temperate grassland ecosystems; (ii) To determine the temperature sensitivity and substrate priming on recalcitrant C decomposition; (iii) To determine microbiological basis underlying temperature sensitivity of recalcitrant C decomposition; and (iv) To develop integrated bioinformatics and modeling approaches to scale information across different organizational levels towards predictive understanding of ecosystem responses to multiple climate change factors, which will be collaborated and integrated with the KBase.

Increasing air temperatures has been shown to impact soil biogeochemical processes, although the corresponding changes to the underlying microbial functional communities are not well understood. Alterations in the nitrogen (N) cycling functional component are particularly important as N availability can affect microbial decomposition rates of soil organic matter and influence plant productivity. To assess these changes, the composition of the N-fixing (nifH), and denitrifying (nirS, nirK, nosZ) soil microbial communities was assessed by targeted pyrosequencing of functional genes involved in N cycling in Oklahoma (OK) tallgrass prairie and Alaskan (AK) permafrost soils at sites where the experimental effect of climate warming is under investigation. Raw reads were processed for quality, translated with frameshift correction, and a total of 313,842 amino acid sequences were clustered and linked to a nearest neighbor using reference datasets. The number of OTUs recovered ranged from 231 (NifH) to 862 (NirK). The N functional microbial communities of the prairie, which had experienced a decade of experimental warming were the most affected with changes in the richness and/or overall structure of NifH, NirS, NirK and NosZ. In contrast, the AK permafrost communities, which had experienced only one year of warming, showed decreased richness and a structural change only with the nirK-harboring bacterial community. A highly divergent nirK- harboring bacterial community was identified in the permafrost soils, suggesting much novelty, while other N functional communities exhibited similar relatedness to the reference databases, regardless of site. Prairie and permafrost soils also harbored highly divergent communities due mostly to differing major populations. Lastly, in order to capture a higher percentage of functional gene harboring microbial communities, we have implemented a primer design tool. New nosZ primer sets have been designed, capturing >93% of diversity among 1,600 reference

sequences.

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.

81. Expanding the Toolkit for Metagenomics, Implementing in KBase, and Applying It to the Study of the Effects of Experimental Warming in Midwestern and Alaskan Soils

James R. Cole1* (colej@msu.edu), Qiong Wang1, Jordan Fish1, Mariah Gilman1, Konstantinos T. Konstantinidis2, Luis M. Rodriguez-R2, Liyou Wu3, Zhili He3, Yiqi Luo3, Edward A.G. Schuur4, James M. Tiedje1, Jizhong Zhou3

1Michigan State University, East Lansing; 2Georgia Institute of Technology, Atlanta; 3University of Oklahoma, Norman; 4 Northern Arizona University, Flagstaff

http://ieg.ou.edu/

Project goal: The overall goal of this project is to advance system-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 main objectives of this integrative project are to (i) determine the responses of microbial community structure, functions and activities to climate warming, altered precipitation and soil moisture regime in a tundra and temperate grassland ecosystem; (ii) determine temperature sensitivity on recalcitrant C decomposition; (iii) determine the microbiological basis that is underlying temperature sensitivity of recalcitrant C decomposition; and (iv) develop integrated bioinformatics and modeling approaches to scale information across different organizational levels. This work focuses on project utilization and integration with DOE KBase resources.

As part of this project we have developed a suite of tools to help us understand the impact of multiple climate change factors on soil carbon cycling processes. We are in the process of porting these tools to the KBase environment to allow us to better share data and methods between our geographically-separated team, thereby providing rapid access to our data and tools, through KBbase, to the broader community of researchers involved in understanding microbial community processes relevant to climate change variables.

In May 2014, two of our developers traveled to Argonne National Labs for an intensive two-day work session to gain hands-on experience with the new KBase deployment procedures and the new central KBase infrastructure components: Shock, a data management system for storing and sharing large data, and AWE, a workflow management system for job scheduling. While at Argonne, we were able to integrate a synchronous RDP Classifier (Wang et al., 2007) developmental version into the new KBase infrastructure. This service accepts input files that were uploaded to the Shock server, executes the RDP Classifier command on the server host and returns results back to users. We have also developed a novel tool called Nonpareil to estimate the sequencing coverage of a metagenomic dataset, i.e., what fraction of the total DNA/species diversity of a microbial community was sampled by a metagenomic dataset, and predict the sequencing effort required to achieve nearly complete community coverage based on the redundancy of shotgun metagenomic reads (Rodriguez-R and Konstantinidis, 2014). This and other tools were developed to overcome several challenges we faced when analyzing soil metagenomes (i.e., the tools were developed to fulfill practical needs). More recently, we have worked with KBase personnel to further increase the computational efficiency of Nonpareil for a large metagenomic database (>100 Gb per dataset) using a kmer approach pioneered by the KBase team, and make Nonpareil available to the scientific community as part of KBase. The release of the tool through KBase is scheduled for later in 2015.

We are proceeding with tool development with “ready for KBase” as a top design priority. We are continuing to implement and improve our new and existing big data analysis tools to be faster and more memory efficient (Cole et al., 2014). We continue to make the source code of our tools available on

GitHub (a KBase deployment requirement), and make them easy to access and install. We continue to produce the detailed documentation and workflow examples required by KBase and that allow the tools to be used more efficiently by the user community. In addition to the tools we previously targeted for KBase integration, we have developed a novel method for assembling specific protein-coding genes of interest. Metagenomics can provide important insights into microbial communities. However, assembling even modest metagenomic datasets with traditional methods has proven to be very computationally challenging. This method uses a combined graph structure and assembles more, longer and better quality gene contigs. This method is implemented as an open source software package called Xander (Wang et al., 2015) and is available at https://github.com/rdpstaff/RDPTools.

References

1. Cole, J. R., Q. Wang, J. A. Fish, B. Chai, D. M. McGarrell, Y. Sun, C. T. Brown, A. Porras- Alfaro, C. R. Kuske, and J. M. Tiedje. 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 42(1):D633-642; doi: 10.1093/nar/gkt1244 [PMID: 24288368]
2. Rodriguez-R, L. M., K. T. Konstantinidis. 2014. Nonpareil: a redundancy-based approach to assess the level of coverage in metagenomic datasets. Bioinformatics 30(5):629-635.
3. Wang, Q, G. M. Garrity, J. M. Tiedje, and J. R. Cole. 2007. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. Appl Environ Microbiol.
4. 73(16):5261-5267; doi: 10.1128/AEM.00062-07 [PMID: 17586664]
5. Wang, Q., J. A. Fish, M. Gilman, Y. Sun, C. T. Brown, J. M. Tiedje and J. R. Cole. Xander: gene-targeted metagenomic assembler. Submitted.

Funding Statement: This research was supported by the Office of Science (BER), U.S. Department of Energy, Biological Systems Research on the Role of Microbial Communities in Carbon Cycling Program (DE- SC0010715) with additional contribution from the DOE Great Lakes Bioenergy Research Center (BER DE-FC02-07ER64494).

82. Functional and taxonomic characterization of Alaskan permafrost and Oklahoma prairie soils using comparative metagenomics: implications for responses to climate change

Konstantinos T. Konstantinidis1* (kostas@ce.gatech.edu), Eric R. Johnston1, Chengwei Luo1, Luis M. Rodriguez-R1, Liyou Wu2, Shi Zhou2, Kai Xue2, Zhili He2, Mengting Yuan2, Yiqi Luo2, Edward A.G. Schuur3, James R. Cole4, James M. Tiedje4, Jizhong Zhou2

1Georgia Institute of Technology, Atlanta, GA 30332, USA; 2University of Oklahoma, Norman, OK 73019, USA; 3Northern Arizona University, Flagstaff, AZ 86011, USA; 4Michigan State University, East Lansing, MI 48824, USA.

http:/www.ieg.ou.edu/

Project Goals: The overall goal of this project is to advance system-level predictive understanding of the feedbacks of belowground microbial communities to multiple climate change factors and their impacts on soil C cycling processes. Regarding this goal, we are pursuing the following objectives: (i) To determine the responses of microbial community structure, functions and activities to climate warming, altered precipitation, soil moisture regime and/or clipping in the tundra and temperate grassland ecosystems; (ii) To determine the temperature sensitivity and substrate priming of recalcitrant C decomposition; (iii) To determine the microbiological basis underlying temperature sensitivity of recalcitrant C decomposition; and (iv) To develop integrated bioinformatics and modeling approaches to scale information across different organizational levels towards predictive understanding of ecosystem responses to multiple climate change factors, which will be collaborated and integrated with the KBase.

Under this project, we have investigated microbial communities from Alaskan tundra permafrost (AK) and Oklahoma temperate grassland (OK) soils, both of which have been experimentally warmed 2 to 4°C for five years above ambient temperature in-situ. Our analyses of well-replicated 16S rRNA gene amplicon, meta-transcriptomic, and whole- community shotgun metagenomic datasets from these soils after one year of warming (5- year data are forthcoming) showed small but significant shifts in community composition, gene expression, and functional metabolic potential compared to the control (un-warmed, adjacent communities). The specific microbial populations and gene/pathways enriched by warming were different between AK and OK. Greater taxonomic composition differences were observed at the OK site relative to AK, presumably resulting from longer generation times due to the less optimal conditions for growth at permafrost soils. Analysis of rRNA gene amplicons recovered in the shotgun-metagenomic data revealed no significant shifts in fungal taxa at both sites, but that the ratio of fungi to bacteria decreased with warming, indicating that the warming treatment was more favorable for bacteria, at least in the short term. The most pronounced bacterial taxon shifts observed at OK site, which was somewhat also observed at the AK site, were an increase in abundance of Actinobacteria and a decrease in Planctomycetes, both representing major phyla in soils, particularly involved in C cycling. In terms of functions, the communities of AK warmed plots were enriched in metabolic pathways related to labile carbon mobilization and oxidation whereas fewer of these patterns were observed in the OK communities, indicating that soil C was more vulnerable to microbial respiration at AK. The OK communities were instead enriched in genes involved in heat shock response and cellular surface structures, particularly, trans-membrane transporters for glucosides and ferrous iron. These results, which were consistent with independent physicochemical measurements and process rates determined in-situ, were linked with higher primary productivity of the aboveground plant communities stimulated by warming. By implementing a new contig binning strategy, we recovered large genomic fragments (>500Kbp continuous; 2-8Mbp non-continuous) representing several abundant populations (0.2-2% of the total community) in the AK metagenomes. These populations appeared to be highly conserved across spatial and environmental

gradients at the AK site and apparently played a key role in microbial community response to the warming treatment. Collectively, our findings suggest that microbial communities of grassland soils play important roles in mediating feedback responses of the soil ecosystem to climate change and that even short periods of warming induce significant changes in microbial community function and composition

To enable this research, we have also developed several bioinformatics tools that addressed practical limitations during the comparative analysis of the soil metagenomes such as how to assess the fraction of the community captured by a metagenomic dataset and estimate the sequencing effort required in study design (Nonpareil tool; Rodriguez-R and Konstantinidis, Bioinformatics 2013), how to determine the taxonomic affiliation of a metagenomic sequence (MyTaxa; Luo et al., NAR 2014), how to bin assembled contigs into population genomes based on time-series metagenomes (BinGeR; Luo et al., in preparation), and how to determine differentially present genes between metagenomic datasets (Luo et al., Methods Enzymol. 2013). Altogether, these make up our Microbial Process Toolkit for gene, metagenomic and metatranscriptomic data integration, modeling and visualization. We are in the process of implementing our toolkit in KBase and we will report on these efforts as well.

References

1. C. Luo, L. M. Rodriguez-R, and K. T. Konstantinidis. A User's Guide to Quantitative and Comparative Analysis of Metagenomic Datasets. Methods Enzymol. 531:525-47. 2013.
2. C. Luo, L. M. Rodriguez-R, E. R. Johnston, L. Wu, L. Cheng, K. Xue, Q. Tu, Y. Deng, Z. He,< J. Zhou Shi, M. M. Yuan, R. A. Sherry, D. Li, Y. Luo, E. A. G. Schuur, P. Chain,< J. M. Tiedje, J. Zhou, and K. T. Konstantinidis. Soil Microbial Community Responses to a Decade of Warming as Revealed by Comparative Metagenomics. Applied & Environ. Micriobiol. In press.

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.

83. Metagenomics-Enabled Predictive Understanding of Microbial Communities to Climate Warming: Results from Long Term Soil Incubations

Rosvel Bracho1, Edward A.G. Schuur1*(Ted.Schuur@nau.edu), Susan Natali2, Elaine Pegoraro1, Liyou Wu3, Huaqun Yin3, Kai Xue3, Mengting Yuan3, Jin Zhang3, Ye Deng3, Zhili He3, Jim Cole4, James M. Tiedje4, Konstantinos Konstantinidis5, Junyi Liang3, Katherine Todd-Brown3, Yiqi Luo3 and Jizhong Zhou3

1Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ and Department of Biology, University of Florida, Gainesville, FL; 2Woods Hole Research Center, Falmouth, MA; 3Institute for Environmental Genomics and Department of Botany and Microbiology, University of Oklahoma, Norman, OK; 4Center for Microbial Ecology, Michigan State University, East Lansing, MI; 5Center for Bioinformatics and Computational Genomics and School of Biology, Georgia Institute of Technology, Atlanta, GA.

http://ieg.ou.edu/

Project goal: The overall goal of this project is to advance system-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 main objectives of this integrative project are to (i) determine the responses of microbial community structure, functions and activities to climate warming, altered precipitation and soil moisture regime a in a tundra and temperate grassland ecosystem; (ii) determine temperature sensitivity on recalcitrant C decomposition; (iii) determine microbiological basis that is underlying temperature sensitivity of recalcitrant C decomposition; and (iv) develop integrated bioinformatics and modeling approaches to scale information across different organizational levels.

As part of this project, this study focused on using laboratory incubations of soils as an isolated system to understand the influence of microbial processes on the release of C, and their response to changes in temperature. Long-term aerobic laboratory incubations were performed at two temperatures (15 & 25°C) to determine the temperature sensitivity of microbial respiration (Q10) in soils from a field warmed tundra (Alaska, AK). Three different layers were incubated from the AK site: two surface layers (0 – 15 cm & 15-25 cm with high carbon content) and a layer deeper in the profile that thaws only a few months each year (45-55 cm). Carbon fluxes were measured periodically over the course of the incubation and total C respired at each time (CT) was fitted with a two-pool C model to estimate pool sizes and decay rates of fast (CF) and slow (CS) decomposing C fractions. We estimated Q10 using two different methods: (1) a short-term method (Q10-S), where soils were exposed to six different temperatures ranging from 5 to 30°C over the period of <1 week, while measuring C fluxes over this range at 14, 100 and 280 days during the long-term incubation; (2) the equal carbon method (Q10-EC), which estimates the ratio of time that it takes for a soil to respire a given amount of carbon at each incubation temperature through the incubation experiment. Q10-EC was estimated for fast and slow decomposing C pools. The microbial community phylogenetic and functional composition, structure, and dynamics were analyzed by 16S rDNA sequencing and a functional gene array (GeoChip 5.0).

Results from the AK soils showed that after one year of incubation, CT in the top 15 cm could be as high as 25% and 15% of the initial soil C content at 25°C and 15°C incubations respectively. The fast decomposing C pool (CF) accounted for up to 5% of the initial C content in the top 15 cm soils. Both, CT and CF decreased with depth but no field warming effect was detected in part because these soils were collected from the initial period of warming for the field experiment. Overall average turnover time for CF was ~ 60 days at these laboratory conditions.

Turnover time for CS varied from 10 years in top soils to ~60 years in soils near the permafrost at 15°C incubation and decreased by half at the higher incubation temperature. The overall short term Q10-S for AK soils was 2.55 ± 0.03. Neither treatment nor depth nor day of incubation, nor incubation temperature had any effect on Q10-s; however, interactions of field treatment*depth and treatment*day of incubation were significant. Estimated Q10-EC was 1.2 ± 0.4 and 2.2 ± 0.06 for CF and CS, respectively, with no significant differences with field treatment, varying depth or incubation temperature. The AIC analysis indicated that the two-pool model was the best fit given the incubation data over one year.

The results of the dissimilarity analysis of the 16S rRNA amplicon sequencing data showed a significant warming effect in the soils from 0-15 cm when incubated at 25oC for two weeks. In those samples, Proteobacteria, Actinobacteria, Acidobacteria, Chloroflexi, Firmicutes, and Gemmatimonadetes had higher abundances in the warmed soils, while Planctomycetes,
Chlamydiae, Armatimonadetes, Verrucomicrobia, Bacteroidetes had lower abundances. Significant differences in microbial population abundances between warmed and control soils from 0-15 cm were also observed when incubated at 15oC after three months, and the soils from >35 cm after 9 months. From GeoChip analyses, differences were observed among soil depths, between treatments, incubation temperatures, and incubation time. For the incubations at 25oC, consistent differences in carbon degradation genes between warmed and control soils from 15-25 cm were observed along the incubation time points.

Short-term incubations (14 days) were performed in order to understand the potential influence of added C on the turnover of native organic matter. Alaska soils from surface (15-25 cm) and deep (45-55 cm) layers were amended with U-13C-labeled glucose (66.45‰) solution, at a concentration of 3 mg glucose-C g-1 dry soil. Two analytical replicates for each depth were incubated aerobically at 15°C inside 1L Mason jars. Flux was measured with an Automated Soil Incubation System (ASIS) that utilizes an open path infrared gas analyzer (LI-820) and a CR1000 data logger. A Picarro Cavity Ringdown Spectrometer (G2101) was used to measure 13CO2. The addition of glucose resulted in higher respiration rates in amended soils; however, priming, increased decomposition of native organic matter, was only observed in the deep layer, where 30% more soil-derived C was respired compared to control (unamended) soils. Net priming was not observed in the surface layer, suggesting that microbes were not energy starved. This preliminary data suggests that microbes in deep layers are limited in energy, and the addition of fresh organic C increases native organic matter decomposition in soils with greater fractions of slowly decomposing C. Based on the long-term incubation study that found that the quickly decomposing C pool in permafrost represents <10% of the total soil C, priming in permafrost ecosystems could enhance mineralization rates of slowly decomposing C accumulated over the long-term.

This work is supported by the US Department of Energy, Biological Systems Research on the Role of Microbial Communities in Carbon Cycling Program (DE-SC00010715).

84. Methods for estimating temperature sensitivity of soil organic matter based on incubation data: A comparative evaluation

Junyi Liang1 (jliang@ou.edu), Katherine Todd-Brown1 (ktoddbrown@gmail.com), Dejun Li1, Zheng Shi1, Huaqun Yin1, LiYou Wu1, Zhili He1, James M. Tiedje2, Jizhong Zhou1, Edward G. Schuur3, Konstantinos T. Konstantinidis4, Yiqi Luo1

1University of Oklahoma, Norman, OK 73019, USA; 2Michigan State University, East Lansing, MI 48824, USA; 3Northern Arizona University, Flagstaff 86011, USA; 4Georgia Institute of Technology, Atlanta, GA 30332, USA.

http://ieg.ou.edu

Project Goals: The overall goal of this project is to advance system-level predictive understanding of the feedbacks of belowground microbial communities to multiple climate change factors and their impacts on soil C cycling processes. Towards this goal, we are pursuing the following objectives: (i) To determine the responses of microbial community structure, functions and activities to climate warming, altered precipitation, soil moisture regime and/or clipping in the tundra and temperate grassland ecosystems; (ii) To determine the temperature sensitivity and substrate priming on recalcitrant C decomposition; (iii) To determine microbiological basis underlying temperature sensitivity of recalcitrant C decomposition; and (iv) To develop integrated bioinformatics and modeling approaches to scale information across different organizational levels towards predictive understanding of ecosystem responses to multiple climate change factors, which will be collaborated and integrated with the K-Base.

Abstract: To better understand the temperature sensitivity (Q10) of soil organic matter (SOM) decomposition, we compared several commonly used methods (i.e., one-pool (1P) model, two- discrete-pool (2P) model, three-discrete-pool (3P) model, and time-for-substrate (T4S) Q10 method) plus a new and more process-oriented approach for estimating Q10 of SOM decomposition from laboratory incubation data to evaluate the influences of the different methods and assumptions on Q10 estimation (Liang et al., 2015). The process-oriented approach is a three-transfer-pool (3PX) model that resembles the decomposition sub-model commonly used in Earth system models. The temperature sensitivity and other parameters in the models were estimated from the cumulative CO2 emission using the Bayesian Markov Chain Monte Carlo (MCMC) technique. The estimated Q10s generally increased with the soil recalcitrance, but decreased with the incubation temperature increase. Our results indicated that the 1P model did not adequately simulate the dynamics of SOM decomposition and thus was not adequate for the Q10 estimation. All the multi-pool models fitted the soil incubation data well. The Akaike information criterion (AIC) analysis suggested that the 2P model is the most parsimonious. As the incubation progressed, Q10 estimated by the 3PX model was smaller than those by the 2P and 3P models because the continuous C transfers from the slow and passive pools to the active pool were included in the 3PX model. Although the T4S method could estimate the Q10 of labile carbon appropriately, our analyses showed that it overestimated that of recalcitrant SOM. The similar structure of 3PX model with the decomposition sub-model of Earth system models provides a possible approach, via the data assimilation techniques, to incorporate results from numerous incubation experiments into Earth system models.

However, it is difficult to constrain the transfer parameters between carbon pools based on the soil respiration data alone (Liang et al., 2015). We propose using the change in functional gene quantified over the course of an incubation to further constrain these models. We took soils from a field warming and deep collar experiment in Oklahoma and incubated the cores over 9 months. In addition to traditional bulk soil organic carbon measurements and soil respiration readings, we quantified the DNA associated with 14,000+ genes involved in carbon degradation and metabolism using GeoChip at 2 weeks, 3 months,

and 9 months during the incubation. Genes were selected as 'fast' genes if they did not increase during the incubation and dropped below 50% of their initial expression levels. Fast genes were then averaged and used with CO2 respiration readings to constrain the parameters of the 3PX model. The results showed that each field treatment has a unique microbial community associated with the fast pool. The fast genes were not enriched with labile carbon metabolic genes and instead were roughly proportional to the gene types sampled. Fast pool estimates provided more information on several parameters for the model without resulting in different carbon predictions over 100 years.

References:

1. Liang J, Li D.,Shi Z., Tiedje J.M., Zhou J., Schuur E.A.G., Konstantinidis K.T., Luo Y. (2015) Methods for estimating temperature sensitivity of soil organic matter based on incubation data: A comparative evaluation. Soil Biology & Biochemistry 80: 127-135.

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.

85. Successional Dynamics of Grassland Microbial Communities in Response to Warming, Precipitation Alternation, and Clipping

Mengting Yuan1, Zhou Shi1, Xue Guo1, Xishu Zhou1, Liyou Wu1*(lwu@rccc.ou.edu), Zhili He1, Joy D. Van Nostrand1, Lauren Hale1, Qichao Tu1, Jie Deng1, Jianjun Wang1, Zhenxin Li1, Tong Yuan1, Jim Cole2, James M. Tiedje2, Edward A.G. Schuur3, Konstantinos Konstantinidis4, Yiqi Luo1, and Jizhong Zhou1

1Institute for Environmental Genomics and Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK; 2Center for Microbial Ecology, Michigan State University, East Lansing, MI; 3Center for Ecosystem Sciences and Society, Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ; 4Center for Bioinformatics and Computational Genomics and School of Biology, Georgia Institute of Technology, Atlanta, GA

http://ieg.ou.edu/

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 the underlying microbiological basis; and (iv) develop integrated bioinformatics and modeling approaches to scale information across different organizational levels. As a part of the integrated project, this study focused on field experiments in a temperate grassland to reveal the influences of elevated temperature, altered precipitation and clipping on soil microbial communities and ecosystem processes. This study consists of three major parts: (i) development and evaluation of next generation functional gene arrays for discerning the impacts of climate change factors on microbial community structure; (ii) long-term succession of plant and microbial communities through sequencing analysis of 16S rRNA gene amplicons and plant communities; and (iii) short-term microbial succession through analyses of 16S rRNA amplicons and functional genes. Overall, this study provides valuable new insights into our understanding of the temporal dynamics of soil microbial communities in response to multiple climate change factors.

Development and evaluation of next generation functional gene arrays. To analyze the functional diversity, composition, structure, dynamics, and metabolic potential/activity of soil microbial communities, we have developed a new generation of functional gene array, GeoChip 5.0 with 167,044 probes targeting 395,894 coding sequences from 1,593 functional gene families involved in C, nitrogen (N), sulfur, and phosphorus cycling, electron transfer, metal homeostasis, organic remediation, stress response, secondary metabolism, and virus and virulence activity. As reported previously, GeoChip 5.0 is highly specific (>98% probes have 100-fold higher signal than the corresponding mismatch probes), sensitive (250ng DNA for detection of >50% targets), and quantitative (99.1% probes have good linearity between signal intensity and DNA concentration; $R2>0.81$, $P<0.05$). Also, the reproducibility of GeoChip hybridization (180K probes per array) was compared to 16S rRNA gene amplicon sequencing (very deep, >80,000 reads per sample) and shotgun metagenomic sequencing (average data size of 18 Gb per sample) using the same soil DNA, showing GeoChip had significantly higher reproducibility (>90% by both Sørensen and Bray-Curtis similarities) than 16S rRNA sequencing (62% by Sørensen, 80% by Bray-Curtis), which was in turn higher than shotgun metagenomic sequencing (5% by Sørensen, 55% by Bray-Curtis). These results provided direct experimental evidences for the presence of artifacts associated with random sampling processes, and indicate that closed-format technologies (e.g., microarrays) do lead to higher reproducibility than open-format technologies (e.g., sequencing). In addition, the sources of variation that could affect reproducibility in GeoChip-detected functional gene communities were

estimated using warmed and control soils. DNA extraction and hybridization each accounted for 1% of the community total variance. Differences among technical replicates were insignificant. Compared to the treatment effects and the biological blocking factor (13% and 27% total variance explained, respectively), variations caused by experimental processes for functional gene array hybridization are negligible.

Long-term succession of plant and microbial communities. The long-term experimental site was established on a temperate grassland of Central Oklahoma (OK) in 2008. Environmental factors simulated include air warming (two levels: control and warming), precipitation alteration (three levels: double, normal, and half) and clipping (two levels: unclipped and clipped). Surface layer (0-15cm) soil samples from all plots were collected annually at the summer peak plant biomass season (Sep. or Oct.) from 2009 to 2014. At the same time, plant biomass and community structure were surveyed. Air and soil temperature, soil moisture and the C flux rate were measured in field. Soil pH, and C and N content were determined in lab for each soil sample taken. Since the beginning of operation, soil temperature at 7.5cm depth was 3°C higher ($P<0.001$) in warmed compared to control plots and 0.8°C ($P<0.001$) lower in unclipped plots compared to clipped plots, due to the insulating effect of plant and litter cover. Soil moisture was significantly ($P<0.001$) lower in clipped and in warmed plots, and higher in the double precipitation plots compared with their controls. Plant biomass fluctuated greatly across all years largely due to differences in wet and dry years, with an observable ($P<0.05$) decrease with reduced precipitation. Pointedly, C3 plant biomass was reduced with warming and elevated precipitation, while C4 plants thrived with more precipitation as indicated by their greater biomass. To examine microbial community changes under treatments, the 264 annual soil samples taken from 2009 to 2014 were analyzed by sequencing 16S rRNA gene amplicons. Dissimilarity analysis showed that all treatments alone as well as year affected the microbial community structure. Permutational analysis of variance indicated that over time, warming became the most influential treatment. α-diversities of microbial communities under different treatments were similar in 2009 and 2010, but gradually decreased in warming ($P<0.05$) samples in later years. The community composition was also altered over time with significantly ($P<0.05$) more Actinobacteria and Firmicutes, and less Proteobacteria, Bacteroidetes and Acidobacteria in warmed samples in 2014, compared to 2009 and 2010 when the communities for control and warming plots were similar. Differences in precipitation level and clipping also impacted the community composition. By 2014, the half precipitation soil samples had less abundant proteobacteria and more abundant firmicutes than the double precipitation soils ($P<0.05$), while in clipped plots, Actinobacteria were less abundant. These results indicate that warming, altered precipitation and clipping have differential effects on the diversity, composition and structure of soil microbial communities over time. GeoChip analysis of these samples is in progress. Short-term microbial succession. To understand the short-term dynamics of microbial communities in response to warming, 96 monthly soil samples were collected in 2012 from in both warmed and control plots and analyzed by sequencing 16S rRNA gene amplicons and GeoChip hybridization. Both warming and sampling month significantly ($P<0.005$) affected the soil microbial community structure for both taxonomic groups and functional genes, while the interaction between these variables was not significant. Also, in general, warming increased the relative abundance of Actinobacteria and Crenarchaeota, and N assimilation and denitrification genes, and decreased the abundance of δ-Proteobacteria, Planctomycetes, OD1 and Chlamydiae, and phosphorus and sulfur metabolism genes. C cycling genes were differentially influenced depending on the pathways involved. In addition, a significantly ($P=0.02$) higher β-diversity (Sørensen index) of warmed communities suggested that they were taxonomically more divergent than the control communities. Finally, more taxa/functional genes showed seasonal dynamics compared to the warming effect. Specifically, Euryarchaeota, Actinobacteria, Bacteroidetes, and Alphaproteobacteria reached their highest abundances concurrently with the highest plant biomass.

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.

86. RDP: Data and Tools for Studying Structure and Function of Microbial Communities

Benli Chai* (chaibenl@msu.edu), Qiong Wang, Yanni Sun, C. Titus Brown, James M. Tiedje and James R. Cole

Michigan State University, East Lansing

http://rdp.cme.msu.edu http://fungene.cme.msu.edu

Project Goals: RDP offers aligned and annotated rRNA and important ecofunctional gene sequences with related analysis services to the research community. These services help researchers with the discovery and characterization of microbes important to bioenergy production, biogeochemical cycles, climate change, greenhouse gas production, and environmental bioremediation.

Over the past year, the RDP websites (Cole et al., 2014) were visited, on average, by 10,908 researchers (unique IP) in 21,842 analysis sessions each month. In addition to the tools and sequence data offered on our websites most RDP tools are available through our GitHub repository (https://github.com/rdpstaff/), with step-by-step instructions for each tool. These command-line RDP tools offer additional options not available on the web and are suitable for high-throughput analysis.

We enhanced our FunGene site (Fish et al., 2013) with new important functional genes. These included additional nitrogen cycle genes such as "clade II" nitrite reductase (atypical NosZ), important for understanding both nitrogen utilization in biofuel crops and greenhouse gas production. Use of the FunGene web increased to an average of 2117 researchers per month in 2878 analysis sessions. During the past year, we updated FunGene data releases five times from searches of the primary sequence databases.

In May 2014, two of our developers traveled to Argonne National Labs for an intensive two-day work session to gain hands-on experience with the new KBase deployment procedures and the new central KBase infrastructure components: Shock, a data management system for storing and sharing large data, and AWE, a workflow management system for job scheduling. While at Argonne, we were able to integrate a synchronous RDP Classifier developmental version into the new KBase infrastructure. This service accepts input files that were uploaded to the Shock server, executes the RDP Classifier command on the server host and returns results back to user.

The RDP Classifier has been extended with new features and training sets for new genes. The Classifier has been enhanced to use gene copy number data to adjust relative taxonomic assignment percentages based on copy number. The precompiled Classifier is trained with the 16S gene copy number data from rrnDB (http://rrndb.umms.med.umich.edu). Both the bacteria and archaea 16S rRNA and fungal LSU training sets have been updated in 2014 to incorporate new phyla and increase coverage of basal lineages. In addition to the previous two fungal LSU training sets, we now provide the Classifier trained on two new high-quality fungal ITS taxonomies: Warcup, a hand-curated set with mycological taxonomic names; and UNITE, grouped by "species hypothesis" accession codes. Both ITS training sets achieved high classification accuracy (85% and 88%) at species level with Warcup performing better at assigning novel sequences when the taxon is not present in the training set.

To demonstrate the utility of command-line RDP tools, we developed a detailed tutorial to walk researchers through the steps of data analysis using RDP tools and importing results to third- party tools. We included the R/Bioconductor package Phyloseq in our tutorial for its extensive suite of analysis and visualization functions. The tutorial provides both supervised and unsupervised methods with example

workflows and sample scripts.

RDP's mission includes user support; email rdpstaff@msu.edu or call +1(517) 432-4998.

References

1. Cole, J. R., Q. Wang, J. A. Fish, B. Chai, D. M. McGarrell, Y. Sun, C. T. Brown, A. Porras- Alfaro, C. R. Kuske, and J. M. Tiedje. 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis Nucl. Acids Res. 42(Database issue):D633-D642; doi: 10.1093/nar/gkt1244 [PMID: 24288368]
2. Fish, J. A., B. Chai, Q. Wang, Y. Sun, C. T. Brown, J. M. Tiedje, and J. R. Cole. 2013. FunGene: the Functional Gene Pipeline and Repository. Front. Microbiol. 4:291; doi: 10.3389/fmicb.2013.00291 [PMID: 24101916]

Funding Statement: RDP research was supported by the Office of Science (BER), U.S. Department of Energy, Grant No. DE-FG02-99ER62848; additional funding came from the DOE Biological Systems Research on the Role of Microbial Communities in Carbon Cycling Program (DE- SC0010715) and from the DOE Great Lakes Bioenergy Research Center via Office of Science (BER) U.S. Department of Energy (DE-FC02-07ER64494).

87. Conditions For Duality Between Fluxes and Concentrations in Biochemical Networks

Ronan M.T. Fleming1(ronan.mt.fleming@gmail.com), Nikos Vlassis2, Ines Thiele1, Michael A. Saunders3

1Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 7 avenue des Hauts-Fourneaux, Esch- sur-Alzette, Luxembourg. 2Adobe Research, 345 Park ave, San Jose, CA, USA. 3Dept of Management Science and Engineering, Stanford University, Stanford, CA, USA.

http://www.en.uni.lu/lcsb/research/systems_biochemistry

Project Goals: In experimental systems biology, the majority of high throughput experimental data is of molecular abundance and the minority is of reaction rates. We seek a modeling framework flexible enough to integrate experimental data on both rates and abundance. With explicit rep- resentation of molecular abundance it becomes possible to mechanistically model regulation, e.g., genetic regulation, where the abundance of an active transcription factor modulates the rate of transcription by binding to a sequence motif in competition with other genomic structural proteins. Phenomenological kinetic models (e.g., Michaelis-Menten kinetics) are potentially more biochemi- cally realistic than flux balance models as they simultaneously represent concentration and flux, but their solution for a steady state concentration quickly becomes intractable for networks with over ˜100 reactions. Our goal is to develop the first algorithms for tractable modeling of steady state flux and concentration at genome-scale.

Mathematical and computational modeling of biochemical networks is often done in terms of either the concen- trations of molecular species or the fluxes of biochemical reactions. When is mathematical modeling from either perspective equivalent to the other? Mathematical duality translates concepts, theorems or mathematical structures into other concepts, theorems or structures, in a one-to-one manner. We present a novel stoichiometric condition that is necessary and sufficient for duality between unidirectional fluxes and concentrations. Our numerical ex- periments, with computational models derived from a range of genome-scale biochemical networks, suggest that this flux-concentration duality is a pervasive property of biochemical networks. We also provide a combinatorial characterisation that is sufficient to ensure flux-concentration duality. That is, for every two disjoint sets of molec- ular species, there is at least one reaction complex that involves species from only one of the two sets. When unidirectional fluxes and molecular species concentrations are dual vectors, this implies that the behaviour of the corresponding biochemical network can either be described entirely in terms of concentrations or unidirectional fluxes.

This work was supported by the U.S. Department of Energy, Offices of Advanced Scientific Computing Research and the Biological and Environmental Research as part of the Scientific Discovery Through Advanced Computing program, grant #DE-SC0010429.

88. Consistent Estimation of Gibbs Energy Using Component Contributions

Elad Noor1, Hulda S. Haraldsd´ottir2* (hulda.haraldsdottir@uni.lu), Ron Milo1 and Ronan M. T. Fleming3

1Plant Sciences Department, Weizmann Institute of Science, Rehovot, Israel; 2Center for Systems Biology, University of Iceland, Reykjavik, Iceland; 3Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

http://www.en.uni.lu/lcsb/research/systems_biochemistry

Project Goals: The aim of this project was to develop of a method for consistent estimation of reaction Gibbs energies in metabolic networks. Reaction Gibbs energies can be used in metabolic modeling for applying thermodynamic constraints on reaction rates, metabolite concentrations and kinetic parameters.

The increasing scope and diversity of metabolic models has led scientists to look for genome-scale solutions that can estimate the standard Gibbs energy of all the reactions in metabolism. Group contribution methods greatly increase coverage, albeit at the price of decreased precision. We present here a way to combine the estimations of group contribution with the more accurate reactant contributions by decomposing each reaction into two parts and applying one of the methods on each of them. This method gives priority to the reactant contributions over group contributions while guaranteeing that all estimations will be consistent, i.e. will not violate the first law of thermodynamics. We show that there is a significant increase in the accuracy of our estimations compared to standard group contribution. Specifically, our cross-validation results show an 80% reduction in the median absolute residual for reactions that can be derived by reactant contributions only. We provide the full framework and source code for deriving estimates of standard reaction Gibbs energy, as well as confidence intervals, and believe this will facilitate the wide use of thermodynamic data for a better understanding of metabolism.

This work was supported by the U.S. Department of Energy, Offices of Advanced Scientific Computing Research and the Biological and Environmental Research as part of the Scientific Discovery Through Advanced Computing program, grant #DE-SC0010429.

89. Constraint-based reconstruction and analysis toolbox extensions for parallel network reconstruction and CellDesigner interoperability

Longfei Mao 1*(longfeimao99@gmail.com), Ronan M.T. Fleming 1

1 Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 7 avenue des Hauts-Fourneaux, Esch-sur-Alzette, Luxembourg

http://wwwen.uni.lu/lcsb/research/systems_biochemistry

Project Goals: This project aims at developing software packages for easy metabolic network reconstruction, annotation and visualization and extending the functionality of the established constraint-based modeling software for analysis and bioengineering of cellular metabolisms at systems level.

In the past decade, constraint-based modelling, as implemented in the COBRA toolbox [1], has emerged as an important tool to study the mechanisms behind biological phenotypes and understand physiological and perturbed metabolic states at genome-scale for organisms. However, application of this novel approach to systems biology for different use scenarios is still hindered by the functionality of the current version of the COBRA toolbox. For instances, constraint-based modelling desires a high-quality reconstruction that can demand extensive manual efforts of contributors with different background knowledge, but the existing reconstruction software does not allow synergizing people's efforts in real time towards a reconstruction. In addition, CellDesigner (CD) is a diagram editor commonly utilized to render the network layout and annotate metabolite and reactions, but there are no COBRA functions available to parse and edit the network files created by CD.

To extend the scope of the established COBRA methods for reconstruction, visualization and annotation of a metabolic network, we have developed two COBRA extensions. The first extension is a collaborative metabolic reconstruction platform, rBioNetServer, stemming from a previously published network reconstruction package rBioNet [2]. Different from the ordinary rBioNet, rBioNetServer adopts a client/server model in which the client program communicates with a remote database server, where stores and manages different versions of reconstructions. Based on this architecture, rBioNetServer enables different people to work on the same metabolic reconstruction simultaneously, to inert reaction, metabolites and relevant annotations, and perform quality control procedures according to the established protocol [3]. The second extension is an interface package that can parse an XML file created by CD, change the thickness of reaction links and colours of nodes in the XML file according to the flux values obtained from constraint- based modelling. There are also auxiliary functions not only to compare a network model stored in the CD XML file and a COBRA model in the Matlab file and correct the discrepancies between them, but also integrate different types of omics data into the XML file in line with the Miriam standard. In the future, we envisage that these extensions can facilitate a community-driven effort to ensure a high quality reconstruction of the cellular metabolism for an organism and serve as a bridge between constraint-based modelling and the popular process diagram editor, CellDesigner, for visualisation and annotation of metabolic networks.

References

1. Schellenberger, J., et al., Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nature protocols, 2011. 6(9): p. 1290-307.
2. Thorleifsson, S.G. and I. Thiele, rBioNet: A COBRA toolbox extension for reconstructing high-quality biochemical networks. Bioinformatics, 2011. 27(14): p. 2009-10.
3. Thiele, I. and B.O. Palsson, A protocol for generating a high-quality genome-scale metabolic reconstruction. Nature protocols, 2010. 5(1): p. 93-121.

This work was supported by the U.S. Department of Energy, Offices of Advanced Scientific Computing Research and the Biological and Environmental Research as part of the Scientific Discovery Through Advanced Computing program, grant #DE-SC0010429

90. Globally Convergent Algorithms to Solve Systems of Non-linear Biochemical Reaction Equations

Ronan M. T. Fleming∗ (ronan.mt.fleming@gmail.com) and Francisco J. Aragon

Artacho Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg.

http://wwwen.uni.lu/lcsb/research/systems biochemistry

Project Goals: In experimental systems biology, the majority of high throughput experimental data is of molecular abundance and the minority is of reaction rates. We seek a modeling framework flexible enough to integrate experimental data on both rates and abundance. With explicit rep- resentation of molecular abundance it becomes possible to mechanistically model regulation, e.g., genetic regulation, where the abundance of an active transcription factor modulates the rate of transcription by binding to a sequence motif in competition with other genomic structural proteins. Phenomenological kinetic models (e.g., Michaelis-Menten kinetics) are potentially more biochemi- cally realistic than flux balance models as they simultaneously represent concentration and flux, but their solution for a steady state concentration quickly becomes intractable for networks with over ˜100 reactions. Our goal is to develop the first algorithms for tractable modeling of steady state flux and concentration at genome-scale.

Solving systems of non-linear biochemical reaction equations is an essential part of genome-scale kinetic modeling. We introduce a new class of function, called duplomonotone, which is strictly broader than the class of monotone function. We study some of the main properties of duplomonotone functions and provide an example of a nonlinear duplomonotone function used to model a system of biochemical reactions. We present three variations of a derivative- free line search algorithm for finding zeros of systems of duplomonotone equations, and we prove their linear convergence to a zero of the function [1]. This classification of the key function appearing in deterministic models of systems of biochemical reactions appears to hold more generally, opening up the realistic prospect of globally convergent algorithms for genome-scale kinetic modeling for the first time.

References

1. Aragon Artacho, F. J., and Fleming, R. M. T. Globally convergent algorithms for finding zeros of duplomonotone mappings. Optimization Letters, Sept. 2014.

This work was supported by the U.S. Department of Energy, Offices of Advanced Scientific Computing Research and the Biological and Environmental Research as part of the Scientific Discovery Through Advanced Computing program, grant #DE-SC0010429.

91. Identification of Conserved Metabolic Moieties by Graph Theoretical Analysis of Atom Transition Networks

Hulda S. Haraldsd´ottir∗ (hulda.haraldsdottir@uni.lu) and Ronan M. T. Fleming

Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

http://wwwen.uni.lu/lcsb/research/systems biochemistry

Project Goals: We aim to develop scalable algorithms for estimation of reaction kinetic parameters in metabolic networks from nonstationary isotope labeling experiments [1]. Computational modeling is essential for the design and analysis of isotope labeling experiments. Current modeling frame- works are limited to small metabolic networks as they require algorithms that scale poorly with network size. We propose to simplify these models by reducing the number of atoms that need to be considered for labeling. To acheive this we will take advantage of the fact that not all atoms traverse metabolic networks independently of each other. Groups of atoms, known as conserved moieties, remain intact through all reactions of a network [2]. Atoms belonging to the same moiety do not have to be modeled separately but can be represented as a single model variable.

All conserved moieties in a metabolic network can be represented as nonnegative integer vectors in the left null space of the network's stoichiometric matrix. Algorithms exist for computing such vectors based on reaction stoichiometry [3, 4, 5] but none are globally convergent. Moreover, existing algorithms give no information about moiety structure and may return vectors that do not correspond to any moieties. Here, we present a novel algorithm for computing conserved moieties in metabolic networks by graph theoretical analysis of the corresponding atom transition networks. Our algorithm returns the exact group of atoms belonging to each moiety as well as the corresponding vector in the left null space of the input stoichiometric matrix. Computation time appears to scale linearly with network size. Computation of conserved moieties in a large metabolic network with more than 4,000 internal reactions completed in approximately two hours.

References

1. Wahl SA, Noh K, Wiechert W (2008) 13c labeling experiments at metabolic nonstationary conditions: An exploratory study. BMC Bioinformatics 9: 152.
2. Reich JG, Sel'kov EE (1981) Energy Metabolism of the Cell. A Theoretical Treatise. London, U.K.: Academic Press. 00006.
3. Schuster S, Hilgetag C (1995) What information about the conserved-moiety structure of chemical reaction systems can be derived from their stoichiometry? The Journal of Physical Chemistry 99: 8017–8023.
4. Famili I, Palsson BO (2003) The convex basis of the left null space of the stoichiometric matrix leads to the definition of metabolically meaningful pools. Biophysical Journal 85: 16–26.
5. De Martino A, De Martino D, Mulet R, Pagnani A (2014) Identifying all moiety conservation laws in genome- scale metabolic networks. PLoS ONE 9: e100750.

This work was supported by the U.S. Department of Energy, Offices of Advanced Scientific Computing Research and the Biological and Environmental Research as part of the Scientific Discovery Through Advanced Computing program, grant #DE-SC0010429.

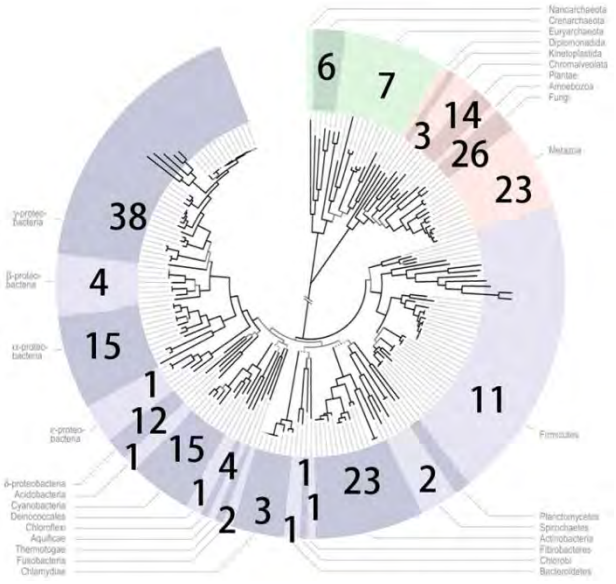
92. The PNNL Biodiversity Library

Samuel H Payne* (samuel.payne@pnnl.gov), Matthew E Monroe, Christopher Overall, Gary R Kiebel, Grant Fujimoto, Bryson Gibbons, Michael Degan, Mary S Lipton, Richard D Smith

Biological Sciences Division, Pacific Northwest National Laboratory, Richland WA

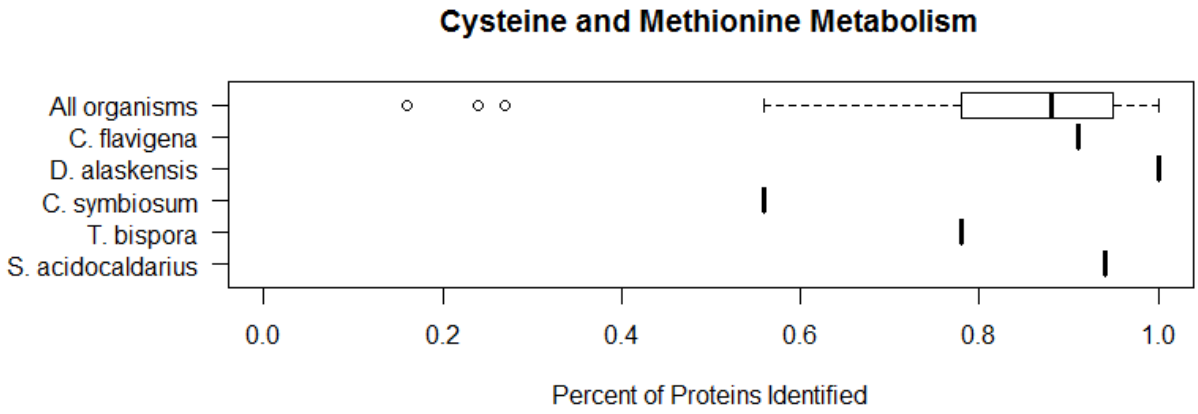
Project Goals: This project is focused on improving algorithms and methods for mass spectrometry data analysis of metaproteomics data. Recent advances in mass spectrometry and biological separations have dramatically increased the depth of proteomic discovery. Unfortunately, traditional computational workflows are in many cases preventing researchers from realizing these benefits for microbial communities. We propose to create a new generation of computational workflows to overcome the sensitivity limitations inherent in status quo data processing schemes.

Unlike laboratory experiments that focus on one single organism, biofuel production is often explored in experiments with either natural or synthetic communities composed of numerous organisms. To understand the active state of these communities it is advantageous to assay the proteome as opposed to the genome. Unfortunately current algorithms for peptide/spectrum matching are reliant on protein databases. The practical implication of this dependence for proteomics of communities (metaproteomics) is that the algorithms often fail to identify sufficient number of peptides and proteins. The major goal of this project is to improve peptide and protein identification in community proteomics datasets. Metaproteomics currently is difficult for environmental samples that either have no protein sequence information (for a project without matched metagenomics) or for projects where the strain heterogeneity hinders complete characterization with metagenomics sequencing. Therefore, we are developing algorithms that work in the absence of sequence databases. Our current approach is to adapt the library search to match spectra of similar, but non-identical, peptides. This requires a very large spectrum library of previously identified peptide/spectrum matches.

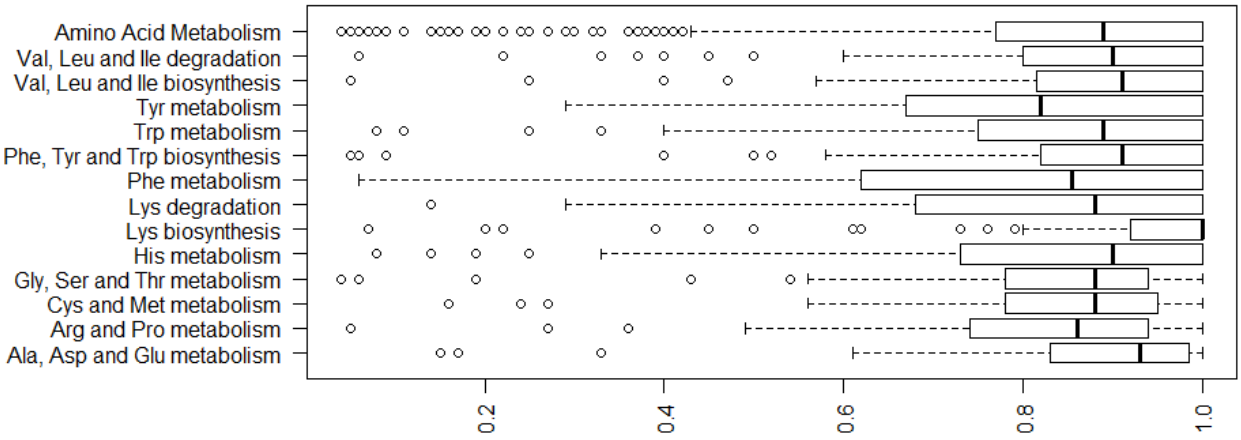


We present here the PNNL Biodiversity Library, a collection of the annotated peptides and proteins from 112 bacterial and archaeal organisms representing 15 phyla into public 3rd party repositories. In addition to the annotation of millions of peptide sequences, we also deposited 13 TB of raw mass spectrometry data. The library contains >2.5 million confidently identified peptides (230,000 proteins) from >60 million spectra. By mapping these proteins onto known pathways via KEGG, we show that proteomics of functionally annotated proteins is consistently high, even for organisms with relatively little data.

As an example of the depth of coverage, we show the coverage of proteins in the cysteine and methionine metabolic pathway. On average, proteomics data for a single organism covers 90% of all annotated proteins in this pathway (see plot below). Specific species are also shown for reference: Desulfovibrio alaskensis, Cellulomonas flavigena, Cenarchaeum symbiosum, Thermobispora bispora, and Sulfolobus acidocaldarius.



This high coverage is typical for amino acid metabolic pathways, and any pathway annotated in KEGG. Thus the Library is both deep in its coverage of annotatedprotein function and broad in its coverage of taxonomic diversity.



93. Inducible Extreme Expression of Enzymes in Poplar

Charleson Poovaiah* (crpoovai@syr.edu), Yao Xiao and Heather Coleman

Syracuse University, Syracuse, New York

http://biology.syr.edu/

Project Goals: The overall goals of the project is to verify in poplar the In Plant Activation (INPACT) technology1, which enables inducible expression of genes and accumulation of proteins at very high levels in planta and to evaluate the ability of cellulases produced through this technology to hydrolyze cellulose to simple sugars for fermentation.

One of the major limitations for the cost-effective production of biofuels is the cost of cellulolytic enzymes. One potential solution is the in planta production of these enzymes within the biofuels crop itself. Utilizing In Plant Activation technology (INPACT), which uses the rolling circle replication of the Gemini virus to produce high levels of gene amplification and protein production under an inducible promoter, this project aims to produce high levels of cellulases in poplar upon induction. Initially, we will verify this technology in poplar for its ability to accumulate recombinant proteins at very high levels. We will then use the technology to express cellulases from three major groups, endoglucanases, exoglucanases and β- glucosidases, in poplar with constitutive and tissue specific promoters. Cellulases from thermophilic organisms have been plant codon optimized and synthesized. These cellulases are being assessed for correct splicing in yeast. The construct harboring the alcohol inducible promoter driving the production of the replication initiation protein (Rep) which causes induction of the INPACT system has been successfully transformed into poplar and mother plants selected based on Rep/RepA gene expression and plant growth before and after alcohol treatment. Mother plants with the highest Rep gene expression after induction and with no expression prior to induction, are currently being transformed with the split gene cassette to assess the INPACT system in poplar. The transgenic poplar plants with the complete INPACT system are being regenerated and will be evaluated for expression in leaf and developing xylem using the GUS reporter system.

References

1. Dugdale, Benjamin, et al. "In Plant Activation: An inducible, hyperexpression platform for recombinant protein production in plants." The Plant Cell Online 25.7 (2013): 2429-2443.

Funding statement: This project is funded by the U.S. Department of Energy.

94. A Systems Biology, Whole-Genome Association Analysis of The Molecular Regulation of Biomass Growth and Composition in Populus deltoides

Annette Fahrenkrog,1 Leandro Neves,1 Juan Acosta,2 Brad Barbazuk,1,3,4 and Matias Kirst1,2,3* (mkirst@ufl.edu)

1Plant Molecular and Cellular Biology Program, University of Florida, Gainesville; 2School of Forest Resources and Conservation, University of Florida, Gainesville; 3Genetics Institute, University of Florida, Gainesville; 4Biology Department, University of Florida, Gainesville.

Project Goals: The overall goal of this project is to identify genetic polymorphisms that regulate biomass productivity and composition in poplar trees. The aims proposed to identify the genetic variants of interest are: 1) optimization of sequence-capture for unbiased, high-throughput and low-cost recovery of target coding and regulatory sequences in P. deltoides; 2) "whole-genome" genotyping of a P. deltoides unstructured population for association mapping; and 3) identification of significant SNP-trait associations with biomass growth and carbon partitioning to define genes and alleles that regulate trait variation. The genes identified with this approach will be used for the implementation of efficient molecular breeding and germplasm selection strategies, accelerating poplar breeding.

Poplars are fast growing trees with high potential to become a major feedstock for the biofuel industry. For these trees to be widely adopted for biofuel production, genotypes with specialized wood quality and growth traits are required. A better understanding of the genetic regulation of biomass growth and wood composition is needed to accelerate the development of new poplar clones through genomics-assisted breeding. Despite their importance, little is known about the genes that regulate these economically important traits. We are using a genome-wide association genetics approach to discover these genes in eastern cottonwood (P. deltoides), a poplar species with a wide distribution in North America and high levels of genetic and phenotypic variation.

In order to genotype a P. deltoides association population composed of 579 individuals, we used the P. trichocarpa version 2.2 reference genome to design probes for sequence capture. A total of 227,943 RNA-based 120 nucleotide long probes were designed to capture 23,835 intergenic regions (one probe every 15 Kb) and 18,153 genes (exons and part of 5'- and 3'-UTRs) previously found to be expressed in vegetative tissues in poplar. The targeted region corresponds to 27.35 Mb (5.7%) of the poplar genome. Sequence capture was performed on pools of 12 samples, and the captured DNA was sequenced using the Illumina HiSeq 2000 platform. We obtained an average of 27.3 million reads per genotype, and mean and median on-target depths of 26.1X and 25.6X, respectively. The percentage of bases sequenced at a depth above 15X (considered sufficient for identification of heterozygous variants) was 59.5%. Capture efficiency was high, with 100% of the genes and 96.3% of the intergenic regions captured.

Three different software (GATK HaplotypeCaller, FREEBAYES and SAMTOOLS) were used to call SNPs in the poplar association population, identifying 1.32 M, 1.31 M and 0.74 M SNPs respectively. In order to generate a high confidence SNP set, 529,628 SNPs identified by all three programs (with matching position and alternative allele calls) were selected. This consensus SNP set contains a high proportion (57.9%) of low frequency variants (MAF < 0.05). Rare variants (MAF < 0.005) constitute 23.3% of the total SNPs identified.

To avoid detection of spurious associations in the genome-wide association study (GWAS) we are carrying out to find significant marker-trait correlations, we assessed the presence of relatedness and population structure in our samples. The relatedness analysis performed with the software KING identified duplicated as well as closely related individuals that were excluded from further analyses. One

possibly hybrid individual was also removed. The remaining 430 unrelated samples were analyzed with the software STRUCTURE and the ΔK methodology, identifying two subpopulations that follow an east-west pattern. This structure needs to be accounted for when performing GWAS.

The association population was grown in a greenhouse and phenotyped for height, diameter, biomass production (separating leaves, stem and roots), and wood composition (lignin percentage, syringyl to guaiacyl lignin ratio and five-carbon and six-carbon sugar content). We are currently carrying out GWAS applying novel methods to detect marker-trait associations for rare variants, as well as traditional methods for common variants.

In summary, the multiplexed exome capture and resequencing protocol optimized in this project's first aim efficiently captured genes and intergenic regions and allowed genotyping the P. deltoides association population for a high number of SNP markers at reduced sequencing costs. This genotyping method is suited for the identification of rare variants. The consensus SNP set obtained after identifying variants with three different programs in the second aim of this project is expected to be of high quality. This approach has been used successfully in other studies to increase variant detection accuracy. With the necessary data already generated (genotypes, phenotypes and assessment of relatedness and structure in the population), the search for significant marker-trait associations underway in the project's third aim is currently under way and will be demonstrated in the investigator meeting.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0003893.

95. Quantitative Redox Proteomics and Ultrasensitive Targeted Proteomics for Functional Proteome Analysis

Jia Guo1, Amelia Y. Nguyen3, Matthew J. Gaffrey1, Tujin Shi1, James J. Moran2, Yi Qu1, Yuqian Gao1, Kim K. Hixson3, Ronald J. Moore1, Stephen R Lindemann1, Jim K. Fredrickson1, David W. Koppenaal3, Mary S. Lipton1, Richard D. Smith1,3, Himadri B. Pakrasi4and Wei-Jun Qian1* (weijun.qian@pnnl.gov)

1Biological Sciences Division, 2Signatures Sciences & Technology Division, and 3Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, 4Department of Biology, Washington University, St. Louis, MO

Project Goals: The main objectives of this early career research project are to develop novel proteomic approaches that will enable quantitative measurements of site-specific regulatory protein posttranslational modifications (PTMs) and targeted quantification approach for monitoring the dynamics of specific pathways. The ability to effectively and quantitatively characterize site-specific PTMs and pathways is essential for understanding the regulation of cellular signaling and protein functions, as well as for enabling a systems biology approach to study organisms as well as communities important for bioenergy or environmental applications. Our developments have been primarily focused on: (1) reversible redox modifications on protein thiols, and (2) ultrasensitive targeted quantification to enable measurements of low-abundance proteins or PTMs in a given pathway or network.

Quantitative redox proteomics: One of the major regulatory mechanisms is redox regulation where nearly all organisms sense subtle changes in their redox homeostasis due to environmental changes. Functional cysteinyl residues in proteins serve as “redox switches” through reversible oxidation, which is recognized as a fundamental mechanism of redox regulation in almost all organisms. We have developed a robust quantitative redox proteomics approach for measuring different types of reversible modifications on individual cysteine thiols to study redox regulation in metabolism or stress conditions of different organisms.1 The general principle of this approach involves the blocking of free thiols, selective reduction, enrichment, and quantification of site-specific redox modifications.

We have applied this redox proteomics approach to profiling the in vivo dynamics of thiol oxidation modulated by light/dark in Synechocystis sp. PCC 6803, an oxygenic photosynthetic prokaryote.2 Reversible protein thiol oxidation is an essential regulatory mechanism of photosynthesis, metabolism, and gene expression in photosynthetic organisms. The redox dynamics of ~2,100 Cys-sites from 1,060 proteins under light, dark, and 3-(3,4-ichlorophenyl)-1,1-dimethylurea (DCMU, a photosystem II inhibitor) conditions were quantified. In addition to relative quantification, the stoichiometry or percentage of oxidation (reversibly oxidized/total thiols) for ~1,350 Cys-sites was also quantified. The overall results revealed broad changes in thiol oxidation in many key biological processes, including photosynthetic electron transport, carbon fixation, and glycolysis. Moreover, the redox sensitivity along with the stoichiometric data enabled prediction of potential functional Cys-sites for proteins of interest. The functional significance of redox-sensitive Cys-sites in NADP-dependent glyceraldehyde- 3-phosphate dehydrogenase, peroxiredoxin (AhpC/TSA family protein Sll1621), and glucose 6- phosphate dehydrogenase was further confirmed with site-specific mutagenesis and biochemical studies. Together, our findings provide significant insights into the broad redox regulation of photosynthetic organisms. In our ongoing work, we have confirmed that S-Glutathionylation is a major form of the redox modification in Synechocystis 6803 and glutaredoxin enzymes play a key role in its regulation. Similarly, we have extended the approach to measure redox changes modulated by light/dark in a more complex microbial system, phototrophic mat-derived unicyanobacterial consortia.
Ultrasensitive targeted proteomics. In addition to global profiling, MS-based targeted proteomics using

selected reaction monitoring (SRM) offers unique advantages in its accuracy of multiplexed quantification, throughput, and sensitivity for enabling the measurements of the dynamics of proteins and PTMs belong to specific gene networks or pathways for both individual organisms and community samples. Traditionally, one of the challenges in targeted quantification is the insufficient sensitivity for low-abundance proteins. Recently, we achieved a significant advance in SRM sensitivity by developing the PRISM (high-Pressure and high-Resolution Separations coupled with Intelligent Selection and Multiplexing) technology3, which offers ~200-fold enhancement in sensitivity. PRISM enables direct quantitation of extremely low-abundance proteins (<100 copies per cell) and PTMs without the need for affinity enrichment. This high sensitivity of PRISM-based targeted quantification makes it ideal for monitoring temporal dynamics of specific sets of proteins or protein isoforms in complex systems such as plants and soil microbial community samples. Preliminary data on quantification of plant enzyme isoforms will be presented.

References:

1. Guo, J.; Gaffrey, M. J.; Su, D.; Liu, T.; Camp, D. G., 2nd; Smith, R. D.; Qian, W. J., Resin-assisted enrichment of thiols as a general strategy for proteomic profiling of cysteine-based reversible modifications. Nat Protoc 2014, 9 (1), 64-75.
2. Guo, J.; Nguyen, A. Y.; Dai, Z.; Su, D.; Gaffrey, M. J.; Moore, R. J.; Jacobs, J. M.; Monroe, M. E.; Smith, R. D.; Koppenaal, D. W.; Pakrasi, H. B.; Qian, W. J., Proteome-wide Light/Dark Modulation of Thiol Oxidation in Cyanobacteria Revealed by Quantitative Site-specific Redox Proteomics. Mol Cell Proteomics 2014, 13 (12), 3270-85.
3. Shi, T.; Fillmore, T. L.; Sun, X.; Zhao, R.; Schepmoes, A. A.; Hossain, M.; Xie, F.; Wu, S.; Kim, J. S.; Jones, N.; Moore, R. J.; Pasa-Tolic, L.; Kagan, J.; Rodland, K. D.; Liu, T.; Tang, K.; Camp, D. G., 2nd; Smith,
4. R. D.; Qian, W. J., Antibody-free, targeted mass-spectrometric approach for quantification of proteins at low picogram per milliliter levels in human plasma/serum. Proceedings of the National Academy of Sciences of the United States of America 2012, 109 (38), 15395-400.

This work is supported by the DOE Early Career Research Award (to WJQ) under the Office of Biological and Environmental Research (BER) in the DOE Office of Science and by the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences grant number DE-FG02-99ER20350 (to HBP). This work uses capabilities developed under the BER-supported Pan-omics project at PNNL and in the Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the DOE. AYN has been supported by an NSF Graduate Research Fellowship.

96. A novel method to retain viability of obligate anaerobic microorganisms during single cell sorting

Anne W. Thompson1, Matthew J. Crow2, Brian Wadey2, Christina Arens1, Serdar Turkarslan1, Sergey Stolyar1, Timothy W. Petersen2, Ger van den Engh3, David A. Stahl4, Nitin S. Baliga2, 5, 6, 7, Adam P. Arkin7, and Paul D. Adams7

1 Institute for Systems Biology, Seattle, WA, USA
2 BD Biosciences, Advanced Cytometry Group, Seattle, USA
3 Center for Marine Cytometry, Birdsview, WA, USA
4 Department of Civil and Environmental Engineering, University of Washington Seattle, WA, USA
5 Departments of Biology and Microbiology, University of Washington, Seattle, WA, USA
6 Molecular and Cellular Biology Program, University of Washington, Seattle, WA, USA
7 Lawrence Berkeley National Lab, Berkeley, CA, USA

Anaerobic microorganisms are major components in numerous environments ranging from naturally-occurring subsurface ground waters, anaerobic hydrocarbon deposits, and anoxic waters and sediments to man---made environments such as wastewater treatment plants and anaerobic digesters. Defining the community structure and the functional ecology of these microbial communities is a key component of research campaigns within the ENIGMA project. Flow cytometry is a powerful tool for high- throughput single cell analysis and sorting of targeted individual cells or populations, however preserving the viability of anaerobic cells during cell sorting has remained a challenge. We have developed a novel technology, called Live Anaerobic Cell Sorting (LAnCS), to address this challenge. Modifications made to the BD Influx high-speed cell sorter create an enclosed path cell sorter that is capable of maintaining an anoxic environment from sample to sort collection tube. We tested LAnCS on a synthetic anaerobic community composed of two organisms. Here, we demonstrate that LAnCS was successful in preserving the viability of both members of this anaerobic community, Desulfovibrio vulgaris Hildenborough (DvH) and Methanococcus maripaludis (Mmp), by growing colonies from single-cell sorts of both organisms on agar plates. This technology will provide the ENIGMA team with a novel tool for addressing campaigns such as Evolution and Assembly of Communities (Stahl), Environmental Microbial Characterization (Deutschbauer), Groundwater Microbiome (Hazen), and Time Series (Zhou). For example, LAnCS could help to investigate the role and importance of low abundance organisms that are not recovered in enrichment cultures, develop strategies to generate viable communities by mixing and matching organisms directly from field samples, and targeting functional groups in complex environmental samples for downstream study of viable cells.

Funding Statement.
Support was provided by the BD Biosciences Advanced Cytometry Group and 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.

97. Altered Biofilm Formation in Desulfovibrio vulgaris Hildenborough by Laboratory-Driven Evolution

Kara B. De León1* (deleonkb@missouri.edu), Grant M. Zane,1 Sara Altenburg,2 Adam P. Arkin,3 Matthew W. Fields,2 Judy D. Wall,1 and Paul D. Adams3

1 Biochemistry Division, University of Missouri, Columbia, MO 65211
2Center for Biofilm Engineering, Montana State University, Bozeman, MT 59715
3Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

http://enigma.lbl.gov

Project Goals: The goal of Ecosystems and Networks Integrated with Genes and Molecular Assemblies (ENIGMA) is to push the frontier of knowledge on the impact of microbial communities on ecosystems. As part of this, we must link genetic capacities to ecological function. Heavy-metal- and radionuclide-contaminated groundwater is common in DOE sites of interest. In these sites, the microbial communities are often predominantly attached to sediment particles as a biofilm. To better understand these communities, we must consider the biofilm growth strategy in a genetic context. Our goal is to determine the mechanisms and genetic requirements of biofilm formation in a common heavy-metal reducing bacterium, Desulfovibrio vulgaris Hildenborough (DvH). In pursuing this goal, inter-laboratory collaboration facilitated by ENIGMA has led to the discovery that two wild-type DvH strains (from MO and MT), both originally from ATCC 29579, have diverged in biofilm formation due to laboratory-driven evolution.

Desulfovibrio vulgaris Hildenborough (DvH) is a sulfate-reducing bacterium present in heavy- metal and radionuclide contaminated sites, often as a biofilm. Yet, the genetic requirements of DvH biofilm formation have not been determined. A biofilm reactor system has recently been established to address this lack of information. After optimization, it became apparent that the DvH-MO is partially deficient in biofilm formation as compared to data published for what should have been the same strain used by our ENIGMA collaborators at Montana State University (DvH-MT). In our reactors, DvH-MT performs as it does in Montana, forming a steady-state biofilm within 96h. DvH-MO biofilm lags for ~120h; however, upon reaching steady-state, the biofilms are compositionally similar. While both strains originated from ATCC 29579, laboratory-specific culturing resulted in a divergence in biofilm formation without direct selection under biofilm growth conditions. The genomes were re-sequenced from planktonic cultures of DvH-MT and DvH-MO, and DvH-MO steady-state biofilm. Twelve single nucleotide polymorphisms (SNPs) were ubiquitous in planktonic DvH-MO but were absent in DvH-MT. In contrast, no SNPs occurred at a frequency >50% in DvH-MT that were absent in DvH-MO. Of note in DvH-MO, a SNP in the ABC transporter (DVU1017) of a type I secretion system (T1SS) has resulted in an Ala to Pro change in a conserved α helix near the ATP-binding site. We hypothesize that this Ala to Pro change inhibits protein transport by the T1SS. DVU1012 and DVU1545 both contain T1SS export protein motifs and have been shown to be abundant in the DvH biofilm matrix. An inhibition of DVU1012 and/or DVU1545 transport may have caused biofilm deficiency in DvH-MO. In-frame deletion mutants of DVU1017, DVU1012, and DVU1545 are being constructed in DvH-MT. These mutants are being screened for biofilm phenotypes. In future work the Pro codon change will be introduced into DvH-MT and restored to Ala in DvH-MO. These results will clarify what has caused this difference in two supposedly identical cultures and determine genes required for biofilm formation in DvH. This discovery emphasizes the importance of monitoring laboratory-driven evolution, especially between collaborating laboratories, and provides insight into the genotype-to-phenotype relationship.

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.

98. Groundwater-fed Bioreactors Show Distinct Colonization and Community-wide Response Dynamics to Perturbations

Dwayne A. Elias1,4* (eliasda@ornl.gov), Andrew J. King1, Kathryn L. Bailey1, Sarah P. Preheim2, Michael S. Robeson II3, Taniya Roy Chowdhury1, Bryan R. Crable1, Richard A. Hurt Jr. 1,4, Anil C. Somenahally5, Stephen Techmann4, Tonia Mehlhorn1, Kenneth A. Lowe1, Andrea Rocha1,4,Anna Zelaya6, Matthew W. Fields6, Adam P. Arkin7, John-Marc Chandonia7, Terry C. Hazen1,4, Eric J. Alm9, Jizhong Zhou10, Tommy J. Phelps1, Craig C. Brandt1, Steven D. Brown1,4, Mircea Podar1,4, Michael W. W. Adams11, David B. Watson1, and Paul D. Adams7

1Oak Ridge National Laboratory, Oak Ridge, TN, USA 2Johns Hopkins University, Baltimore, MD, USA 3Colorado State University, Fort Collins, CO, USA 4University of Tennessee, Knoxville, TN, USA 5Texas A&M AgriLife Research, Overton, TX, USA, 6Montana State University, Bozeman, MT, USA, 7Lawrence Berkley National Laboratory, Berkley, CA, USA, 8University of Washington, Seattle, WA, USA, 9Massachusets Institute of Technology, Cambridge, MA, USA, 10University of Oklahoma, Norman, OK, USA, 11University of Georgia, Athens, GA, USA

ABSTRACT
Laboratory bioreactors have long been used for investigating the characteristics of a microorganism or simple synthetic community. However, for studying natural or in-situ microbial communities, discontinuous “snapshot in time” sampling has mainly occurred. In this study an in-field bioreactor system was developed to temporally monitor and manipulate the in-situ microbial community while maintaining the in-situ community structure.

Three above ground, in-field reactors were continuously fed microaerobic (0.2% O2) groundwater directly from an existing well at the Oak Ridge Field Research Center, Oak Ridge, TN, for 12 weeks. Each bioreactor contained 800 ml of groundwater and 8 replicate biofilm coupons filled with sterilized site sediment to monitor both the planktonic and biofilm communities. Gas influx was varied from aerobic (weeks 1-7, & 9) to anaerobic (weeks 8 & 10) to confirm that manipulation of bioreactor microorganisms was tractable. Samples from the incoming groundwater and from each bioreactor were taken every two days to match the dilution rate of the reactors. Each sample was analyzed by 16S rRNA sequencing at an average of 10,000 reads and key biogeochemical properties were measured including pH, dissolved oxygen, ORP, conductivity, 12 organic acids 14 anions and 3 sugars. At every third time point 53 different metals were also measured.

Community structure and diversity was highly similar across all three bioreactors according to 16S rRNA sequencing, representing 30-65% of the groundwater OTUs overall and 50-85% of high abundance groundwater OTUs. Biofilm coupons captured a unique subset of the groundwater OTUs but on average were only 27% similar to groundwater and 48% similar to the bioreactor planktonic samples. Community beta-diversity patterns indicated bioreactors were more different to the groundwater than expected if no growth was occurring, thereby suggesting growth in the reactors.

Correlations between organic acid profiles and bacterial clades revealed that the metabolic function was conserved across all three bioreactors and the in-situ groundwater community.

Transitions to anoxic conditions and subsequent lowering of the pH at weeks 7 and 10 resulted in strong, repeatable bacterial community and individual clade shifts toward the groundwater composition. However, not all bacterial groups in the bioreactors mirrored those in the groundwater. In fact, known metal- and organic acid-metabolizing clades increased in abundance in the bioreactors when the incoming groundwater increased in solute concentration despite no change in the incoming clade abundance.

Similarly, co-occurrence relationships of OTUs known for syntrophic and predator- prey interactions were observed over time in the bioreactors. This type of in-field bioreactor system allows for discreet temporal monitoring of microbial community structure and function simultaneously while allowing community responses to be determined during the testing of new strategies for environmental amendment or restoration on a small and affordable scale.

Keywords. Bioreactor, Microbial Community, Groundwater, Diversity, Anaerobic, Time Course

Funding Statement: The work conducted 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, 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.

99. High-throughput genetic characterization of environmental bacteria

Adam Deutschbauer1* (AMDeutschbauer@lbl.gov), Morgan N. Price1, Kelly M. Wetmore1, Jayashree Ray1, Jennifer V. Kuehl1, Robert J. Waters1, Romy Chakraborty1, Gareth Butland1, Matthew Blow1, James Bristow1, Adam P. Arkin1 and Paul D. Adams1

1Lawrence Berkeley National Laboratory, Berkeley, CA

http://enigma.lbl.gov

Project Goals: The Ecosystems and Networks Integrated with Genes and Molecular Assemblies (ENIGMA) program broadly seeks to understand the interactions between environmentally relevant microorganisms and their environment. One aim of this large interdisciplinary project is to rapidly bring environmental bacteria to model-organism status to enable systems-level investigations into microbial metabolism, regulation, stress response, and interactions under defined laboratory conditions. Here, we describe our development of random barcode transposon site sequencing, a high-throughput tool for screening of bacterial phenotype, and its application for gene function discovery in diverse environmental bacteria.

Transposon mutagenesis with next-generation sequencing (TnSeq) is a powerful approach to annotate gene function in bacteria, but existing protocols for TnSeq require laborious preparation of every sample before sequencing. Thus, the existing protocols are not amenable to the throughput necessary to identify phenotypes and functions for the majority of genes in diverse bacteria. Here we present a method, random barcode transposon-site sequencing (RB-TnSeq), that increases the throughput of mutant fitness profiling by incorporating random DNA barcodes into Tn5 and mariner transposons and by using barcode sequencing (BarSeq) to assay mutant fitness. RB-TnSeq can be used with any transposon and TnSeq is performed once per organism instead of once per sample. Each BarSeq assay requires only a simple PCR and 48-96 samples can be sequenced on one lane of Illumina HiSeq. We demonstrate the reproducibility and biological significance of RB-TnSeq with Escherichia coli, Phaeobacter inhibens, Pseudomonas stutzeri RCH2, Shewanella amazonensis SB2B, and Shewanella oneidensis MR-1. To demonstrate the increased throughput of RB-TnSeq, we performed 387 successful genome-wide mutant fitness assays and identified 5,196 genes with significant phenotypes across the five bacteria. To support ENIGMA science, we have applied RB- TnSeq to 3 genetically diverse Pseudomonas strains isolated from the ORNL iFRC site and identified phenotypes for over 1,000 genes from each isolate. Lastly, in collaboration with ENIGMA researchers, we are applying RB-TnSeq to dissect mechanisms of microbial interactions, identify genetic determinants of metal metabolism, and identify epistatic (genetic) interactions.

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.

100. Microbial Interactions with Native Natural Organic Matter in Groundwater and Sediment from the Oak Ridge FRC

Romy Chakraborty1*(rchakraborty@lbl.gov), Angelica Pettenato1, Xiaoqin Wu1, Sindhu Jagadamma2, Terry C. Hazen2, Matthew Fields3, Trent Northen1, Stefan Jenkins1, W Andrew Lancaster4, Mike W.W Adams4, Adam P. Arkin1 and Paul D. Adams1

1Lawrence Berkeley National Laboratory, Berkeley, CA; 2Oak Ridge National Laboratory, Oak Ridge, TN; 3Montana State University, Bozeman, MT; 4University of Georgia, Athens, GA,

http://enigma.lbl.gov

Project Goal: Natural organic matter (NOM) is central to microbial food webs; however, little is known about the interplay between physical and chemical characteristics of NOM and it's turnover by microbial communities in groundwater and sediments of Oak Ridge Field Research Center (FRC). Microbe mediated molecular and geochemical mechanisms control the flow of carbon that support evolution and maintenance of a community within a given environment. To gain insight into the cycling of carbon, and how the turnover of NOM regulates the development of indigenous microbial community at the FRC, it is critical to extract NOM, identify its characteristics, and correlate this to the structure and composition of microbial communities and to the metabolic potential of that community.

In this FY15 Discovery project, we focused on NOM from the background well, FW305. Preliminary experiment initiated with FW305 sediment enriched with 10ppm fulvic acid (as part of the Microparticles Campaign), showed rapid evolution of CO2. Isolations from this fulvic acid enriched samples yielded strains belong to Sphingomonas, Pseudomonas, Undibacterium, Rugamonas, Rhodococcus genera. The same sediment sample (FW305) was tested for NOM extraction methods using three mild solvents, e.g., phosphate buffered saline (PBS), pyrophosphate, and warm de-ionized water with shaking or mild agitation. The pH was kept close to in-situ conditions, 6.5–7.0. We also tested the efficacy of glass beads as a physical abrasive to aid in extraction. The dissolved organic carbon (DOC) in the extracted NOM was in the range of 8.8–28 mg/L. Sonication with PBS extracted higher levels of DOC, while shaking with pyrophosphate extracted higher levels of metals such as iron, zinc, cobalt and manganese when compared with other tested methods. A spectrophotometric scan of the sample with highest DOC content showed high absorbance between 300-380nm, indicative of presence of functional groups with high aromaticity such as carboxylic and phenolic groups. The extracted NOM was fractionated using molecular weight cut-off filters (100 kDa). In ongoing experiments, we are testing existing isolates from FW 305 for their ability to transform the NOM, and identify the transformation product. In addition, we will add this fractionated NOM to feed native microbes present in the groundwater of well FW305 to identify the microbial isolates/community that preferentially grew with this NOM.

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

101. Mutations in the Gene Regulatory Network Facilitate Adaptation to Variable Environments

Serdar Turkarslan1, David. J. Reiss1, Sergey. Stolyar1, Nina. Arens1, Jason Flowers2, Grant. M. Zane3, Thomas R. Juba3, Judy D. Wall3, David A. Stahl2, Adam P. Arkin4, Nitin S. Baliga1*(Nitin.Baliga@systemsbiology.org), and Paul D. Adams4

1Institute for Systems Biology, Seattle WA; 2University of Washington, Seattle, WA; 3University of Missouri, Columbia, MO; 4Lawrence Berkeley National Lab, Berkeley CA

http://networks.systemsbiology.net/dvh

Project Goals: The main goal of this project is to delineate regulatory network readjustments that drive adaptation to variable environments and impact trajectory of evolution. We are investigating the fitness benefit of accumulating mutations in a GRN or the precise mechanism(s) by which they facilitate evolution of a generalist. Ultimately, we aim to develop predictive understanding of organismal and community fitness under dynamically changing environments.

Abstract: Upon subjecting generalists to long-term directed evolution in uniform growth conditions, mutations in regulatory genes often appear early, and prior to erosion or specialization of metabolic functions. Given that regulation is important for effecting physiological readjustments, we predicted that regulatory mutations are less likely to be adaptive to a variable environment. Unexpectedly, the generalist Desulfovibrio vulgaris Hildenborough (DvH) rapidly went to extinction when it was required to repeatedly switch between two principal ecologically-relevant lifestyles: an independent sulfate-respiring lifestyle and a lifestyle that does not require sulfate-respiration but depends on syntrophic interaction with a methanogenic archaeon -- Methanococcus maripaludis (Mmp). Disruptive mutations in the DvH regulatory network, by contrast, stabilized this dynamic lifestyle switching, significantly prolonging time to extinction. Global transcriptome analyses demonstrated that the regulatory mutations facilitated adaptation by potentially minimizing the energetic burden associated with up and down regulating genes during environmental shifts. Even under uniform selection pressure, the serial transfers that are necessary for long-term evolution experiments also impose repetitive environmental shifts, explaining why regulatory mutations appear early.

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

102. Novel single-cell sequencing methods link target genes with bacterial hosts

Sarah J. Spencer1* (sjspence@mit.edu), Manu V. Tamminen2, Sarah P. Preheim3, Zhili He4, Andrew Lancaster5, Michael Thorgersen5, Michael Adams5, Aindrila Mukhopadhyay6, Jizhong Zhou4, Eric J. Alm1, Adam P. Arkin7, and Paul D. Adams6

1Massachusetts Institute of Technology, Cambridge, MA; 2ETH Zürich, Switzerland; 3Johns Hopkins University, Baltimore, MD; 4University of Oklahoma, Norman; 5University of Georgia, Athens, 6Lawrence Berkeley National Laboratory, Berkeley, CA; 7University of California, Berkeley

Project Goals: We aim to deconstruct microbial ecosystems by creating assays at the scale of individual microbes. Current methods to profile single cells often restrict throughput and require expensive, technically challenging equipment. To fill this technological gap we developed a benchtop protocol that uses emulsion droplets as microreactors for targeted gene amplification. With the ability to link genomic information from single cells, a variety of applications open up, including linking functional genes to species and linking phage or plasmids to their host cells. High-throughput single-cell assays will add a missing dimension to bulk microbial surveys, which are unable to dissect flexible, dynamic microbial genomes. This will in turn inform models and efforts to observe, predict, and direct microbial function at perturbed environmental sites.

Our ability to understand and manipulate microbial ecosystems is restricted by a lack of information at scales relevant to individual microbes. This lack of information stems from the technology available; bulk sequencing techniques lose the mapping between species and their genomic content, while single-cell sequencing limits throughput. To provide an alternative, our group developed a novel method to amplify pairs of genes from single cells, while maintaining high throughput. The technique, termed Emulsion, Paired Isolation, and Concatenation PCR (epicPCR) uses emulsion droplets as nanoliter reactors to assay up to 1 million cells in parallel, with cells that can be loaded directly from environmental samples.

Our first application of single-cell gene fusions links together functional genes with phylogenetic indicator genes. We validated the protocol by assaying for sulfate reduction genes in a well- documented aquatic ecosystem. We targeted the dissimilatory sulfite reductase gene (dsrB) and physically linked it to a species-indicator gene (16S rRNA) within the same cell. The resulting data showed clean selection for only species carrying the dsrB gene, and these species all fell into well-documented sulfate-reducing clades including deltaproteobacteria. This demonstrates our ability to load a complex community, and then selectively amplify the species which carry a target gene of interest.

Expanding from our proof-of-principle, we have projects to link function to phylogeny in samples relevant to multiple campaigns. Together with Mike Adams at the University of Georgia, we are investigating metal metabolism at contaminated aquifer sites. The metabolism of heavy metals is dependent on molybdenum-utilizing enzymes, but these enzymes appear widely distributed within diverse, contaminated sampling sites. epicPCR can identify molybdenum-utilizing species within aquifers as well as within engineered bioreactors managed by the Elias Lab at Oak Ridge National Laboratory in Tennessee.

In addition, we are engineering a scale-up of this method to profile a large number of functional genes in parallel – providing a way to capture the niches of species within a community. Together with Jizhong Zhou at the University of Oklahoma, we compiled a list of prevalent metabolic genes in perturbed aquifer communities. These target genes contribute to an epicPCR panel linking functional genes to species. This panel, termed PuLSE (Panel Linking Species to Ecology), provides a detailed accounting of species along with the metabolic roles they are fulfilling within a given site.

The versatility of our emulsion-based protocol also allows for assays of spatial relationships beyond single-cell genomes. Instead of looking at how two genes from a single cell co-locate, we can assay how two genes from different cells (or plasmids, or phage) co-locate. By creating droplets containing plasmids or phage of interest, we are querying microscale host relationships. A discovery project, in collaboration with Aindrila Mukhopadhyay at Lawrence Berkeley National Laboratory, explores the host relationships of extrachromosomal plasmids. Our lab is also leading the effort to link phage functional genes (terminase, integrase, etc.) with bacterial host genes. Both projects support multiple groups from our SFA, informing studies of the groundwater microbiome as well as basic questions in evolution and community assembly.

Our ultimate goal is to offer consortium investigators a window into the myriad microscale relationships that complicate the study of complex environmental samples. We present a technique that not only removes inference from community functional analysis, but offers novel opportunities to study the spatial relationships between species and mobile genetic material.

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.

103. Screening for Genetic Interactions in Desulfovibrio vulgaris Hildenborough

Valentine Trotter1 , Maxim Shatsky1, Avneesh Saini1, Morgan N. Price1, Kelly M.Wetmore1, Thomas R. Juba2, Adam P. Arkin1, Judy D. Wall2, John-Marc Chandonia1, Adam M. Deutschbauer1, Gareth P. Butland1* (GPButland@lbl.gov) and Paul D. Adams1

1Lawrence Berkeley National Laboratory, Berkeley CA; 2University of Missouri, Columbia, MO.

http://www.enigma.lbl.gov

Project Goals: The Biotechnology component of ENIGMA seeks to develop and apply a suite of technological tools in order to expand our capability to obtain systems-levels insights into microbial activity. Such insights will form the foundation for predictive models of key microorganisms from a single environment that will serve as valuable resources for assessing ecological questions relevant to microbial community structure and function. Using environmentally relevant metal and sulfate-reducing bacteria, ENIGMA researchers have developed tools which enable evidence-based annotation of gene function using high-throughput mutagenesis and extensive phenotyping of mutants under stress conditions; evidence-based annotation of transcripts using tiling microarrays and RNAseq; protein complex isolation and identification; mass spectrometry-based proteomics and metabolomics analysis; and high resolution imaging. Here we expand our capabilities by applying a recently developed sequencing based fitness profiling approach (RB-TnSeq) to rapidly screen for genetic interactions (fitness changes in double mutant strains) across multiple growth conditions.

Gene fitness profiling of single mutant strains can routinely identify a statistically significant phenotype for many genes. However, many mutants display only subtle phenotypes which are hard to interpret. This phenomenon is likely attributable to partial functional redundancy between genes in different pathways masking effects of single mutations. Genetic Interactions (GI) describe epistatic relationships between genes and are the basis for functional redundancy between specific pathways and processes within a cell. A common way to screen for genetic interactions is to introduce a second mutation into a collection of single mutants and to assess the compounded effect of the second mutation. Combinations of mutations resulting in a greater or reduced fitness defect compared to either single mutation alone are genetically interacting. Genetic interactions can therefore not only result in stronger, more specific fitness defects due to removal of redundant processes, but may provide more detailed information about gene function such as the order in which genes function in a pathway. We are utilizing a TnSeq- based procedure in which each transposon contains a unique 20 nt molecular barcode (RB- TnSeq; A. Deutschbauer). When applied to strains containing mutations, this approach will allow us to generate double mutant libraries rapidly and detect both synthetic lethal and conditional GI by monitoring the abundance of barcodes under stress conditions. Initial experiments with Desulfovibrio vulgaris Hildenborough (DvH) wild type have successfully resulted in a library with 329,755 transposon insertions across 181,020 distinct locations. A total of 303 genes were found to have no transposon insertion within 5-85% of their coding sequence and were considered as potentially essential genes. We will present a comparison of these data with results from previous methods (Fels et al. AEM 2013). The DvH genome encodes at least three formate dehydrogenase (FDH) isozymes (DVU0587-0588, DVU2481-2482 and DVU2811-2812) and FDH activity is essential for cells to grow using formate as a primary carbon source. A pilot study is ongoing using strains harboring unmarked deletions (JW710 Δupp background) in these FDH enzyme-encoding genes clusters. We have performed preliminary experiments to generate double mutant libraries and have observed synthetic lethal GI based on the differences observed between DvH wild-type and JW710 strains. Indeed, we have been unable to interrupt a number of genes annotated as being involved in de novo synthesis of UMP in JW710-based strains. The inviability of these double mutants is likely to be the result of synthetic lethal GI between these genes and upp which encodes uracil phosphoribosyltransferase,

a key enzyme in the salvage pathway for UMP synthesis from uracil. This provides a gold standard, positive control for further GI screening in the FDH mutant background and additional markerless deletion mutants.

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

104. Temporal Variation in Groundwater Geochemistry and Microbial Community Structure at the Oak Ridge Field Site

Andrea M. Rocha1, 2* (rochaam@ornl.gov), Tonia L. Mehlhorn1, Jennifer E. Earles1, Kenneth A. Lowe1, Dawn M. Klingeman1, David B. Watson1, Dominique C. Joyner1, 2, Julian L. Fortney1, 2, Sindhu Jagadamma1,2, Jizhong J. Zhou3, Joy D. Van Nostrand3, Michael W. W. Adams4, Farris L. Poole4, W. Andrew Lancaster4, Romy Chakraborty5, Dwayne Elias1, Paul D. Adams5, Adam P. Arkin5, Eric J. Alm6, and Terry C. Hazen1, 2

1 Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, 2 University of Tennessee, Knoxville, TN, 3University of Oklahoma, Norman, OK, 4University of Georgia, Athens, GA, 5Lawrence Berkeley National Laboratory, Berkeley, CA, 6Massachusetts Institute of Technology, Cambridge, MA

http://enigma.lbl.gov

Project Goals: The goal of the ENIGMA field microbiology component is to identify key microbial populations and determine the community events and mechanisms of these populations that impact and control environmental activities of interest, ultimately predicting how perturbations of the environment may affect community structure and function. From these results, we can develop models that can be applied to microbial populations overlaid with geochemical parameters and engineering controls. An example of such a model is described in the 100-well survey where we demonstrated the ability to utilize natural bacterial communities as in situ environmental sensors that capture environmental perturbations and elucidate key systems biology features (Smith et al., submitted). Here, we expand upon the survey to capture the microbial community response to temporal changes in the groundwater geochemistry to (1) provide for temporal tuning of predictive models, and (2) to determine if and to what extent geochemical variation in groundwater affects microbial community, activity, and genetic diversity along different well depths.

Across aquatic and terrestrial environments, numerous studies have sought to characterize key microbial communities and to identify factors that drive changes in microbial community structure and activity. While these studies enable us to further understand and potentially uncover key correlations between the composition of microbial communities and their environment, information regarding temporal community dynamics is often limited or in many cases lacking. One such example is the Oak Ridge Field Research Center (ORFRC) where there has been a large focus to characterize the spatial distribution of groundwater and soil microbial communities across different geochemical transects (e.g. Uranium-Nitrate-pH). In this time- series study, we aim to bridge this gap by capturing the spatio- and temporal variation of geochemistry on the overall structure, function, and genetic diversity of the groundwater microbial communities in the groundwater wells at the ORFRC. Here, we present our findings from two pilot studies conducted during November 2013 and from November 2014 - January 2015, respectively, at the ORFRC background field site.

One of our main objectives of the temporal study is to determine how resilient (or volatile) microbial communities are to daily and weekly changes in groundwater geochemistry. To capture changes in microbial community structure and geochemical constituents, we initially sampled two deep and two shallow groundwater wells over the course of three weeks during November 2013. For each well and time-point, groundwater samples were collected for geochemical and microbial communities analyses. Nucleic acids were collected by filtering water through a 10.0µm pre-filter and 0.2µm membrane filter and then extracted using a Modified Miller method. Results from the study showed that geochemical measurements across all the wells remained fairly stable over the course of the study. However, a decrease in pH and increase in conductivity measurements was observed in the shallow wells during

small rain events. Unlike the geochemistry, the 16S rRNA sequencing of the microbial community structure within each well varied on a daily basis in both the 0.2 µm and 10.0 µm size fraction. Statistical analysis of the 16S data using Adonis indicated that there were statistically significant differences (p-value = 0.0001) in the community structure between wells throughout the study. Nonmetric multidimensional scaling analysis of the community structure did not show distinct differences between communities present in shallow versus deep well depth. However, analysis indicated that communities present in wells FW-300 (shallow) and GW-460 (deep) were much more variable throughout the time course.

Due to the stability of the geochemistry in the 2013 pilot study, we hypothesize that the variation in microbial communities is a result of both a sampling effect and from the inclusion of communities associated with biofouling in the well casing and sloughing off during sampling. To determine if the daily variation in groundwater microbial community profiles in the first study were naturally occurring within the aquifer or if (and to what extent) the variation in community structure is a result of factors associated with biofouling or sampling artifacts, we sampled six wells from the background site from November 2014 – January 2015. Of the six wells, we physically and chemically cleaned four wells to remove biofilm and attached particulates from the well casing. The remaining two wells that weren't cleaned, served as controls. Prior to cleaning, all wells were sampled to establish a baseline microbial community profile. Post- cleaning, each well was sampled a total of twelve times. For each well and time-point, groundwater samples were collected for geochemical and microbial community analyses.
Currently, the microbial and geochemical data are still being analyzed although the preliminary results indicate geochemical variation in response to rain events during the course of the study.

Overall, results from both pilot studies suggest evidence of geochemical and microbial response within select wells in response to rain events. Findings from these will enable ENIGMA campaigns for more specialized questions on microbial community structure, provide for temporal tuning on environmental models, and further our understanding of the natural temporal variations versus external factors, such as biofouling or sampling-related effects on microbial communities.

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

105. Towards Rapid Characterization of Microbial Communities, Interactions, and Community Assembly

N.B. Justice1* (nbjustice@lbl.gov), S. Kaur2, S. Kosina1, A. Sczesnak2, A.M. Deutschbauer1, T. Northen1, A.P. Arkin1,2, P.D. Adams2

1Lawrence Berkeley National Laboratory, Berkeley, CA; 2University of California, Berkeley, CA.

Project Goals: Our major goals are 1) the development of a high-throughput method to accurately measure the absolute abundances of multiple species in microbial assemblages, 2) the identification of microbial interactions amongst cultivated FRC isolates, 3) the demarcation of the fundamental niche of these isolates across multiple environmental dimensions, and 4) a predictive understanding of interactions and niche behavior based on genetic elements.

Abstract: Microbial communities are shaped by a complex network of interaction between different species (e.g., competition or mutualism) as well as selection from the environment (e.g., geochemical parameters). Understanding how these forces influence community assembly remains a difficult task—limiting our ability to predict community response to perturbation. Towards this end, we are developing a high-throughput assay to rapidly generate time-resolved measures of absolute abundance of multiple species in co-culture. The method is based on 1) a set of dual-indexed primers amplifying variable regions of the 16S rRNA, and 2) application of a "spike-in" standard organism and species-specific standard curves which allow quantitation of the number of cells of a given organism within each assemblage. The primers used amplify a ~450 bp variable region of the 16S rRNA gene (V4/V5) and contain 5' adapter sequences with everything necessary for binding and sequencing on Illumina MiSeq and HiSeq platforms. We present a general overview of the efficacy of this approach and the efforts that have been made to reduce both random and systematic bias from the measurement. We further present a demonstration of the method in the context of several co-culture experiments using different numbers of organisms within different environments. Together with time-resolved population measurements, we applied exometabolomic footprinting to explore the mechanisms of interspecific interactions and degrees of resource utilization overlap. Ultimately, we aim to use exometabolomic footprinting and high-throughput genetics together with precise population tracking to understand the mechanisms and principles that shape the ecology of complex microbial assemblages across an increasingly realistic set of controlled environmental parameter regimes.

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

106. Understanding the Role of Pelosinus In Uranium-Contaminated Environments

Andrew Lancaster1, Michael P. Thorgersen1, Alexey E. Kazakov4, Farris L. Poole1, Kara De Leon2, Sagar Utturkar3, Pavel S. Novichkov4, Judy D. Wall2, Steven D. Brown3, Kelly S. Bender5*(bender@micro.siu.edu), Michael W. W. Adams1, Adam P. Arkin4, and Paul D. Adams4

1University of Georgia, Athens, GA; 2University of Missouri, Columbia MO; 3Oak Ridge National Laboratory, Oak Ridge, TN; 4Lawerence Berkeley National Laboratory, Berkeley, CA; 5Southern Illinois University, Carbondale, IL

http://enigma.lbl.gov

Project Goals: The overarching goal of Ecosystems and Networks Integrated with Genes and Molecular Assemblies (ENIGMA) is to understand environmentally relevant microbial communities and their environmental interactions. With many environments being impacted by varying concentrations of a range of metals, it is critical to elucidate the mechanisms that microbial communities employ to tolerate, to assimilate, and to metabolize metals. Our goal in this interdisciplinary project is to understand the genomics, regulation, metabolism, and uranium assimilation of Pelosinus, a dynamic genus that includes U(VI)- and Cr(VI)-reducing strains isolated from DOE sites of interest.

Pelosinus is a genus of the strictly anaerobic phylum Firmicutes possessing strains with U(VI)- and Cr(VI)-reducing capabilities that have been isolated from the Hanford Nuclear Reservation (HNR) 100-H and Oak Ridge Field Research Center (ORFRC)1,2,3. To further understand the environmental significance and metabolic capabilities of Pelosinus, our group has selected two of these strains, UFO1 and JBW45, for further genomic and phenotypic analysis. These were isolated from the HNR and ORFC, respectively2,3. Comparative analysis of the complete UFO14 and JBW45 genome sequences2 indicated genome sizes of 5.1-5.3 MB containing a range of genes that suggest diverse metabolic capabilities. Phenotypic analyses indicated that both strains could be cultured in the laboratory on R2 media utilizing fumarate as a substrate.

Because previous U(VI)-reducing studies on Pelosinus sp. strain UFO1 have shown both extracellular binding of U and intracellular deposition of U precipitates3, we performed metal analyses to further understand U assimilation in strain UFO1. While this analysis indicated that strain UFO1 is inhibited by 200 μM U, it also resulted in the isolation of U- binding membrane proteins. These proteins were identified as UFO_4202-4203, which contain S-layer domains and combine to form a major protein found both in the membrane and secreted into the growth medium. Subsequent metal analysis of strain JBW45 also identified an extracellular U-binding protein containing an S-layer domain with 78% and 42% similarity to UFO_4202 and UFO_4203, respectively. Based on the metabolic potential and unique metal biochemistry of UFO1 and JBW45, regulatory analysis was also performed. A comparative study of transcription regulatory genes from the families of the metal sensing regulators revealed a substantial difference between UFO1 and JBW45 strains that may contribute to phenotypic diversity of these strains.Only four of ten MerR family regulators from UFO1 are conserved in JBW45, and only three of seven ArsR family regulators from JBW45 are present in UFO1. We also applied a conservative propagation procedure to build 26 draft regulons in UFO1 using a collection of manually curated regulons from 11 Bacillales genomes from the RegPrecise5 database. Genome analysis of UFO1 also identified 52 potential riboswitch elements. These elements include switches with ykkC/yxkD and yybp/ykoy motifs that have been shown to control efflux pumps/multi-drug resistance and manganese homeostasis/tellurium resistance in other bacteria, respectively. To further understand the biology and regulation of strains UFO1 and JBW45, genetic system development is currently underway with mutant construction by conjugation showing promise. Future collaborative work to elucidate the unique U metabolism of UFO1

and JBW45 includes further characterization of U-binding proteins, RNA-Seq to identify U-responsive genes, regulons and sRNA analyses, as well as isolation of new strains from wells containing high levels of U from the ORFRC.

References

1. Beller, H.R. et al., 2013. Genomic and physiological characterization of the chromate- reducing, aquifer-derived Firmicute Pelosinus sp. Strain HCF1. Appl. Environ. Microbiol. 79: 63-73.
2. De Leon K.B., Utturkar S.M., Camilleri L.B., Fields M.W., Brown S.D., and Wall J.D. 2015. Complete genome sequence of Pelosinus fermentans JBW45 (in preparation).
3. Ray, A.E. et al., 2011. Evidence for multiple modes of uranium immobilization by an anaerobic bacterium. Geochim Cosmochim. Acta 75: 2684-2695.
4. Brown, S.D. et al., 2014. Complete genome sequence of Pelosinus sp. strain UFO1 assembled using single-molecule real time DNA sequencing technology. Genome Announc. 2: e00881-14.
5. Novichkov, P.S., et al., 2013. RegPrecise 3.0--a resource for genome-scale exploration of transcriptional regulation in bacteria. BMC Genomics. 1;14:745.

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.

107. 'Species Filter' Effects on Sediment Biofilms and Groundwater Source Diversity

Zelaya1,2, K. Bailey3, P. Zhang4, S.P. Preheim5, J. Van Nostrand4, D.A. Elias3,7, E.J. Alm5,7, J. Zhou4,7, P. Adams6,7, A.P. Arkin 6,7, and M. W. Fields1,2,7 (matthew.fields@biofilm.montana.edu)

1Center for Biofilm Engineering, Bozeman, MT; 2Department of Microbiology & Immunology, Montana State University, Bozeman, MT; 3Division of Environmental Sciences, Oak Ridge National Laboratory, Oak Ridge, TN; 4Institute for Environmental Genomics, University of Oklahoma, Norman, OK; 5Department of Biological Engineering, MIT, Boston, MA, 6Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA; 7ENIGMA

http://enigma.lbl.gov/

Project Goals: The goal of Ecosystems and Networks Integrated with Genes and Molecular Assemblies (ENIGMA) is to push the frontier of knowledge on the impact of microbial communities on ecosystems and to elucidate microbial assembly rules. As part of this, we must link genetic capacities to ecological function (e.g., heavy metal and radionuclide---contaminated groundwater in DOE sites of interest). In these sites, the microbial communities are often predominantly attached to sediment particles as a biofilm. We have characterized temporal dynamics of bacterial groundwater communities and compared to the establishment of particle-associated bacterial communities on native sediments incubated down-well. In pursuing this goal, inter-laboratory collaboration facilitated by ENIGMA has led to the identification of aquifer population distributions between the liquid/solid boundaries in situ.

Understanding the factors that determine microbial assembly, composition, and function in subsurface environments are critical to assessing contributions to biogeochemical processes such as carbon cycling and bioremediation. However, these factors are still not fully understood. In this study, surrogate sediment samples were incubated for 3 months in 3 wells (FW301, FW303, FW305) within the background site of the Oak Ridge Field Research Center in Oak Ridge, TN. Local sediment biofilm communities were compared to those of the groundwater (source diversity). Groundwater samples from each well were collected approximately 3 times a week. Multiple sediment samples (n=12) were used per well to determine inter--- and intra---well variation. Spatial and temporal community analysis of local and source samples via ss---rRNA paired---end sequencing and distribution---based clustering revealed higher richness, diversity, and variability in source groundwater communities compared to sediment---associated communities. Ordination analysis grouped the newly formed local communities as more similar to those of the groundwater than to those of the original parent sediment. The predominant groundwater sequences per well were Curvibacter, Delftia, and Acidovorax for FW-301, Aquabacterium, Oleiphilus, and Bradyrhizobium for FW---303, and Acidovorax, Curvibacter, Caulobacter, and Elusimibrobium for FW---305. Other sequences displayed transitory predominance for different wells. The community composition was different between wells, and FW---305 (a younger well) showed more diversity over time. In sediment samples, 20---40% of the communities consisted of populations that were abundant at less than 5% of the total sampled diversity. The sediment biofilms from each well were also distinct from each other. Intra---well sediment biofilms showed much less variability, with the exception of FW---305. Sediment biofilm communities were distinct from corresponding groundwater communities, with some populations becoming predominant in the biofilm (e.g. Aquabacterium, Perlucidibaca, and Paraperlucidibaca); however, different OTUs were respective to each well. These results indicate a shift in local community structure that is influenced by the available source community as well as hydrology.

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.

108. ENIGMA Sediment Metabolomics

Author(s): Stefan Jenkins1* (sjenkins@lbl.gov), Richard Baran2, Romy Chakraborty1, Adam M. Deutschbauer1, Adam P. Arkin1, Trent R. Northen1, Paul D. Adams1

1Lawrence Berkeley National Laboratory, Berkeley, CA; 2Thermo Scientific, San Jose, CA.

http://enigma.lbl.gov

Project Goals: 1) Sediment metabolomic methods development 2) Examine the diverse substrate pool available to support growth in FRC sediment

Abstract: One aim of the ENIGMA SFA is to understand the relationship between microbial community structure and environmental parameters. A critical yet poorly understood parameter is the composition of the organic pool that serves as electron donors at our FRC field site. Here we describe our efforts to use liquid chromatography mass spectrometry (LC/MS) to characterize the available soluble substrate pool from FRC sediment samples. To maximize the number of metabolites detected we used two types of stationary phases, hydrophilic interaction liquid chromatography (HILIC) for polar metabolites and reverse phase (RP) for non-polar metabolites. Sediment metabolomic methods were initially developed using a library of 102 compounds representative of multiple classes of metabolites (organic and amino acids, sugars, lipids, fatty acids, vitamins, etc.). These methods were then used to characterize aqueous extractions of chloroform fumigated FRC sediment. This revealed that FRC sediments contain a wide range of mono-, di- and oligosaccharides, amino acids, fatty acids, lipids, nucleobases, nucleotides and many novel metabolites. These results are informing the design of medias for investigation of microbial assemblies ultimately to determine the role of the substrate pool in structuring FRC communities.

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

109. Characterizing the Pseudomonas Sulfur Regulome

Sarah Zerbs, Peter Korajczyk, Frank R. Collart, and Peter E. Larsen* (plarsen@anl.gov)

Biosciences Division, Argonne National Laboratory, Lemont, IL

http://www.bio.anl.gov/molecular_and_systems_biology/protein_expression

Project Goals: The Argonne "Environment Sensing and Response" Scientific Focus Area (SFA) program seeks to identify the molecular basis of cellular transport and sensory pathways that mediate the response of terrestrial ecosystems to environmental nutrients. The mechanistic links between and within ecosystems comprised of plants, fungi, and soil bacteria involved in the production of biomass for fuel are currently very poorly defined. The effects of nutrient availability, closely linked to climate, on those mechanistic links, are also inadequately understood. This program will address this knowledge gap by mapping transport and sensor proteins to specific environmental compounds to define their function and biological roles and establish a series of defined connections between the environment and the cell. The knowledge will facilitate the development of system-level models predictive of cellular response to changes in environmental conditions.

In soil communities with sufficient N, P and C inputs available sulfur (S) can become limiting due to competition between community members. Like Phosphorus, the preferred sulfur source is inorganic sulfate, which composes only ~5% of total S present in soil. The remaining sulfur is present as other inorganic species and various organosulfur compounds. Organic sulfur compounds are structurally diverse, with common functional groups including thioesters, thioethers, thiols, sulfate esters and sulfones. Previous studies have shown that the soil sulfur pool is very dynamic, with S rapidly transformed between many inorganic and organic S compounds.

Plants can only take up S as inorganic sulfate or as amino acids methionine and cysteine so there is a requirement for microbial activity to transform complex molecules into plant-available forms. In many cases bacteria can remove a sulfur groups from organic sulfur compounds even when the carbon skeleton cannot be metabolized. In many organisms, elucidation of the transporters that mediate uptake or enzymes associated with sulfur-processing pathways has been limited as bacteria are frequently cultured in media that contains inorganic sulfur at concentrations which fully repress organic sulfur metabolism pathways. Furthermore, even under de-repressed conditions sulfur-metabolizing proteins are often expressed at low levels and are not detected by proteomics-based identification methods. This presents an advantage to transcriptomics-based interrogation of these microbial pathways as RNA-seq is sensitive to small changes in transcript levels. In this study, we characterize the expression profiles associated with bacterial growth on a variety of sulfur compounds to determine the linkage to specific metabolic pathways and transporters. The goal is to develop a modeling approach useful in predicting the Pseudomonas sulfur regulome for a variety of conditions.

Transcriptomics
Transcriptome libraries were generated for P. fluorescens SBW25 cultured on minimal growth media supplemented with 2-Aminoethyl hydrogen sulfate, Cysteine, Glutathione, 4-Nitrophenyl sulfate, Methionine, Methionine Sulfone, α-keto-γ(methylthio)butyricacid, Sodium Sulfate, or Taurine as sole sulfur sources. Pseudomonas was able to utilize all of these sulfur compounds in liquid culture but showed variations in total biomass after 24 hours of growth. 327 genes were found to be significantly differentially regulated in response to sulfur media type, of which 88 genes coded for metabolic enzymes, 457 for transmembrane transporters, and 14 for transcription factors. Functional analysis of differentially regulated transporters indicate that amino acid transporters and transporters that maintain osmotic balance

are associated with the Pseudomonas sulfur regulome. Changes in expression of genes for metabolicenzymes indicate that the net effect of transcriptional regulation of metabolic processes do not favor any particular sulfur metabolic pathway, but instead act to maintain metabolic homeostasis. On sulfur media that was least favorable for growth (2-Aminoethyl hydrogen sulfate and Methionine Sulfone) gene expression was significantly negatively correlated with translated protein sulfur content. On sulfur media that was most favorable for growth (Cysteine, Glutathione, and Sodium Sulfate) gene expression was significantly positively correlated with translated protein sulfur content. This indicates that Pseudomonas actively regulates it proteome to conserve sulfur when bioavailable sulfur is limited. It also suggests that for some organosulfur compounds the bacteria are unable to match S liberation rates with protein synthesis demands resulting in S stress even though growth is permitted on that substrate.

Modeling the Pseudomonas Sulfur Regulome
The Pseudomonas sulfur regulome was modeled as an Artificial Neural Network. In the ANN, gene expression of the significantly differentially expressed transcription factors, enzymes, and transmembrane transporters are predicted as a function of 27 Quantitative Structure-Activity Relationship (QSAR) features of sulfur nutrients. The ANN model accurately predicted gene expression as functions of sulfur media QSAR for transcription factor expression (Pearson's Correlation Coefficient 0.99), enzymes (0.98), and transmembrane transporters (0.96). The ANN was used to predict the transcriptome and biomass for Pseudomonas cultured on 25 additional sulfur media types for which growth data was available and for which the sulfur compound can be described using the set of selected 27 QSAR parameters. Modeled sulfur regulome predicted biomass of Pseudomonas cultured on different sulfur media with an average percent error of 23% (Minimum error 0.8%, maximum error 90%). Accuracy of biomass prediction is dependent upon the similarity of modeled sulfur compound to one of the 9 sulfur compounds used to build the ANN model.

Deconvolution of the Sulfur Regulome
While the ANN has the capacity to predict microbial behavior as a function of nutrient QSAR parameters, the structure of the underlying ANN provides insights into the mechanisms of the Pseudomonas sulfur regulome. The topology of the ANN has three distinct but interlinked subnetworks. Each subnetwork responds to unique QSAR parameters, is driven by the activity of one main transcription factor per subnetwork, and regulated clusters of genes that are specific in their composition of transporters and metabolic functions.

These results indicate that this modeling approach can be used to accurately predict the Pseudomonas sulfur regulome for a variety of conditions. The model can predict the regulome for novel sulfur source with the accuracy of the model for a particular nutrient dependent on the structural similarity of that nutrient to the nutrient used to train the model. This allows for the capacity to understand, predict, and engineer microbial systems to detect and respond to arbitrary environmental compounds.

The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357.Brookhaven National Laboratory is operated under Contract No. DE-AC02-98CH10886.This contribution originates from the "Environment Sensing and Response" Scientific Focus Area (SFA) program at Argonne National Laboratory. This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Genomic Science Program (GSP).

110. Effect of Nitrogen and Phosphorus Limitation on the Growth and Morphology of Engineered Microbial-Plant Communities.

Leland Cseke2, Shalaka Desai1, Sarah Zerbs1, Peter E. Larsen1, Geetika Trivedi2, Avinash Sreedasyam2, Peter Korajczyk1, and Frank R. Collart1 (fcollart@anl.gov)1

1Biosciences Division, Argonne National Laboratory, Lemont, IL
2Department of Biological Sciences, University of Alabama in Huntsville, Huntsville, AL

http://www.bio.anl.gov/molecular_and_systems_biology/protein_expression

Project Goals: The Argonne "Environment Sensing and Response" Scientific Focus Area (SFA) program seeks to identify the molecular basis of cellular transport and sensory pathways that mediate the response of terrestrial ecosystems to environmental nutrients. The mechanistic links between and within ecosystems comprised of plants, fungi, and soil bacteria involved in the production of biomass for fuel are currently very poorly defined. The effects of nutrient availability, closely linked to climate, on those mechanistic links, are also inadequately understood. This program will address this knowledge gap by mapping transport and sensor proteins to specific environmental compounds to define their function and biological roles and establish a series of defined connections between the environment and the cell. The knowledge will facilitate the development of system-level models predictive of cellular response to changes in environmental conditions.

Plant microbial communities play an important role in terrestrial ecosystem carbon (C) storage. These communities are typically comprised of symbiotic fungi and diverse species of bacteria that are modulators of plant response to the environment and change carbon budgeting in forest ecosystems. Carbon allocation (biomass) and partitioning (biomass quality) in plants are affected by nutrient availability in soils and are influence by the microbial community. Changes in plants C allocation are elicited by nutrient availability. To understand the molecular mechanism underlying these plant microbe interactions, we examined the molecular and phenotypic changes of aspen seedlings that were inoculated with mycorrhizal fungi and / or mycorrhizal helper bacteria, evaluated under either nitrogen or phosphorous nutrient stress conditions.

To determine the colonization-induced changes in plant signaling or metabolism that alter P / N use efficiency, aspen seedlings were inoculated with mycorrhizal fungi and / or mycorrhizal helper bacteria. The plant transcriptome was sequenced from the plant root samples colonized with and without the microbial community for five weeks in controlled conditions. Expression profiles of the plant microbial community demonstrated 2482 genes were significantly differentially regulated in response to bacterial colonization (Pseudomonas fluorescens SBW25) which contributes 6% of total number of genes. Plant transporters and transcription factors were highly up-regulated during bacterial colonization, whereas enzyme related genes were down-regulated during the interaction. Over 3000 plants genes were significantly differentially regulated due to mycorrhizal colonization by Laccaria bicolor.

We also examined the effect of nitrogen and phosphorous nutrient stress on Aspen phenotype after

microbial colonization. Aspen trees were subjected to low levels of nitrogen (2.94 mM, 1.47 mM, 0.98 mM and 0.49 mM) and phosphorus (0.236 mM, 0.1 mM, 0.075 mM and 0.041 mM) along with optimum nitrogen (14.7 mM) and phosphorus levels (1.25 mM). Experimental setup included appropriate controls consisting of the one- and two-component combinations of P. fluorescens Pf-01 and Laccaria with the aspen trees. Phenotypic data was collected that included fresh and dry weights, leaf number, and stem length along with root tissue scans for analysis of root branching.

Phenotypic analysis indicated there was a decrease in plant biomass with reducing levels of nitrogen, while significant increases in plant biomass were observed when colonized by either Laccaria alone or Laccaria in combination with P. fluorescens strain Pf-01. However, increase in biomass was most significant at optimum nutrient levels, 56% increase in biomass with Laccaria alone and 74% with combination of Laccaria and Pf-01. At 2.94 mM nitrogen there was a 10% increase in plant biomass with Laccaria alone and no effect on biomass when in combination with Pf-01. Below 2.94 mM nitrogen there was no significant increase in biomass; however, in each of these nitrogen-limiting conditions there was a slight increase in biomass when Laccaria was in combination with Pf-01. It is evident from the study that Laccaria alone or in combination with Pf-01 is not mobilizing nitrogen from growth medium with very low nitrogen levels (below 2.94 mM).

Under phosphorus limitation there was significant increases in plant biomass when colonized by either Laccaria alone or Laccaria in combination with Pf-01. An increase in biomass was most significant at 1.25 mM phosphorus (optimum level), 30% increase in biomass with Laccaria alone and 80% with combination of Laccaria and Pf-01. Also, at 0.236 mM and 0.1 mM phosphorus, there was around 10-20% increase in biomass with Laccaria alone or in combination with Pf-01. At lower phosphorus levels (0.236 mM, 0.1 mM and 0.075 mM) Pf-01 alone had negative impact on plant health, i.e. decrease in biomass. Overall, it appears that the Laccaria treatment alone or in combination with Pf-01 was alleviating phosphorus stress.

This knowledge will facilitate the development of more accurate systems-level models predictive of root response to environmental conditions or changes and potential for ecological niche enhancement by soil symbioses. The elucidation of function, regulation, and system response will support the Genomic Sciences Program goal of achieving a genome-based, dynamic systems-level understanding of organism and community function.

The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. Brookhaven National Laboratory is operated under Contract No. DE-AC02-98CH10886.This contribution originates from the "Environment Sensing and Response" Scientific Focus Area (SFA) program at Argonne National Laboratory. This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Genomic Science Program (GSP).

111. Methionine Importers in Soil Bacteria: Potential for Transporter-Component Crosstalk

Eliza L. Zielazinski, Sarah Zerbs, Frank R. Collart, and Philip D. Laible* (laible@anl.gov) Biosciences Division, Argonne National Laboratory, Lemont, IL

http://www.bio.anl.gov/molecular_and_systems_biology/protein_expression

Project Goals: The Argonne "Environment Sensing and Response" Scientific Focus Area (SFA) program seeks to identify the molecular basis of cellular transport and sensory pathways that mediate the response of terrestrial ecosystems to environmental nutrients. The mechanistic links between and within ecosystems comprised of plants, fungi, and soil bacteria involved in the production of biomass for fuel are currently very poorly defined. The effects of nutrient availability, closely linked to climate, on those mechanistic links, are also inadequately understood. This program will address this knowledge gap by mapping transport and sensor proteins to specific environmental compounds to define their function and biological roles and establish a series of defined connections between the environment and the cell. The knowledge will facilitate the development of system-level models predictive of cellular response to changes in environmental conditions.

A typical bacterial ABC transporter contains three protein components: a solute binding protein (SBP), a transmembrane domain (TMD) dimer in the inner membrane, and an ATPase dimer in the cytoplasm. In bacteria, roughly 2-5% of the genome codes for components of ABC type transporters (1). Certain types of soil bacteria contain an even higher percentage of ABC transporters, such as Pseudomonas flourescens, a species that contains well over 350 transport protein components [as identified by Transport DB (1)]. Interestingly though, approximately 30% of the ABC transporter proteins in P. fluorescens PF-5 are 'orphans' not genomically co-located with other transporter component genes and thus difficult to assign to a specific complex. The abundance of ABC transporter component genes and the presence of orphan genes suggest the potential for crosstalk amongst members. While studies have indicated that TMDs can interact with multiple SBPs with varying affinity and specificity (2), there is limited information regarding the specificity of TMD and ATPase protein-protein interactions. The promiscuity of TMD-ATPase interactions may also affect promiscuity of SBP interactions with the transporter complex and refine ligand specificity of the overall transporter. To address this knowledge gap, a family of methionine transporter complexes from four Pseudomonas fluorescens strains was investigated and the potential for crosstalk between components was assessed.

Genomic analyses of four Pseudomonas fluorescens strains (PF-5, Pf0-1, SBW25, and WH6) have identified a set of ten ATP-binding cassette (ABC)-type amino acid transporters with high sequence similarity to the structurally characterized methionine importer MetNI from Escherichia coli. Recombinant, dual-vector expression strategies yield intact complexes of the transmembrane domain (TMD) and ATPase components (Fig. 1) of the P. fluorescens gene targets in E. coli. Expression experiments systematically varying the combination of ATPase genes with one TMD gene revealed promiscuity of certain ATPases, where stable "hybrid" complexes could be formed with both a TMD from the same strain as the ATPase or with a TMD from a different P. fluorescens strain. This type of functional "crosstalk" between ATPases and TMDs could play a role in rapid nutrient exchange between Pseudomonas soil bacteria and other rhizosphere inhabitants. The ability of different ATPases to recognize the same binding site of a TMD also provides an opportunity to study the molecular basis of recognition between the domains of the transporter core.

It has been established for MetNI that ligand binding of methionine to the ATPase domain causes transinhibition, where transport is suppressed in a concentration-dependent manner by the substrate. Few ABC transporters with transinhibition regulatory features have been characterized so it is unkown if the regulatory small molecule always matches the transported small molecule. Utilizing ATPase activity inhibition assays, it is possible to identify ligands that have an inhibitory effect on transporter complex function. We determined probable transporter substrates by testing the specificity of associated MetQ-family SBPs using a fluorescence-based thermal shift assay. The ATPase inhibitors were compared with the SBPs ligand-binding profiles to determine if there were patterns between small molecules transported from the environment and the internal regulators of transport activity. Cross-talk at both the external and internal membrane interfaces may have a role in expanding transporter capacity while retaining specificity. Association of non-native ATPases with a TMD could also enable activity regulation by alternative regulatory substrates. Initial comparison of the results for the activity inhibition assay and the thermal shift assay for a set of methionine-derivative ligands for native and non-native transporter complexes support these hypotheses. These results demonstrate how multiple complete transporters and orphan components may function together to afford functional advantages to bacteria in complex and highly competitive environments such as the plant rhizosphere.

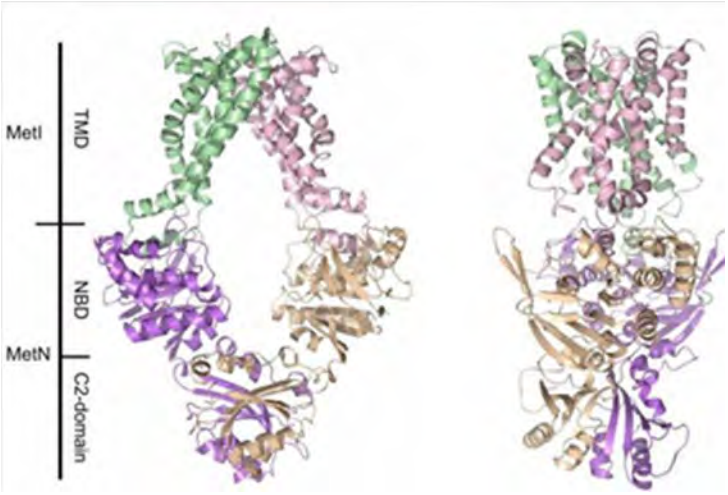


Fig. 1. **Structure of MetNI.** Representative bacterial ABC importer (*3*).

References

1. Ren, Q. H., Chen, K. X., and Paulsen, I. T. (2007) TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels, Nucleic Acids Res 35, D274-D279.
2. Thomas, G. H. (2010) Homes for the orphans: utilization of multiple substrate-binding proteins by ABC transporters. Mol. Microbiol. 75, 6-9.
3. Kadaba, N. S., Kaiser, J. T., Johnson, E., Lee, A., and Rees, D. C. (2008) The high-affinity E-coli methionine ABC transporter: Structure and allosteric regulation, Science 321, 250-253.
4. Ferreira, M. J., and de Sa-Nogueira, I. (2010) A Multitask ATPase Serving Different ABC-Type Sugar Importers in Bacillus subtilis, J Bacteriol 192, 5312-5318.

This abstract has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. This contribution originates from the "Environment Sensing and Response" Scientific Focus Area (SFA) program at Argonne National Laboratory. This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Genomic Science Program (GSP).

112. Genetic and molecular anayses of the regulation of flowering in the temperate grass, Brachypodium distachyon

Daniel P. Woods1,3,4 and Richard Amasino1,3

1Department of Biochemistry, University of Wisconsin, Madison, WI USA 3Great Lakes Bioenergy Research Center, United States Department of Energy 4Laboratory of Genetics, University of Wisconsin, 425-G Henry Mall

amasino@biochem.wisc.edu

Abstract:Proper timing of flowering is essential for reproductive success in plants and is a major determinant of biomass yield. A key adaptation to seasonal variation in temperate climates is the evolution of a vernalization requirement. Vernalization is the process by which competence to flower is achieved only after prolonged exposure to winter cold. A vernalization requirement is found in a range of flowering plant families, yet relatively little is known about the molecular nature of this phenomenon outside the eudicot model Arabidopsis thaliana. We are using the small temperate grass, Brachypodium distachyon as a model to study flowering in the grasses. Here we report the characterization of the vernalization response across diverse Brachypodium accessions, and have begun to assess the genetic basis of the variation in flowering in populations derived from crosses between accessions that vary in their vernalization requirement. There is considerable variation in the vernalization and flowering time responses across the Brachypodium accessions tested, and natural variation studies suggest a few genes with large effects control flowering in Brachypodium. Additionally, we have screened for vernalization-insensitive and rapid- flowering mutants. We have mapped several mutants within our collection using whole genome sequencing followed by the use of the bioinformatics mapping pipeline, CLOUDmap. Several novel flowering time genes have been identified, many of which appear to be negative upstream regulators of the key vernalization gene in grasses, VERNALIZATION1. Relative to the large genomes, long generation times, and domesticated backgrounds of cultivated cereals, Brachypodium offers a powerful opportunity to understand flowering time at a deeper molecular level in grasses.

113. Modulating plant cell wall for biofuel at GLBRC

Sang-Jin Kim1, Starla Zemelis1, Kenneth Keegstra1,2, and Federica Brandizzi1,2

1Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI 48824;
2MSU-DOE Plant Research Lab, Michigan State University, East Lansing, MI 48824.

https://www.glbrc.org/research/plants

Projects goals : Great Lakes Bioenergy Research Center (GLBRC) is focused on sustainable production of crops with desirable biofuel traits and efficient conversion of biomass into fuels and chemicals. To achieve the goals, GLBRC has four research areas – Plants, Sustanability, Deconstruction and Conversion. The ultimate goal of plants area is to increase both the amount and quality of biomass by modulating cell wall digestibility through lignin modification and the metabolic and genetic system that control accumulation of oils and other easily digestible, energy-rich compounds in plant tissues.

To increase easily digestible hemicellulose (mixed-linkage glucan; MLG) in plant stem parenchyma cells, a Plants team in GLBRC is focused on understanding the factors required to accumulate large quantities of MLG including genetic regulation of MLG synthesis, identification of tissue specific promoters and characterization and engineering of MLG synthases.

To achieve the goals, we are using Brachypodium distachyon as a model grass species. MLG is one of the major components of cereal grains, and occupies up to 80% of cell walls of the Brachypodium endosperm. The MLG biosynthesis depends on the biochemical activity of membrane spanning glucan synthases encoded by the CSLH and CSLF cellulose synthase-like gene families. However, relatively little is known about their topology with respect to the biosynthetic membranes and requirement for producing two different linkages in the MLG glucan backbone. As the first step on the project, we have demonstrated the topology of CSLF6 protein derived from Brachypodium (BdCSLF6) using heterologous expression systems. Using live cell imaging and immuno-electron microscopy analyses of tobacco epidermal cells expressing BdCSLF6, we demonstrate 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. By implementing protease protection assays of BdCSLF6 expressed in the yeast Pichia pastoris, we also demonstrate that the catalytic domain, the N-terminus and the C- terminus of the protein are exposed in the cytosol. Furthermore, we found that BdCSLF6 is capable of producing MLG not only in tobacco cells but also in Pichia, which generally does not produce MLG. Together, these results support the conclusion that BdCSLF6 can produce both of the linkages present in the (1,3; 1,4)-β-D-glucan chain of MLG and that the product is channeled at the Golgi into the secretory pathway for deposition into the cell wall.

This initial approach allowed us to set up the condition for analyzing MLG synthases. To accomplish our goals, we are currently engineering BdCSLF6 to improve its ability to produce large quantity of MLG.

The Great Lakes Bioenergy Research Center is supported by the Office of Biological and Environmental Research in the DOE Office of Science (DOE BER Office of Science DE-FC02- 07ER64494).

114. Using Natural Variation in Switchgrass to Identify Genes Relevant for Biomass Production and Improved Breeding Strategies

Joseph Evans1, Megan Bowman1, Jeongwoon Kim1, Kevin Childs1, Emily Crisovan1, Brieanne Vaillancourt1, Aruna Nandety2, Shawn Kaeppler, Michael Casler4, and C. Robin Buell1* (buell@msu.edu)

1Michigan State University, Dept. of Plant Biology, East Lansing MI; 2University of Oklahoma, Department of Plant Biology and Microbiology, Norman, OK; 3University of Wisconsin, Dept. of Agronomy, Madison, WI; 4USDA---ARS, U.S. Dairy Forage Research Center, 1925 Linden Dr., Madison, WI

https://www.glbrc.org

Project Goals

Switchgrass (Panicum virgatum L.) is a polyploid, perennial grass species that is native to North America, and is being developed as a biofuel feedstock crop. Switchgrass is present primarily in two ecotypes: a northern upland ecotype composed of mixed tetra--- and octoploid switchgrasses, and a southern lowland ecotype composed of primarily tetraploid switchgrasses. While the ecotype designations are relatively well understood due to previous work using genotyping---by---sequencing and chloroplast simple sequence repeat markers, the development of switchgrass as a biofuel feedstock crop will require more detailed genetic information that can drive improvement efforts including accelerated and targeted breeding efforts. We developed a custom repeat library to mask repetitive sequences in the genome, and then utilized a high---coverage exome capture sequencing approach with the switchgrass v.1.1 genome to sequence a diversity panel composed of 537 individuals spanning 45 upland and 21 lowland switchgrass populations. From these data, we identified over 12 million polymorphic loci, from which, we selected a highly filtered and robust subset of 1.5 million high fidelity loci. From these polymorphisms, we were able to identify distinct population groups within the upland and lowland ecotypes, a result that was further supported through genetic distance analysis. We were able to also identify several large effect SNPs in flowering time genes that appeared to be ecotype specific, which may help contribute to the large phenotypic differences between ecotypes. We also identified 16,971 up---copy number variants, 108,475 down---copy number variants, and 14,178 presence---absence variants. Using these copy number variants in a genetic distance analysis, we were able to confirm our earlier population clustering results, which provides evidence that there may be large groups of both ecotype and population specific copy number variants. These polymorphisms are being incorporated into genome wide association studies and genome selection experiments to identify individuals with improved biomass yield and biofuel feedstock traits.

This work was funded by the Department of Energy Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494).

115. The Great Lakes Bioenergy Center Material Production Chain: Integrated Production and Characterization of Feedstocks, Pretreated Biomass, and Hydrolysate as Standardized Reagents for Center-Wide Research

David M. Cavalier (cavalie8@msu.edu)1, David Benton2, Ying Gao2, Rebecca G. Ong1, Ruwan Ranatunga2, Yaoping Zhang2

1 Michigan State University, East Lansing, Michigan; 2 University of Wisconsin, Madison, Wisconsin

https://www.glbrc.org

The goal of the Great Lakes Bioenergy (GLBRC) Material Production Chain (MPC) is to produce consistent high quality feedstocks, Ammonia Fiber Expansion (AFEXTm)-treated biomass, and AFEX-treated corn stover and switchgrass hydrolysates as reagents for center-wide research.

A major challenge in conducting research in a large center is providing researchers with significant amounts of standardized research materials. The GLBRC MPC is a highly integrated project with researchers and support staff from Plants, Deconstruction, Conversion, Sustainability, Core Facilities, and Scientific Computing Areas working together to design and implement standard operating procedures (SOPs) and quality control (QC) measures to produce standardized materials for center---wide research. These materials include a variety of plant feedstocks (corn stover, mixed---prairie, miscanthus, poplar, sorghum, and switchgrass), ammonia fiber expansion (AFEXTM) treated biomass, and hydrolysates generated from AFEX--- treated corn stover and switchgrass. All materials are produced by adhering to SOPs and tested to ensure they fall within the boundaries defined by statistically robust QC measures. A custom laboratory information system (LIMS) was developed to catalogue and track materials, their relevant data and experimental details as they move through MPC, and to allow researchers to seamlessly order materials. Here we present the main features of the MPC and provide examples data collected at each step.

This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494).

116. Structural studies of β-aryl ether lignin degradation

Kate E. Helmich1,2* (helmich@wisc.edu), Jose Henrique Pereira,3,4 Richard A. Heins,3 Daniel Gall,1,2 Ryan P. McAndrew,3,4 Craig Bingman,1 Timothy J. Donohue,1,2 Daniel R. Noguera,1,2 Blake A. Simmons,3,5 Kenneth L. Sale,3,5 George N. Phillips Jr.,6 John Ralph,1,2 Paul D. Adams3,4

1University of Wisconsin, Madison, Wisconsin; 2Great Lakes Bioenergy Research Center; 3Joint BioEnergy Institute; 4Lawrence Berkeley National Laboratory, Berkeley, California; 5Sandia National Laboratories, Livermore, California; and 6Rice University, Houston, Texas

https://www.glbrc.org
https://www.jbei.org

Project Goals: To use biochemical and structural methods to characterize the lignin β-aryl ether degradation pathway from Sphingobium sp. SYK-6. The information gained can be applied to enhance the degradation of lignin in the processing of second-generation biofuels and to develop valuable aromatics and other renewable chemicals from lignin streams.

Lignin is both a major obstacle to the sugar release from lignocellulosic biomass for the production of second-generation biofuels and a potential source of aromatics and other valuable chemicals. Degradation of lignin has been relatively well characterized in fungi, and is becoming better understood in bacteria. A catabolic pathway for the enzymatic cleavage of β-aryl ether linkages, which account for 50–70 % of all inter-unit linkages in lignin, has been previously identified in the bacterium Sphingobium sp. SYK-6. Here we present a structural characterization of the two β-etherase enzymes in this degradation pathway, the glutathione S-transferase (GST) enzymes LigE and LigF. LigE and LigF catalyze the regiospecific, glutathione-dependent cleavage of the β-ether bond in (βR)- or (βS)-(3'-methoxyphenoxy)-γ-hydroxypropiovanillone (MPHPV) to release guaiacol and (βS)- or (βR)- glutathionyl adduct of γ-hydroxypropiovanillone, respectively. Although both structures follow the canonical GST fold architecture, an N-terminal thioredoxin- like binding domain and a C-terminal all α-helical domain, each forms a unique dimer interface resulting in distinctly different substrate binding pockets on opposite sides of the active site channel. Understanding the mechanism of β-aryl ether cleavage has great potential for the effective breakdown of lignin in biofuels processing. This new information can enhance our ability to efficiently degrade lignin and to enhance its use as a source of valuable aromatics and other renewable aromatic chemicals.

The Great Lakes Bioenergy Research Center is a U.S. Department of Energy Bioenergy Center supported by the DOE Basic Energy Research Office of Science (DE-FC02-07ER64494); the Joint BioEnergy Institute is supported by the DOE Office of Science, Office of Biological and Environmental Research (DE-AC02-05CH11231).

117. Diversity and Stability in Experimental Fields of Perennial Bioenery Crops

Karen A. Stahlheber1,2* (nikolaka@msu.edu), Randall D. Jackson2,3, and Katherine L. Gross1,2

1W. K. Kellogg Biological Station, Michigan State University, 2Great Lakes Bioenergy Research Center, Madison, WI, 3University of Wisconsin, Madison

https://www.glbrc.org/research/sustainability

In this project our goal was to determine the effect of species richness on productivity and its change through time in perennial cropping systems designed for bioenergy production. For growers of biofuel feedstocks, yield stability may be an important secondary ecosystem service of diverse plantings, especially in the context of climate change and inter-annual variability in growing conditions. Experiments in community ecology have identified several mechanisms linking higher diversity to increased stability, yet these have not been tested in agro-ecosystems, which differ in many important ways from the communities used in these experiments. The strength of a relationship between plant diversity and stability may also vary with environmental context, thus comparing multiple sites is critical. By testing the importance of ecological mechanisms for achieving stable yields at two replicate sites, we will generate general recommendations for feedstock production.

Evidence from both community ecology and studies of candidate bioenergy crops suggests that diverse plantings may not outperform the most productive monocultures in terms of yield. Mixtures of many different species, however, may offer benefits for long-term sustainability that may offset smaller yields. One such ecosystem service is stability, or the consistency of production over time. The presence of multiple species intrinsically reduces temporal variance in the same way that grouping stocks together in funds reduces risk to investors; a phenomenon termed the portfolio effect in ecology. In addition to this statistical mechanism, however, the biological interactions between species can enhance stability. For example, different species (or functional groups) may respond in alternate ways to environmental change such that one may dominate in some conditions, and another in others. This niche complementarity between species increases stability. Lastly, in mixtures that are highly dominated by one species, the response of this species can determine the pattern over time. By examining biofuel plantings for evidence of these mechanisms, we aim to understand where and when diversity is stabilizing.

This project is a part of the Great Lakes Bioenergy Research Center's biofuel cropping systems experiment (BCSE), which has two replicate sites at Kellogg Biological Station (Michigan) and Arlington Agricultural Research Station (Wisconsin). In this experiment, eight managed agroecoystems were established for candidate crops whose biomass may be used as feedstock for bioeful production. Several of these candidate cropping systems included perennial plants that are native to the Midwest region of the USA. Following establishment, small plots in these fields have been sub-sampled in July to determine ANPP (annual net primary productivity) and plant composition. These data are used for our analysis of temporal change and diversity.

In this poster we present results from the first five years of biomass production. Over this period, switchgrass was the most productive crop, although this was more dramatic at the Wisconsin site compared to Michigan. Each perennial cropping system developed a distinct plant community composition which differed from all others within a site, and across sites. We also observed significant change in the plant communities over time as the abundances of different species shifted. During the initial five years, the most diverse cropping system in terms of species richness (a mix of native prairie species) was not necessarily the most stable overall. Species richness, however, did have a weakly positive effect on stability at the Michigan site. By contrast, negative covariance between functional

groups (a measure of niche complementarity) always increased stability. Plots with a dominant species that produced a consistent amount of biomass year after year were also very stable overall. This suggests that including multiple functional groups and a stable dominant will be important for achieving stability in yield.

Funding Soure: This work was funded with support from the USA Department of Energy's Great Lakes Bioenergy Research Center.

118. Candidate Gene Discovery and Validation by Analysis of Natural Variation for Cell Wall Compositional Traits in Maize

Marlies Heckwolf1*(mheckwolf@wisc.edu), German Muttoni2,3, Nicholas Santoro4, Shane Cantu4, Candice Hirsch5, Brieanne Vaillancourt4,6, C. Robin Buell4,6, Natalia de Leon1,2 and Shawn Kaeppler1,2

1DOE Great Lakes Bioenergy Research Center, Madison, WI, 2University of Wisconsin- Madison, WI, 3Current address: Monsanto Company, Lebanon, IN, 4DOE Great Lakes Bioenergy Research Center, East Lansing, MI, 5University of Minnesota, MN, 6Michigan State University, MI

Goals: The main goal of this project is to utilize naturally occurring genetic variation in maize as a gene discovery tool to identify novel genes that are associated with cell wall compositional traits and therefore potentially reduce plant biomass recalcitrance.

Maize is a relevant source of biomass for biofuel production in the near term, and also an important model for other biofuel feedstock grass species. Our project exploits naturally occurring phenotypic variation for non-grain biomass traits in maize coupled with extensive genetic information to identify novel genomic regions that can reduce biomass recalcitrance. We have evaluated cell wall bound glucose and pentose release from cores taken from the second lowermost stalk internode of 563 diverse maize inbred lines using an automated digestibility platform. These inbred lines were evaluated in 2010, 2011 and 2012 in field experiments with a randomized complete block design and two replications per year. Three representative plants per genotype and field replication were sampled each year.

We observed significant genetic variation for both traits ranging from 4.1 to 20.2 % of sugar per mg dry biomass for glucose and 1.2 to 9.02 for pentose release across genotypes. The genetic basis of these traits was analyzed using a set of 438,222 single nucleotide polymorphisms in a genome wide association study. We identified natural alleles of different genes that are promising candidates to alter biomass digestibility. Based on their biological relevance and expression profiles, genes were prioritized for validation. These candidates are annotated as having a role in xylan biosynthesis and nuclear processes that relate to cell wall biosynthesis. Our current focus is on using transgenic plants and transposon alleles to further explore the potential of these genes in reducing biomass recalcitrance.

This work was funded by the Department of Energy Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494).

119. Chemical genomic analysis reveals feedstock specific differences in the microbial response to lignocellulosic hydrolysates

Jeff Piotrowski1 (jpiotrowski@wisc.edu), Li Hinchman1, Quinn Dickinson1, Scott Bottoms1, Rebecca Garlock Ong2,3, Yaoping Zhang1, Robert Landick1

1-Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI; 2-Michigan State University, East Lansing, MI; 3-Michigan Technological University, Houghton, MI

https://www.glbrc.org

Project Goals
Lignocellulosic derived fermentation inhibitors represent a challenge to biofuel production. We developed a method to biologically assess the quality and inhibitor landscape of hydrolysates using chemical genomics. We are using this approach to understand the different biological responses to hydrolysates produced using different methods and feedstocks. Herein, we specifically investigated the biological response of Saccharomyces cerevisiae to hydrolysates produced from corn stover and switchgrass and discovered a unique chemical genomic profile of each hydrolysates. We have extended this analysis to other systems using E. coli and Z. mobilis chemical genomics to provide a rapid method of assessing the biological response to hydrolysates that yield rich functional insight into fermentation inhibitors, and can inform hydrolysate production strategies.

Abstract text
Lignocellulosic biomass derived fuels and chemicals provide a suite of sustainable bioproducts. Before biomass can be converted to fuel or compounds through biological conversion, it must be converted to fermentable sugars via pre-treatment and hydrolysis, and these sugars converted to fuels by microorganisms. Both pre-treatment and hydrolysis can imbue the resultant hydrolysates with toxicity arising from residual pre-treatment chemicals or biomass derived inhibitors1, which throttle fermentation rates at a substantial economic cost2. Rapid assessment of hydrolysate quality is a key requirement not only to assess the content of fermentable sugars and inhibitors, but also the variation that can arise during production. Analytical chemistry methods can provide detailed compositional analysis of the substrate, but given the complex nature of the plant biomass and chemistry occurring during pre-treatment, it is impossible to quantify all compounds present in the hydrolysate, and ultimately the effects of the composition on the microbial biocatalysts must be inferred. What is needed is an “analytical biology” method to assess the quality of starting substrates, which can provide functional information of the microbial response to a medium and inform process improvements.

Chemical genomics is a reverse-genetics approach that uses genome-wide mutant collections to gain functional insight into the modes-of-actions and cellular targets of bioactive compounds3. Pooled mutant collections are grown in the presence of a bioactive compound, and the individual fitness of these mutants can be assessed by sequencing of their molecular barcodes. Sensitive mutants give clues into the mode-of-action of fermentation inhibitors, and resistant mutants can provide points of rational engineering of tolerance.

We applied chemical genomics to investigate the biological response to hydrolysates produced from either corn stover or switchgrass, which show different fermentation profiles. We challenged the yeast deletion collection with a variety of hydrolysates produced from either ammonia fiber expansion (AFEX) treated switchgrass or corn stover. While variation between batches was low, indicating consistency in production, obvious differences were apparent between hydrolysates made from the different feedstocks.

Deletion mutants in amino acid related processes were more resistant in corn stover produced hydrolysate compared to switchgrass, indicating potential amino acid deficiencies in switchgrass hydrolysates. We also found that deletion mutants in genes involved in ergosterol biosynthesis (ERG6, ERG3) were sensitive to switchgrass hydrolysates. Mutations in these genes affect membrane composition and confer sensitivity to acetic acid. We found that the acetate content of switchgrass hydrolysates was significantly higher ($p<0.001$) than corn stover. These data suggest that the slower fermentation rates found in switchgrass hydrolysates could arise from amino deficiencies and acetic acid stress.

Chemical genomic profiling of hydrolysates provides detailed information on the biological response of fermentative microbes to lignocellulosic hydrolysates. Using similar mutant collections in E. coli4 and Z. mobilis5, we have further extended this method to diverse industrially relevant microbes. We are presently exploring the effects of more diverse feedstocks and interannual variability on the biological response of fermentative microbes.

References

1. Piotrowski, J. S. et al. Death by a thousand cuts: the challenges and diverse landscape of lignocellulosic hydrolysate inhibitors. Front. Microbiol. 5, (2014).
2. Keating, D. H. et al. Aromatic inhibitors derived from ammonia-pretreated lignocellulose hinder bacterial ethanologenesis by activating regulatory circuits controlling inhibitor efflux and detoxification. Microb. Physiol. Metab. 5, 402 (2014).
3. Ho, C. H. et al. Combining functional genomics and chemical biology to identify targets of bioactive compounds. Curr. Opin. Chem. Biol. 15, 66–78 (2011).
4. Otsuka, Y. et al. GenoBase: comprehensive resource database of Escherichia coli K-12. Nucleic Acids Res. 43, D606–D617 (2015).
5. Skerker, J. M. et al. Dissecting a complex chemical stress: chemogenomic profiling of plant hydrolysates. Mol. Syst. Biol. 9, 674 (2013).

This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494).

120. Introducing readily cleavable bonds into the lignin backbone: The Zip-Lignin™ Strategy

Steven D. Karlen1,2,* (skarlen@wisc.edu), Fachuang Lu,1,2 Dharshana Padmakshan,1,2 Matt Regner,1,2 Timothy Pearson,1,2 Yimin Zhu,1,2 Jorge Rencoret,1,2 Rebecca A. Smith,1,2 Saunia Withers,1,3 Emily Frankman,1,3 Ji-Young Park,1,4 Eliana Gonzales-Vigil,1,4 Debora L. Petrik,1,5 Cynthia L. Cass,1,5 Cliff Foster,1,3 Nicholas Santoro,1,3 Heather Free,6 Bronwen Smith,6 Philip Harris,6 John H. Grabber,7 John Sedbrook,1,5 Curtis G. Wilkerson,1,3 Shawn D. Mansfield,1,4 and John Ralph1,2

1 Great Lakes Bioenergy Research Center (GLBRC); 2 University of Wisconsin, Madison, Wisconsin; 3 Michigan State University, East Lansing, Michigan; 4 University of British Columbia, Canada; 5 Illinois State University, Normal Illinois; 6 University of Auckland, New Zealand; 7 US Dairy Forage Research Center, USDA-ARS, Madison, Wisconsin

https://www.glbrc.org

Project Goals: To provide plants with the necessary pathways to produce monolignol ferulate conjugates and export them to the cell wall where they introduce readily cleavable ester linkages (‘zips’) into the lignin polymer backbone in ways that significantly improve biomass processing energetics.

Lignin is a complex and irregular phenolic polymer that fortifies plant cell walls, and is responsible for the majority of the recalcitrance to industrial processing of plant biomass. The biosynthesis of lignin has been found to tolerate a large variety of perturbations to the lignin monomer pool. We have utilized this plasticity to incorporate ester linkages into the lignin polymer backbone by introducing a FERULOYL-CoA MONOLIGNOL TRANSFERASE (FMT) gene fused to a xylem-specific promoter into poplar trees. The FMT enzyme produces monolignol ferulate conjugates that integrally incorporate into the lignin polymer backbone. Under mild alkaline pretreatment conditions the ester bonds readily cleave resulting in an improved digestibility for the FMT-poplar over wild-type poplar. With improved methods recently developed, we have found that many plants seem to naturally incorporate low levels of these monolignol ferulate conjugates. In a screening of over 30 plant species, mainly cash crops, we found evidence for the conjugates in grain crops (corn, sorghum, rice, wheat) as well as some hardwoods (poplar, aspen, eucalyptus, balsa), but not in softwoods (cedar or spruce). The finding of such naturally occurring zip-monomers opens up new opportunities for engineering or breeding crop plants with cell walls that are better designed for deconstruction.

Publication:
“Monolignol ferulate transferase introduces chemically labile linkages into the lignin backbone” Wilkerson, C. G.; Mansfield, S. D.; Lu, Fachuang; Withers, S.; Park, J.; Karlen, S. D.; Gonzales- Vigil, E.; Padmakshan, D.; Unda, F.; Rencoret, J; Ralph, J Science, 2014, 344, 90−93.

The Great Lakes Bioenergy Research Center is a U.S. Department of Energy Bioenergy Center supported by the DOE Basic Energy Research Office of Science (DE-FC02-07ER64494).

121. Laccases Direct Lignification in the Discrete Secondary Cell Wall Domains of Protoxylem

Mathias Schuetz,1 Anika Benske,1,2 Rebecca A. Smith1,2* (rasmith29@wisc.edu), Yoichiro Watanabe,1 Yuki Tobimatsu,3,4 John Ralph,3 Taku Demura,5 Brian Ellis,2 A. Lacey Samuels1

1Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4
2Michael Smith Laboratories University of British Columbia Vancouver, British Columbia, Canada V6T 1Z4
3Department of Biochemistry and United States Department of Energy Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, University of Wisconsin, Madison, Wisconsin, USA 53726
4Graduate School of Agriculture, Kyoto University, Kyoto 606---8502, Japan 5Nara Institute of Science and Technology, Nara 630---0192, Japan

https://www.glbrc.org

Project Goals: To determine how the specificity of localized lignin deposition and cell wall cross---linking are dictated by the cell.

Plants precisely control lignin deposition in spiral or annular secondary cell wall domains during protoxylem tracheary element (TE) development. Because protoxylem TEs function to transport water within rapidly elongating tissues, it is important that lignin deposition is restricted to the secondary cell walls in order to preserve the plasticity of adjacent primary wall domains. The Arabidopsis (Arabidopsis thaliana) inducible VASCULAR NAC DOMAIN7 (VND7) protoxylem TE differentiation system permits the use of mutant backgrounds, fluorescent protein tagging, and high---resolution live---cell imaging of xylem cells during secondary cell wall development. Enzymes synthesizing monolignols, as well as putative monolignol transporters, showed a uniform distribution during protoxylem TE differentiation. In contrast, the oxidative enzymes LACCASE4 (LAC4) and LAC17 were spatially localized to secondary cell walls throughout protoxylem TE differentiation. These data support the hypothesis that precise delivery of oxidative enzymes determines the pattern of cell wall lignification. This view was supported by lac4lac17 mutant analysis demonstrating that laccases are necessary for protoxylem TE lignification. Overexpression studies showed that laccases are sufficient to catalyze ectopic lignin polymerization in primary cell walls when exogenous monolignols are supplied. Our data support a model of protoxylem TE lignification in which monolignols are highly mobile once exported to the cell wall, and in which precise targeting of laccases to secondary cell wall domains directs lignin deposition.

Publication
“Laccases Direct Lignification in the Discrete Secondary Cell Wall Domains of Protoxylem” Schuetz, M.; Benske, A.; Smith, R. A.; Watanabe, Y.; Tobimatsu, Y.; Ralph, J.; Demura, T.; Ellis, B.; Samuels, A. L. Plant Physiology, 2014, 166, 798-807.

This research was funded by the U.S. Department of Energy Great Lakes Bioenergy Research Center (grant no. BER DE-FC02-07ER64494).

122. Changes in biophysical climate regulation services from converting native grassland to bioenergy production in the US Midwest

Xuesong Zhang1,2*, Kaiguang Zhao1,2, Michael Abraha2,3, Ilya Gelfand2,3, Allison M. Thomson1,2, Steve K. Hamilton2,3, Jiquan Chen2,5, Roberto C. Izaurralde2,4, G. Philip Robertson2,3

1 Joint Global Change Research Institute, Pacific Northwest National Laboratory & University of Maryland, College Park Maryland 20740, USA; 2 Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, Michigan 48824, USA; 3 W.K. Kellogg Biological Station, Michigan State University, Hickory Corners, Michigan 49060, USA; 4 Department of Geographical Sciences, University of Maryland, College Park, Maryland 20740, USA; 5 Department of Geography, Michigan State University, East Lansing, Michigan 48824, USA

Principal Investigator: G. Philip Robertson

https://www.glbrc.org

Project Goals: Promoting our understanding and quantification of changes in biophysical climate regulation services as a result of converting native grassland to bioenergy production in the US Midwest. Abstract: Converting unmanaged grassland to managed bioenergy crops not only alters biogeochemical cycles, but also modifies surface biophysics, such as albedo, surface roughness, rooting depth, stomatal conductance, and leaf area. These biophysical perturbations subsequently change radiation budget at land surface and land-atmosphere exchange in sensible heat, evapotranspiration (ET), and momentum fluxes, which consequently influence atmospheric temperature, moisture, and circulation patterns. Climate regulation services from these biophysical effects are comparable to biogeochemical benefits of bioenergy production, with particular significance at the local and regional scales. Such biophysical effects, however, are often neglected in current climate policies; there is growing recognition that they should be additionally included to complement GHG values of land-based mitigation activities.

In this research, we combine in situ field measurements and remote sensing observations to improve our understanding of changes in biophysical climate regulation services from converting grassland to perennial bioenergy crops. In the US Midwest, albedo change as a result of cultivating native grassland for cellulosic bioenergy feedstocks could enhance the net greenhouse gases (GHGs) mitigation benefit of cellulosic bioenergy production (116.5 MgCO2 ha-1) by 20% over a time horizon of 50 years. With an integrated climate-agroecosystem model, parameterized with in situ and remote sensing data, we further demonstrate that considering interactions between agroecosystem processes and atmospheric circulation could result in noticeable difference in simulated regional climate (e.g. precipitation, temperature, and radiation budget), highlighting the importance of additionally including biophysical climate services in evaluating land-based mitigation activities, such as bioenergy production.

Funding statement:
This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494, DOE BER Office of Science KP1601050, DOE EERE OBP 20469-19145).

123. Nitrous oxide emissions from Midwest row-crops: Comparing presence and absence of winter cover crops

Neville Millar1,2, Dean G. Baas3, and G. Philip Robertson (robert30@msu.edu)1,2*

1W.K. Kellogg Biological Station, Michigan State University, Hickory Corners, MI; 2Dept. of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI; 3Michigan State University Extension, Michigan State University, Centreville, MI.

https://www.glbrc.org/research/sustainability http://cris.nifa.usda.gov/cgi-bin/starfinder/28112/crisassist.txt

Project Goals: Quantify multi-year emissions of the potent greenhouse gas (GHG) nitrous oxide (N2O) from cover crop systems under varying management to 1) improve the accuracy of inventories of crop based agricultural GHG emissions, and 2) evaluate the potential of cover crops as a GHG mitigation strategy.

Nitrous oxide (N2O) is the largest contributor to the greenhouse gas burden of cropping systems in the US, with emissions primarily due to N fertilizer inputs and other soil management activities. The practice of including winter cover crops in corn-based row-crop systems is not widely adopted but is increasing. The beneficial impacts of cover crops, that include reduced soil erosion, increased SOM, and weed suppression, are well known. However, very few studies have investigated the effect of including cover crops in row-crop rotations on N2O emissions.

Here, we will present three years of N2O emissions data from a corn-soybean-winter wheat rotation, either with or without cover crops, situated at the Kellogg Biological Station in SW Michigan. Results will explore the effects of cover crop presence or absence, cover crop type (grass, legume, Brassica), and cover crop termination time (early vs. late) on emissions of N2O.

Over the three years, results show that: 1) including cover crops did not have much affect on N2O emissions compared to no cover crops (Fig 1); 2) cover crop type and cover crop termination time had no overall affect on total N2O emissions (Fig 2); 3) N2O emissions were highest in corn (vs. wheat and soybean) in all years (Fig 3); and, 4) highest fluxes occurred following N fertilization, soil disturbance, and rainfall (soil wetting).

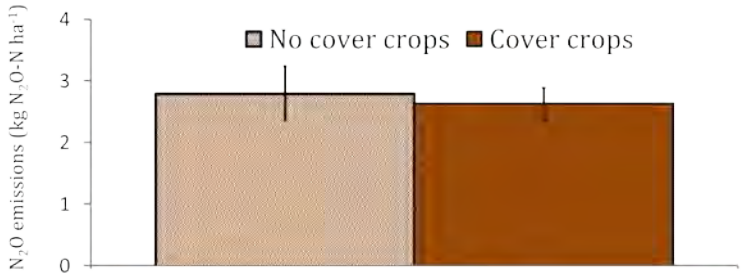


Fig 1. Average total N2O emissions from all treatments with and without cover crops (2012 – 2014).

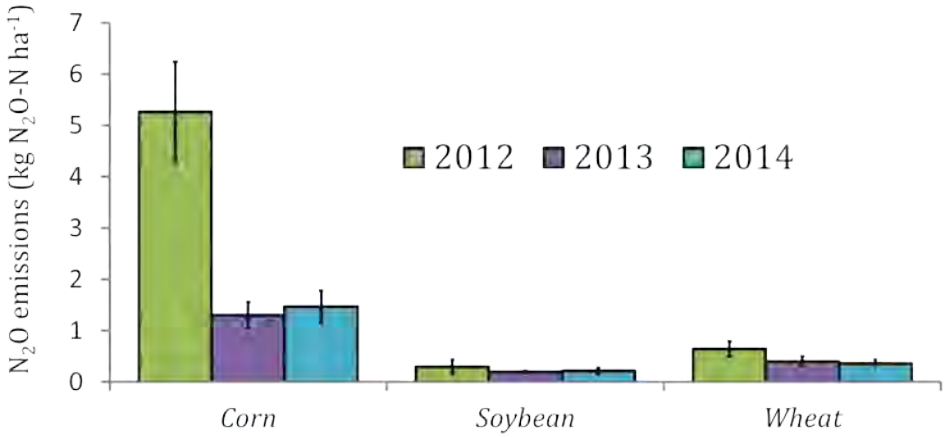


Fig 2. Average total N2O emissions from cash crops with cover crops (2012 – 2014).

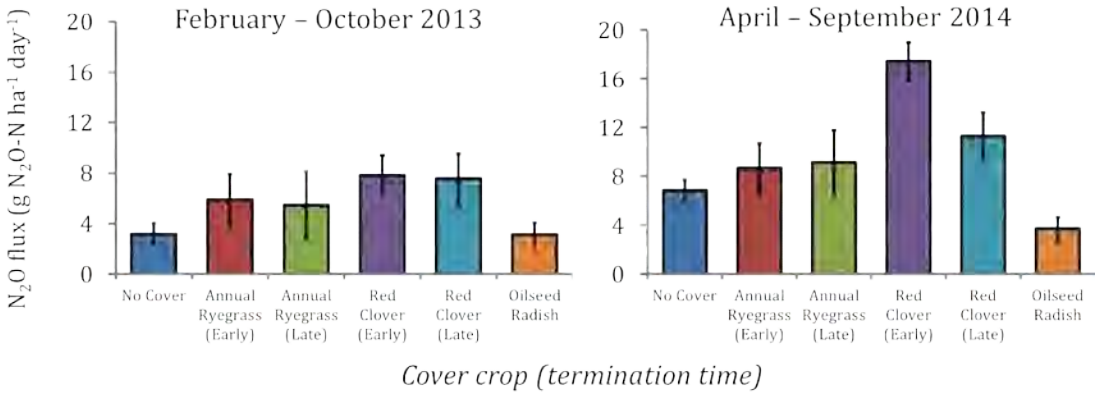


Fig 3. Average daily N2O fluxes from no cover crop and cover crop treatments during corn phase following cover crop termination (2013 and 2014).

Funding statement: This work was funded in part by the Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494), the DOE OBP Office of Energy Efficiency and Renewable Energy (DE-AC05-76RL01830), and the USDA NIFA grant ORG: 2011-51106-31046.

124. Switchgrass Augments Its Nitrogen Supply with Fixed Nitrogen

Sarah S. Roley1*(roleysar@msu.edu), Chao Xue2, James M. Tiedje2, G. Philip Robertson1,2

1W.K. Kellogg Biological Station, Michigan State University, Hickory Corners, MI;
2Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI

https://www.glbrc.org/research/sustainability

Project Goals: The overall goal of this project is to delineate the nitrogen (N) cycle of perennial grasses grown for conversion to cellulosic ethanol. As part of this overall goal, we sought to determine: 1) if N fixation can be detected in switchgrass, in the field; 2) the rate of fixation; 3) the location of fixation (plant tissue or soil); and 4) the identity of bacteria responsible for fixation.

Abstract text: Switchgrass (Panicum virgatum) is a candidate species for conversion to cellulosic ethanol. Native to North America, it is a warm-season grass that has high water-use and nitrogen (N)-use efficiency. At the Kellogg Biological Station site of the Great Lakes Bioenergy Research Center, we have observed a lack of switchgrass yield response to N fertilizer addition. This remarkable N-use efficiency has also been observed at other locations, and we hypothesized that switchgrass was obtaining N via biological nitrogen fixation (BNF).

Plants capable of BNF are potentially a sustainable choice for biofuels production systems. Fertilizer addition is one of the main sources of greenhouse gases from these systems; its production and field application both result in carbon dioxide emissions and soil emissions of nitrous oxides. Plants capable of BNF require less fertilizer and thus will result in less greenhouse gas production.

To detect BNF in the field, we injected 15N2 gas into the rhizosphere of switchgrass plants grown in unfertilized and fertilized plots. Prior to and after the injection, we sampled switchgrass leaves, stems, roots, and rhizosphere soil from both the plant receiving the 15N2 gas and from a plant with an unenriched atmosphere (control plant) in the same plot. We analyzed the N isotope composition of each, with post-injection enrichment evidence of fixation. Preliminary data show enrichment in switchgrass leaves and roots grown in unfertilized plots (Fig. 1).

To determine the rate of fixation, we placed four switchgrass plants into an airtight chamber in the greenhouse. We enriched the chamber with 15N2 gas for two days, and then analyzed the isotopic composition of the plant parts. We found enrichment in all plant tissues (leaves, stems, roots, and soil), relative to control plants that were incubated in a chamber with 14N2 gas. An average of 38% of total plant N was from fixation, and the average fixation rate was 8 mg N fixed/g plant/d. On an annual basis, these rates scale up to an average of 48 kg N/ha/yr, which is comparable to rates measured in sugar cane and soy beans.

To determine the location of fixation, we incubated individual plant tissues (leaves, stems, washed roots, and surface-sterilized roots) and sieved soil in gas-tight vials that were enriched with 15N2 gas. After 10 days of incubation, we found enrichment in roots and sieved soils (t-test, $p < 0.05$), and marginally significant enrichment in stems (t-test, $p < 0.1$), but not in leaves or sterilized roots, indicating that fixation was occurring in soils and on root surfaces (Fig. 2).

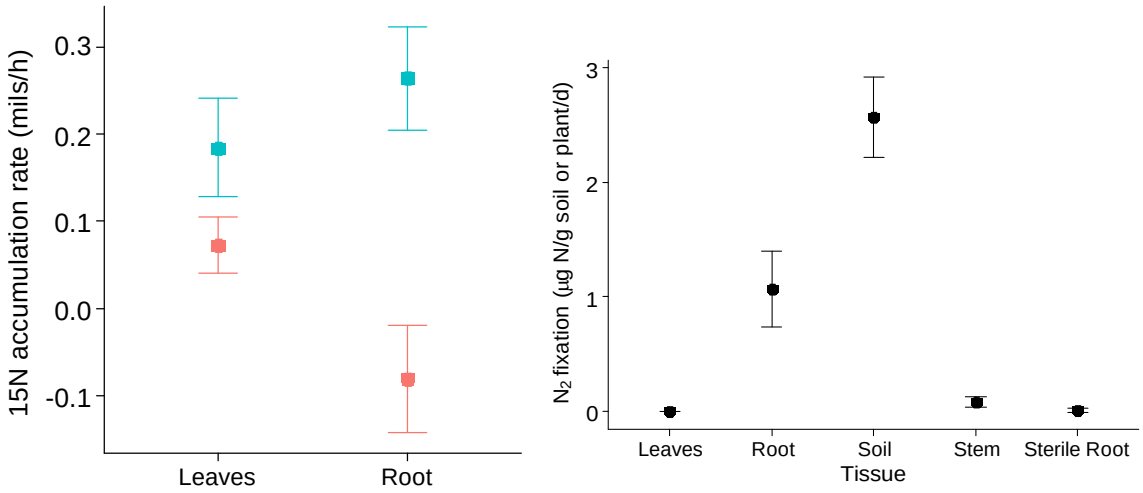


Figure 1 (left). 15N accumulation rate of switchgrass leaves and roots in plants that received 15N2 gas (treatment plants, blue dots) and in control plants that did not receive any gas inputs (red dots). Enrichment in the treatment plants, relative to controls, is evidence of N2 fixation.

Figure 2 (right). N2 fixation rates in various switchgrass tisues and sieved rhizosphere soils.

For the microbial analyses, we surface-sterilized all plant parts, and then extracted and sequenced DNA from each of the tissues, using nifH primers to separate out N2 fixers. We found a large diversity of N2-fixing microbes in switchgrass soils, roots, leaves, and stems; all tissues had at least 38 distinct genera present. The unfertilized soil diazotroph community was distinct from the community in the fertilized soils, and the unfertilized roots had a community distinct from the other plant tissues, as calculated via Bray-Curtis distance. Leaves and stems were not distinct from one another or between fertilizer treatments.

Overall, this project demonstrates that switchgrass is able to augment its N supply with fixed N, and that the rates are of agricultural significance. Fixation likely occurs by microbes that are closely associated with the surface of switchgrass roots, as well as microbes living freely in the soil. Switchgrass may have lower N fertilizer needs than previously assumed, potentially making it a sustainable choice for cellulosic biofuels cropping.

125. Comparative Fermentation of Saccharomyces cerevisiae and Zymomonas mobilis in Lignocellulosic Hydrolysates Produced from Corn Stover and Switchgrass to Investigate the Effect of Interannual Climate Variability on Hydrolysate and Microbial Performance

Yaoping Zhang1 (yzhang8@wisc.edu), Rebecca Garlock Ong,2,3 Jeff Piotrowski, 1 Gregg Sanford, 1 Dustin Eilert, 1 Dave Cavalier, 2 and Donna Bates 1

Great Lakes Bioenergy Research Center, 1University of Wisconsin-Madison, Madison, WI, 2Michigan State University, East Lansing, MI, and 3Michigan Technological University, Houghton, MI

https://www.glbrc.org

Project Goals: Different biomass feedstocks will be necessary for sustainable biofuel production from lignocellulose. Therefore, it is crucial to understand microbial responses to hydrolysates produced from different feedstocks. We produced lignocellulosic hydrolysates using AFEXTM-pretreated corn stover and switchgrass harvested from three different years with significantly different levels of precipitation and then performed comparative fermentations to study microbial responses to these different feedstock hydrolysates. The chemical compositions of these lignocellulosic hydrolysates were analyzed and the comparative fermentation experiments were carried out using Saccharomyces cerevisiae and Zymomonas mobilis strains. The growth, glucose and xylose utilization, and ethanol production were monitored during fermentation, and RNA samples were collected for RNAseq analysis to generate high sensitivity profiles and further define microbial responses in these different hydrolysates.

In order to identify and overcome key barriers in the sustainable conversion of lignocellulosic biomass to biofuels, we have produced more than 21 batches of hydrolysates using AFEXTM- pretreated corn stover and switchgrass harvested from three different years with significantly different levels of precipitation (Y2010 a wet year, Y2012 a dry year, and Y2013 an average year). A three-tiered strategy has been implemented to understand microbial responses to these diverse feedstock hydrolysates: Tier I - Hydrolysate Compositional Analysis. After hydrolysates were produced using our standard protocol, the concentrations of sugars, phenolics, organic acids, aldehydes, and other well-known inhibitors in these hydrolysates are analyzed by HPLC, GC/MS and LC/MS. This provides us basic knowledge about the variation of compounds in different hydrolysates. Tier II - Chemical Genomics for Fingerprinting of Hydrolysates. S. cerevisiae and Z. mobilis deletion collections have been used for chemical genomics to generate biological fingerprints of the hydrolysates. Tier III - First-pass Multiomics Fermentation. We conduct fermentation experiments using engineered yeast and Z. mobilis strains on hydrolysates that show significant differences in Tier I and II studies. Cell growth is monitored, and samples for endproduct (HPLC-RID) and RNAseq are collected. The endproduct analysis will determine the efficiency of converting sugars (both glucose and xylose) into ethanol whereas the RNAseq will generate high sensitivity gene expression profiles for responses of S. cerevisiae and Z. mobilis in these different hydrolysates. In collaboration with JGI, RNAseq data will be analyzed to determine how microbial responses (e.g., stress responses, expression of efflux pumps and regulators, etc.) vary among the different feedstocks.

Hydrolysate compositional analysis showed that overall corn stover hydrolysates (ACSH) contain higher levels of lignocellulose-derived inhibitors than switchgrass hydrolysates (ASGH). Chemical genomics with a S. cerevisiae deletion mutant library indicated distinct chemical genomics profile when different feedstock hydrolysates were used. Fermentation experiments also showed that S. cerevisiae and Z. mobilis have different responses to these different hydrolysates. Specifically Z. mobilis showed very similar growth and glucose/xylose utilization in ACSH and ASGH, while S. cerevisiae grew significantly

better in ACSH than ASGH. There are also effects of interannual climate variability on microbial fermentation performance, especially when S. cerevisiae was used for fermentation. In collaboration with JGI, RNAseq data analysis is underway to determine how microbial responses (e.g., stress responses, expression of efflux pumps and regulators, etc.) vary among the different feedstocks. This information will help us to identify the extent to which different feedstocks produce different microbial responses and how bottlenecks for the conversion of biomass to biofuel vary among different feedstocks.

We are currently expanding our studies to other feedstocks, including miscanthus, sorghum, and mixed-prairie. The chemical compositional analysis and comparative fermentation with these diverse feedstocks to determine microbial responses will generate knowledge to enable development of successful and sustainable lignocellulosic biofuel technologies.

This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494).

126. Ultrastructure of plant cells by electron microscopy: towards increased biofuel production

Purbasha Sarkar1,2, Maria Mueller1* (mmueller@lbl.gov), Michael Kowalczyk3, Salil Apte2, Edgar Yap2, Paul Adams1, Manfred Auer1,2

1Joint BioEnergy Institute, Physical Biosciences Division & 2Life Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA; 3Edinburgh Napier University, Edinburgh, United Kingdom

Project Goals: Determine mechanical cell wall properties and thylakoid membrane organization in chloroplasts for improving lignocellulosic biomass production.

Plant cell walls are the storehouse of complex sugars that if separated out efficiently from the wall lignins, can be fermented to produce commercially viable biofuel. There have been several approaches to genetically re-engineer the cell walls in various feedstocks with the ultimate goal of making the walls more accessible to deconstruction processes. At the same time, the modified cell walls need to maintain enough strength for the plants to grow and be functional. We pioneered a structural approach to determine the mechanical properties of plant cell walls. We are using qualitative and quantitative analysis of cryo electron tomography images and 3D computer-aided design (CAD) programs to develop realistic 3D macromolecular resolution models of various cell walls. Results of static stress-strain simulations of cell wall models allow an assessment of cell wall properties, which can be used for wall design improvement and optimization.

In addition to the cell wall project, we are studying structure and dynamics of thylakoid networks in chloroplasts of land plants. These organelles define plants, and they are responsible for photosynthesis as well as numerous other functions. Optimization of photosynthesis and carbon-fixation are key processes for biomass generation. Chloroplasts harvest the light and produce the biochemical energy that allows cell wall biomass production. To carry out this complex task, in both plants and algae, chloroplasts contain an elaborate architecture of complex lamellar membrane systems, also known as photosynthetic thylakoid membranes. These sheet-like membrane-bound compartments are arranged into regularly spaced stacks, called grana and non-stacked stroma thylakoids. To understand the interplay between the molecular processes of light harvesting and carbon-fixation and the morphology of thylakoid network, it is necessary to know how these networks develop, what the three-dimensional (3D) organization of thylakoid membranes at different time points of differentiation looks like and what kind of the structural changes take place under changing light conditions. Our approach aims at combining cryo immobilization techniques with focused ion beam/scanning electron microscopy (FIB/SEM). Cryo- immobilization can be achieved by high-pressure freezing (HPF), which allows for the vitrification of cells or whole tissues. Combined with a recentlydeveloped quick protocol for freeze-substitution (FS) and resin embedding, this method leads to sample preservation that is very close to the native state without imposing size limitations. With the exception of a few very small cells, more complex eukaryotic specimens must be thinned to make them transparent to the electron beam. In contrast, using FIB/SEM, there is no need for thin samples. This technology can be used to directly cross-section and image biological samples, such as cells or tissues for 3D reconstructions. By combining these two methods, we plan to obtain 3D volumes of chloroplasts in a close to native condition and in their native cellular context.

This work was supported by This work was conducted at Joint BioEnergy Institute and was supported by the Office of Science, Office of Biological and Environmental Research, U.S. Department of Energy under Contract No. DE- AC02-05CH11231.

127. Substantial Improvements in Methyl Ketone Production in E. coli from Rational Metabolic Engineering and Random Mutagenesis

Ee-Been Goh,1 Edward Baidoo,1 Xiaoliang Cheng,1 Helcio Burd,1 Taek Soon Lee,1 Jay D. Keasling,1 Trent R. Northen,1 and Harry R. Beller1,* (HRBeller@lbl.gov)

1Joint BioEnergy Institute (JBEI), Emeryville, CA

Project Goals: The Joint BioEnergy Institute (JBEI) aims to produce a chemically diverse suite of biofuels from lignocellulosic biomass. Some JBEI fuels use fatty acids as precursors, as these biomolecules are highly reduced, aliphatic compounds that, when modified (e.g., decarboxylated), can have properties comparable to those of petroleum- derived fuel components. The goal of the project presented here is to engineer E. coli to produce diesel-range methyl ketones in the gram-per-liter range with yields of at least 40% of maximum theoretical yield.

We have engineered Escherichia coli to overproduce saturated and monounsaturated aliphatic methyl ketones in the C11 to C15 (diesel) range; this group of methyl ketones includes 2- undecanone and 2-tridecanone, which have favorable cetane numbers and are also of importance to the flavor and fragrance industry. We have made specific improvements that resulted in more than 10,000-fold enhancement in methyl ketone titer relative to that of a fatty acid-overproducing E. coli strain, including the following: (a) overproduction of β-ketoacyl-coenzyme A (CoA) thioesters achieved by modification of the β-oxidation pathway (specifically, overexpression of a heterologous acyl-CoA oxidase and native FadB, and chromosomal deletion of fadA) and (b) overexpression of a native thioesterase (FadM). The first generation of engineered E. coli (Goh et al. 2012) produced ~380 mg/L of methyl ketones in rich medium. We have subsequently made additional genetic modifications that included balancing overexpression of fadR and fadD to increase fatty acid flux into the pathway, consolidation of the pathway from two plasmids into one, codon optimization, and knocking out key acetate production pathways (Goh et al. 2014). These modifications have led to a methyl ketone titer of 1.4 g/L with 1% glucose in shake flask experiments, which represents 40% of the maximum theoretical yield, and also attained titers of g/L after ~45 h of fed-batch glucose fermentation (the best values reported to date).

In addition to rational metabolic engineering, we have conducted chemical (MNNG) mutagenesis studies of a high methyl ketone-producing strain and screened using NIMS (Nanostructure-Initiator Mass Spectrometry) technology. Shake-flask incubations of 10 of the most promising mutants resulted in as much as 3- to 4-fold increases in titer relative to the control (un-mutated) strain. We are currently carrying out re-sequencing of the mutants to identify which mutations resulted in the greatest increases in methyl ketone production.

We have also conducted in vitro assays with purified pathway enzymes, which confirmed that a decarboxylase is not required to convert β-keto acids into methyl ketones and that FadM is promiscuous and hydrolyzes not only β-ketoacyl-CoAs but also other CoA-thioester pathway intermediates. These in vitro results have provided insight on how to fine-tune expression of pathway genes for further optimization of methyl ketone production.

References

1. Goh, E.-B., E. E. K. Baidoo, J. D. Keasling, H. R. Beller. 2012. Engineering of bacterial methyl ketone synthesis for biofuels. Appl. Environ. Microbiol. 78:70-80.
2. Goh, E.-B., E. E. K. Baidoo, H. Burd, T. S. Lee, J. D. Keasling, H. R. Beller. 2014. Substantial improvements in methyl ketone production in E. coli and insights on the pathway from in vitro studies. Metabolic Engineering 26:67-76.

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.

128. Quantitative Metabolic Modeling at Fuel Synthesis Division (JBEI)

Hector Garcia Martin1,2*(hgmartin@lbl.gov), Jennifer Gin1,2, Chris Shymansky1,2, Amit Ghosh1,2, Garrett Birkel1,2, Victor Chubukov1,2, Aindrila Mukhopadhyay1,2, Adam Arkin1,3, Jay Keasling1,2,3,4

1Fuel Synthesis Division, Joint BioEnergy Institute, Emeryville, California; 2Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA; 3Department of Bioengineering, University of California, Berkeley, CA; 4Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA

http://www.jbei.org

Project Goals : The goal of the Quantitative Metabolic Modeling directorate at Fuels Synthesis at the Joint BioEnergy Institute (JBEI) is to develop models of metabolism which are both quantitative and predictive, in order to improve biofuel production in a rationally directed fashion. We use experimental, computational and mathematical tools to achieve this goal.

Our efforts are divided into three main areas: Flux-based models of metabolism: We have developed a method to use 13C carbon labeling experimental data to constrain fluxes for genome-scale models. The measurement of fluxes for comprehensive genome-scale metabolic model allows us to improve the yield of fattyl-acid derived biofuels in S. cerevisiae.

Data mining of high-throughput -omics data: We are using the data coming from high-throughput proteomics experiments to adjust promoter strength and improve production of limonene and bisabolene. Development of web-based tools for -omics data visualization and storage: We have created two tools: the Experiment Data Depot (EDD) and the Multiomics Visualization Tool (MvT). The EDD provides not only a storage site for metabolic data but also a visualization tool and an easy way to download data in SBML format, which can be used for flux analysis and integrated with (e.g.) the DOE KBase. MvT allows for the visualization of metabolic maps at different scales of resolution using an interactive interface.

129. DIVA: More Science, Less DNA Construction

Joanna Chen1,2, Hector A. Plahar1,2, Manjiri Tapaswi1,2, Garima Goyal1,2, Nina Stawski1, Jay D. Keasling1, and Nathan Hillson1,2* (njhillson@lbl.gov)

1Fuels Synthesis and 2Technologies Divisions, Joint BioEnergy Institute, Emeryville, CA

https://diva.jbei.org

Project Goals:
DNA construction is vital to a broad range of biological research, yet it is predominantly an inefficient diversion from researchers' core research goals and expertise. We have developed a Design, Implementation, and Verification Automation (DIVA) platform to liberate researchers from building DNA, enabling them to instead focus on designing and testing their experiments of interest. A number of software tools have been developed to aid researchers design and build DNA constructs. The DIVA platform seeks to integrate these tools into a seamless and useful workflow to further save researchers time and effort. An initial version of the DIVA platform has been deployed at JBEI and is under continued development. We are also working closely with the Joint Genome Institute towards deploying DIVA to better serve their user community.

References

1. Chen, J., Densmore, D., Ham, T.S., Keasling, J.D., Hillson, N.J. (2012) DeviceEditor visual biological CAD canvas. J Biol Eng 6 (1).
2. Ham, T.S., Dmytriv, Z., Plahar, H., Chen, J., Hillson, N.J., Keasling, J.D. (2012) Design, implementation and practice of JBEI-ICE: an open source biological part registry platform and tools. Nucleic Acids Res 40 (18):e141.
3. Hillson, N.J.*, Rosengarten, R.D., and Keasling, J.D. (2012) j5 DNA Assembly Design Automation Software. ACS Synthetic Biology 1 (1), 14-21.
4. Linshiz, G., Stawski, N., Goyal, G., Bi, C., Poust, S., Sharma, M., Mutalik, V., Keasling, J.D., and Hillson, N.J. (2014) PR-PR Cross-Platform Laboratory Automation System. ACS Synthetic Biology 3 (8), 515–524.

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.

130. Advanced pathways for microbial production of branched C5 alcohols

Aram Kang, Kevin George, Edward Baidoo, George Wang, Jay Keasling, Taek Soon Lee* (tslee@lbl.gov)

Fuels Synthesis Division, Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, CA 94608, USA

Project Goals: The Joint BioEnergy Institute (JBEI) aims to produce a chemically diverse suite of biofuels from lignocellulosic biomass. Isoprenoid-based biofuels have been of great interest due to their superb fuel properties such as low freezing temperature and high octane number. Mevalonate (MVA) pathway is one of the major biosynthetic pathways of isoprenoid fuel production, and the engineering of this pathway is a key approach to achieve higher production of these biofuels. Various engineering strategies and tools have been explored to identify the bottlenecks of the pathway and to understand the pathway enzymes better, but the intrinsic energy demands of this pathway and the operational cost for aeration to meet the energy demands have been still a problem when a production in large scale using fermentor is exploited. In this work, we present modified version of the MVA pathway that will address these issues for isoprenoid biofuel production. Isoprenoids are considered as one of the most promising advanced biofuels. Among the isoprenoid compounds, branched five carbon (C5) alcohols have been tested as good biofuel compounds with favorable combustion properties and octane numbers to gasoline. A synthetic pathway for C5 alcohols production has been reported previously in E. coli using the heterologous MVA pathway (Chou and Keasling, 2012), and further metabolic engineering efforts on this synthetic pathway have led to about 50% theoretical yield in the C5 alcohol production (George et al. 2014). Even though the MVA pathway has been known to be less efficient than the methylerythritol phosphate (MEP) pathway in carbon and redox balance as well as in energy balance, the MVA pathway has been extensively used for microbial production of a range of valuable isoprenoids due to its tractability and the high titers it can provide. However, the energy demands of this pathway and the operational cost for aeration to meet these energy demands have been a problem, especially when a production in large scale using fermentor is exploited. In this work, we present modified pathways for C5 alcohol production that will address these issues of the traditional MVA pathways. We engineered advanced pathways for C5 alcohol production that reduce cellular costs for isopentenol production. One of the modified pathways showed that isopentenol could be produced via decarboxylation of mevalonate monophosphate to isopentenyl monophosphate by the promiscuous activity of the decarboxylase. The titer and the growth of the engineered strains with this modified pathway were better than those with the original pathway, and currently the efficiency of this modified pathway is tested under microaerobic condition. The decarboxylase enzyme engineering and the pathway optimization of this modified pathway would lead a microbial C5 alcohol production more economically feasible, especially for large scale industrial application.

References

1. Chou, H.H., Keasling, J.D., Synthetic pathway for production of five-carbon alcohols from isopentenyl diphosphate. Appl. Environ. Microbiol. 2012, 78, 7849-7855
2. George, K.W., Chen, A., Jain, A., Batth, T.S., Baidoo, E., Wang, G., Adams, P.D., Petzold, C.J., Keasling, J.D., Lee, T.S., Correlation analysis of targeted proteins and metabolites to assess and engineer microbial isopentenol production, Biotech. Bioeng. 2014, 111, 1648-1658

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.

131. Tackling biomass recalcitrance with novel biological approaches

Yan Liang1, Aymerick Eudes1, Vimalier Reyes-Ortiz1, Edward Baidoo1, Sasha Yogiswara1,2, Veronica Teixeira Benites1, Clarabelle Cheng-Yue1, George Wang1, Ming Hammond2,3, Jay Keasling1, Dominique Loqué1* (dloque@lbl.gov)

1Joint BioEnergy Institute, Physical Biosciences division, Lawrence Berkeley National Laboratory, 1 Cyclotron road, Berkeley, CA 94720, USA; 2University of California, Department of Chemistry, Berkeley, CA, 947203; 3University of California, Department of Molecular and Cell Biology, Berkeley, CA, 947203

http://www.jbei.org/research/divisions/feedstock/cell-wall-engineering/

Project Goals: The main goals of this research are to develop and validate technologies that reduce biomass recalcitrance

The development of alternative transportation fuels that can meet future demands while reducing global warming is critical to the national, environmental, and economic security of the United States. Currently, sugars used for biofuels production are largely derived from sucrose extracted from sugarcane and starch from corn, but there is a large, untapped resource (more than a billion tons per year) stored in plant biomass that could be utilized for liquid fuels production. However, cell wall recalcitrance hampers the development of cost-effective and energy-efficient processes to convert this biomass into fermentable sugars. Plant cell walls are mainly composed of cellulose, hemicellulose and lignin and their recalcitrance to enzymatic hydrolysis is primarily caused by lignin followed by cellulose crystallinity. Both are essential for plant development as they provide structural support to plants and protect plants against multiple stresses, which render their manipulations more challenging.

We developed non-conventional approaches and used synthetic biology to tackle both bottlenecks. We first developed and applied biological switches to express a defective cellulose synthase protein in fiber cells to reduce cellulose crystallinity without affecting plant growth. In parallel we elaborated a novel strategy to biologically inhibit lignin biosynthesis and primarily targeted hydroxycinnamoyl transferase (HCT) activity. We screened and identified HCT competitive inhibitors and reconstituted in situ metabolic pathways to inhibit lignin biosynthesis in secondary cell wall. In planta characterization revealed that both approaches were successful in reducing biomass recalcitrance and increasing sugar yield.

This work was part of the DOE Early Career Award and 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.

132. Universal expression tools to improve nutrient acquisition of energy crops

Nasim Mansoori Zangir1* (nmansoorizangir@lbl.gov), Carlos Hernandez-Garcia1, Patrick Shih1, Charlee Vuu1, Garima Goyal1, Sangeeta Nath2, Nathan J. Hillson1,2, Samuel Deutsch2, Dominique Loqué1

1Joint BioEnergy Institute, Physical Biosciences division, Lawrence Berkeley National Laboratory, 1 Cyclotron road, Berkeley, CA 94720, USA; 2DOE Joint Genome Institute, 2800 Mitchell Drive, B310-346A, Walnut Creek, CA 94598, USA

http://www.jbei.org/research/divisions/feedstock/cell-wall-engineering/

Project Goals: The main goals of this research are the generation of "universal" expression tools for plant root engineering and to utilize them to improve nutrient acquisition in energy crops.

Plant growth and development relies on roots as a means to anchor the plant body as well as to absorb water and nutrients such as N, P and K. As deficiency in these nutrients negatively impacts photosynthesis, plant growth and yield, chemical fertilizers have been used to compensate and meet the growing demand for plant material. Unfortunately the excessive use of fertilizers has come at high environmental and economic costs; and its production utilizes a substantial proportion of worldwide energy consumption. The current development of growing crops on marginal land (low water content, low nutrient supply, vulnerability to erosion and heavy-metal pollution) for bioenergy will reduce competition with food crops and the pressure on high-quality arable lands utilization. As sufficient nutrient uptake by the plant root is vital especially when soil conditions restrict its availability, root systems in energy crops can be engineered to acquire/ accumulate the required nutrients.

Our focus is to design "universal" expression tools for plant root engineering functional across diverse plant species. Through targeted-cell engineering and synthetic biology we aim to engineer plant root systems to generate metabolic pathways and improve nutrient acquisition/accumulation in energy crops. This aim will be carried out through development and validation of a large set of root-specific and nutrient-responsive promoters. In house transcriptome analysis will help identify root specific genes that are constitutively expressed, induced or repressed through conserved metabolic responses to N, P and Fe starvation in hydroponic conditions across a widely diverse selection of plant species. Promoters from the identified genes will be isolated and characterized for spatiotemporal expression patterns and expression levels across different taxonomic classes using composite plants and conventional transgenic approaches. Subsequently monocot and dicot "universal" promoter libraries of the detected genes alongside those previously reported will be generated using a semi-high-throughput promoter cloning approach and yeast homologous recombination assembly method and ultimately used for tissue specific metabolic pathway engineering to improve nitrogen use efficiency in bioenergy crops.

We expect that the results of this research will generate a diversity of building blocks for plant engineering and will directly contribute to advance the DOE's mission for the sustainable production of bioenergy.

This work was part of the DOE Early Career Award and 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 U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

133. Genetic Parts Screening and Artificial N-glycosylation Motif Engineering for

Heterologous Protein Production in Aspergillus niger

Jinxiang Zhang1, Saori Amaike-Campen1, Sam Deutsch2, Matthew Hamilton2, Ljiljana Paša- Tolić3, Erika Zink3, Jon Jacobs3, Blake Simmons1, Scott Baker1, John Gladden1, Jon Magnuson1* (jon.magnuson@pnnl.gov)

1Joint BioEnergy Institute, Emeryville, CA; 2Joint Genome Institute, Walnut Creek, CA; 3Environmental Molecular Sciences Laboratory, Richland WA

http://www.jbei.org/

Project Goals
The principal goal of the Fungal Biotechnology Team is building an efficient fungal platform for heterologous gene expression in fungi, which accommodates the array of largely prokaryotic enzymes with superior properties discovered and characterized in the other teams of the Deconstruction Division at JBEI. The main platform organism is Aspergillus niger and we are applying systems biology approaches to understanding heterologous enzyme production, and employing that understanding for the engineering of a better protein production platform.

Abstract
Aspergillus niger is a genetically tractable model organism for scientific discovery and a platform organism used in industry for the production of enzymes. Expression of secreted native enzymes at tens of grams per liter have been discussed by those in industry, but high level production of heterologous enzymes remains elusive. Strategies to increase production include the use of strong promoters, protease-deficient strains, fusion proteins, multiple gene copies, etc. However, yields of heterologous proteins are still lower than desired. In collaboration with the Environmental Molecular Sciences Laboratory (EMSL), we generated proteomics data from secretome samples of A. niger grown on a variety of minimal and rich media, with the goal of identifying useful genetic elements for increasing heterologous protein production. Twenty promoters, six signal sequences and four introns from the most highly secreted proteins were identified as candidate genetic elements to enhance heterologous gene expression. These candidate elements were tested for their ability to drive expression of a prokaryotic glycoside hydrolase. A vector was designed to target integration of the modified expression cassette to the native glucoamylase gene locus by homologous recombination. Considerable diversity was seen in heterologous protein production driven by these various elements. Interestingly, one promoter for a membrane protein and a non-classical signal peptide showed promising results. In addition, based on the 3D structure of a heterologous glycoside hydrolase, we generated seven individual artificial N-glycosylation motifs on the surface of the heterologous protein and had the genes synthesized in a collaboration with the Joint Genome Institute (JGI). Analysis of these modified proteins for glycosylation, using EMSL's top-down proteomics capability, is underway.

Correlation of glycosylation with any changes in enzyme kinetic and thermodynamic properties of the altered proteins is the goal of this aspect of the research.

134. Heterologous Expression of Ionic Liquid Tolerant Cellulases in Aspergillus niger

Saori A. Campen,1 John M. Gladden1,3* (jmgladden@lbl.gov), Stephanie J. Sibert,1 Sneha Srikrishnan,1 Jed Lynn,1 Pallavi Phatale,1 Jinxiang Zhang,1 Joel M. Guenther,1 Taya Feldman,1 Jennifer Hiras,1 Steven W. Singer,1 Kenneth L. Sale,1,3 Blake A. Simmons,1,3 Scott E. Baker,1,2 Jon K. Magnuson,1,2

1 Joint Bio Energy Institute, Emeryville, CA 94608; 2 Pacific Northwest National Laboratory, Richland, WA 99352; 3 Sandia National Laboratory, Livermore, CA 94551-0969

http://www.jbei.org/research/divisions/deconstruction/fungal-biotechnology/

Project Goals: Our objective is to understand the mechanisms that drive high titer enzyme production in A.niger and use this knowledge to further its development as a heterologous expression host for high titer production of designer cellulase cocktails.

Cheap and efficient deconstruction of plant biomass is critical for the successful commercialization of lignocellulosic biofuels. Biomass pretreatment with ionic liquids (ILs) has been demonstrated to dramatically increase saccharification efficiency under reduce cellulase enzyme loadings, and therefore shows great promise in this regard. However, some ILs that are good for pretreatment, such as 1-ethyl-3-methylimidazolium acetate, strongly inhibit commercial cellulase enzymes, and therefore excessive amounts of water are required to remove the IL from the biomass prior to saccharification. To address this issue, several IL-tolerant bacterial cellulases have been discovered and assembled into an IL-tolerant cellulase mixture, called JTherm, capable of hydrolyzing pretreated biomass in the presence of 1-ethyl-3-methylimidazolium acetate.

Currently, we are attempting to produce high titers of these IL-tolerant enzymes in Aspergillus niger to enable further research and to demonstrate commercial viability of this technology. A. niger is an ascomycete filamentous fungus commonly used in industry to produce citric acid a high titers of the glycoside hydrolase glucoamylase. The fact that this organism is already used in industry, is amenable to genetic manipulation, and has been demonstrated to product high titers of a single enzyme makes it a good choice to develop into a host for heterologous enzyme production. To understand how amenable this fungi is to expressing heterologous IL-tolerant enzymes, thirty-two bacterial and fungal genes, encoding beta- glucosidases, cellobiohydrolases and endoglucanases were screened for expression in A. niger. To probe the functionality of the enzymes that were successfully expressed in A. niger, we used the Jsalsa (the JBEI suite for automated lignocellulosic saccharification) platform to obtain activity profiles with respect to temperature, pH and ionic liquid and compared them to profiles of the same enzymes expressed in E.coli. It was found that many of the enzymes that expressed well in E. coli also expressed well in A. niger, and the enzymes that were profiled on Jsalsa appear to be functionally equivalent. Finally, we used enzymes produced by A. niger to assemble an IL-tolerant cellulase cocktail that was able to hydrolyze IL-pretreated biomass.

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.

135. GONST1 and 2 are GIPC Specific GDP-Mannose Transporters

Beibei Jing,1 Toshiki Ishikawa,2 Noriko Inada,3 Xiaolan Yu,4 Jeemeng Lao,1 Taku Demura,3,5 Maki Kawai-Yamada,2 Paul Dupree,4 Jenny C. Mortimer1,4,5* (jcmortimer@lbl.gov)

1Joint BioEnergy Institute, Berkeley CA, USA; 2Graduate School of Science and Engineering, Saitama University, Japan; 3Graduate School of Biological Sciences, NAIST, Japan; 4Department of Biochemistry, University of Cambridge, United Kingdom; 5Cellulose Production Research Team, BMEP, CSRS, RIKEN Yokohama, Japan

http://www.jbei.org/research/divisions/feedstock/plant-systems-biology/

Project Goals:
We want to understand how glycosylation is controlled in plants, and to use this knowledge to engineer plants with improved glycan profiles for bioenergy, industry and human health. Our focus is the mannan biosynthetic pathway. Mannan is a C6 polysaccharide, the most abundant polysaccharide after cellulose in softwoods such as pine. However, in most crop plants, as well as the model plant Arabidopsis, it is a relatively minor polymer. The glycosyltransferases responsible for its synthesis have been identified, but attempts to engineer increased amounts have had limited success.

We are identifying and characterizing other components of the mannan synthetic pathway. This allows us to use JBEI-developed plant synthetic biology tools to re-engineer the cell wall. Our initial goal is to develop plants which have an increased accessible hexose content in biomass. This will reduce downstream processing costs for bioenergy, due to microbial preferences for hexoses over pentoses during fermentation.

In this project, the aim was to identify Golgi-localized nucleotide sugar transporters which provide substrates for mannan biosynthesis. Characterization of these transporters will allow us to boost substrate availability for the glucomannan synthases.

Abstract:
Nucleotide sugar transporters (NSTs) translocate the substrates for cell wall biosynthesis and other glycosylation processes into the Golgi from the cytosol. NSTs responsible for the cell wall precursors GDP-Fuc, GDP-Glc, GDP-Man and GDP-L-Gal have yet to be identified. The Arabidopsis thaliana protein GOLGI-LOCALIZED NUCLEOTIDE SUGAR TRANSPORTER 1 (GONST1) and its close homolog GONST2 have been both previously identified as GDP-Man transporters. In vitro, GONST1 and GONST2 can both rescue a yeast GDP-Man transporter mutant1. In a liposome assay, GONST1 can transport all four plant Golgi GDP sugars (GDP-Man, GDP-Glc, GDP-Fuc, and GDP-L-Gal)2. GONST1 and 2 were therefore predicted to supply substrates for glucomannan biosynthesis in vivo. However, our analysis showed that the plants were unaffected in glucomannan biosynthesis, or any other cell wall polymer.

Sugars also decorate glycosphingolipids, and this process is also believed to occur in the Golgi. Glycosylinositolphosphoceramides (GIPCs) are the most abundant sphingolipid in the plant plasma membrane. We demonstrated that Arabidopsis GIPCs contain mannose sugar decorations that are dramatically decreased in gonst1. gonst2 shows a small reduction in GIPC mannosylation and gonst1 gonst2 plants completely lack mannosylated GIPCs, which indicated GONST1 and GONST2 have redundant function in GIPC mannosylation. We conclude that GONST1 and GONST2 specifically transport GDP-Man as substrates for GIPC biosynthesis. gonst1 displays a constitutive hypersensitive response. gonst2 has normal growth but gonst1 gonst2 is severely stunted with leaf lesions and an early

senescence phenotype. The characterization of theses mutants demonstrates that the loss of mannose from GIPCs can have a severe effect on plant development and immunity.

The link between GIPC glycosylation and immunity was previously unrecognized. The identification of new members of the biosynthetic pathway will contribute to the understanding of GIPC function.

References

1. Handford et al. 2004, Mol Gen Gen
2. Mortimer et al. 2013, Plant Cell

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. Work conducted at RIKEN was funded by a RIKEN FPR fellowship award to JCM, and work at Cambridge was funded by the BBSRC award BB/G016240/1 to PD.

136. Development of a High Throughput Platform for Screening GH Enzymes and Methyl Ketones based on Oxime-NIMS (Nanostructure-Initiator Mass Spectrometry)

Kai Deng1,2*(kdeng@lbl.gov), Xiaoliang Cheng1,3, Huu Tran1,2, Joel Guenther1,2, Vimalier Reyes-Ortiz1,3, Taichi Taksuka4, Ee-Been Goh1,3, Harry Beller1,3, Ben Bowen3, Henrik Geertz-Hansen1, Samuel Deutsch5, Blake Simmons1,2, Paul Adams1,3,6, Anup Singh1,2, Brian Fox4 , Trent Northen1,3

1Joint BioEnergy Institute, Emeryville, CA; 2Sandia National Laboratories, Livermore, CA; 3Lawrence Berkeley National Laboratory, Berkeley, CA; 4GLBRC, University of Wisconsin, Madison, WI; 5Joint Genome Institute, Walnut Creek, CA; 6University of California, Berkeley, CA

http://www.jbei.org

Project Goal: Our goal is to develop a high throughput platform to study the specificity and activity of large-scale GH enzymes and biofuel strains.

Abstract: Cost effective conversion of biomass into sugars for biofuel production requires high-performance low-cost enzyme cocktails that are active under the process conditions and biofuel producing strains with high-titers of the desired products.

Development of the requisite enzymes and strains requires large-scale studies of specificity and activity. However, the existing high throughput methods (e.g. DNS assays) lack the specificity needed to understand enzyme kinetics, which is crucial to identify bottle-neck step during the biomass deconstruction and to develop the required enzyme cocktails.

Here we describe a high throughput screening platform for glycosyl hydrolysases (GHs) activities and fatty acid methyl ketones that is based on bioconjugation chemistry and nanostructure-initiator mass spectrometry. The bioconjugation involves the oxime formation between a unique chemical probe and aldehyde (glycans) and ketone groups. The quantitative analysis is enabled by incorporation of C13 labeled internal standards. The key to the high throughput analysis is the novel integration of nanoliter-scale acoustic sample deposition with nanostructure-initiator mass spectrometry analysis (R&D100). This allows thousands of samples to be printed on a single mass spectrometry chip for rapid laser-desorption based analysis.

References

1. De rond et al Current Opinion in Biotechnology 2015, 3, 1-9.
2. Deng, K. et al. ACS Chemical Biology 2014, 7, 1470-1479.
3. Greving, M. et.al. Analytical Bioanalytical Chemistry 2012, 403, 707-11.
4. Reindl, W. et.al. Energy.Environ.Sci. 2011, 4, 2884-2893. 5. Northen, T. R. et.al. PNAS 2008, 105, 3678-3683.

Acknowledgements

The DOE Joint BioEnergy Institute and DOE Great Lakes Bioenergy Research Center are supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 and through contract DE-FC02-07ER64494, respectively. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's Nuclear Security Administration under contract DE-AC04- 94AL85000.

137. Applications of a high-throughput targeted proteomics assay for analysis of engineered Escherichia coli

Leanne Jade G. Chan,1,2 Melissa Nhan,1,2 Pragya Singh,1,2 Huu Tran,1,2 Jay D. Keasling,1- 4 Paul D. Adams,1,2,4 and Christopher J. Petzold1,2* (cjpetzold@lbl.gov)

1Joint BioEnergy Institute, Emeryville, California; 2Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California; 3Department of Chemical & Biomolecular Engineering, University of California, Berkeley; 4Department of Bioengineering, University of California, Berkeley

http://www.jbei.org

Project Goals:
Transformation of engineered Escherichia coli into a robust microbial factory is contingent on precise control of metabolism. Targeted proteomics is well suited for metabolic engineering due to rapid detection and accurate quantification of proteins. We created a high-throughput proteomics toolkit that provides a large number of validated and optimized targeted proteomics methods for absolute quantification of over 400 E. coli proteins. To facilitate high sample throughput, we have developed a fast chromatography method (10 minute total run time) and an automated sample preparation procedure that performs cell lysis, protein quantification, and tryptic digestion of hundreds of samples in a few hours. We demonstrate the utility of the toolkit by characterizing proteins in various metabolic pathways from exponentially growing E. coli cultured with glucose or xylose as the carbon source in aerobic and anaerobic conditions. From this work, principle component analysis using proteomics (PCAP) and correlation analysis methods that use targeted proteomic data have been developed to aid pathway optimization.

138. Genomic Resources for the Study of Cell Wall Biosynthesis in Grasses

Nhan T. Pham1, 2, *(nhpham@ucdavis.edu), Manoj K Sharma1,2,3, Rita Sharma1,2,4, Peijian Cao5, Mitch Harkenrider1, Jerry Jenkins6,7, Jane Grimwood6,7, Jiyi Zhang8, Jeremy Schmutz6,7, Michael K. Udvardi8, Rashmi Jain1, 2, Guotian Li1, 2, Mawsheng Chern1, 2, Liangrong Jiang1, Syed Mehar Ali Shah1, Wendy Schackwitz7, Joel Martin7, Henrik Scheller1, and Pamela Ronald1, 2

1Joint BioEnergy Institute, Emeryville, California, USA; 2Department of Plant Pathology and the Genome Center, University of California, Davis, California, USA; 3Present Address: Department of Biotechnology, Jawaharlal Nehru University, New Delhi, India; 4Present Address: Department of Life Sciences, Jawaharlal Nehru University, New Delhi, India; 5China Tobacco Gene Research Center, Zhengzhou Tobacco Research Institute, Zhengzhou, China; 6HudsonAlpha Institute of Biotechnology, Huntsville, Alabama, USA; 7Department of Energy Joint Genome Institute, Walnut Creek, California, USA; 8Plant Biology Division, Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, Oklahoma, USA; 9Center for Integrative Genomics and Department of Molecular and Cell Biology, University of California, Berkeley, California, USA.

www.jbei.org

Project Goals: To create genomic resources to identify genes predicted to control grass cell wall biosynthesis and modification in grasses

Due to its large mass, low input requirements and ability to adapt to a wide range of growing conditions, switchgrass is a candidate bioenergy crop. Cost efficient conversion of the lignocellulosic biomass in switchgrass and other grasses into biofuels will require basic knowledge of the genes that control grass cell wall biosynthesis and modification. Our lab is advancing these goals using two different approaches. First, we have identified switchgrass homologs of rice genes known to control biomass and stress-response related traits. We generated a switchgrass BAC library, screened 96,000 BAC clones, and in collaboration with the Joint Genome Institute (JGI), we have sequenced 311 BAC clones, corresponding to 51.7 Mb.

We identified 695 unique switchgrass genes predicted to control cell wall biosynthesis and stress responses. Using microarray and comparative phylogenomic analysis, we established the phylogenetic and functional relationships of the 87 glycosyltransferase 2 (GT2) family genes in rice and switchgrass.

We are also advancing rice genetic resource as an efficient method for identifying agronomically relevant grass genes. We created a collection of 4000 Kitaake rice mutants by fast-neutron mutagenesis. In collaboration with the JGI, we sequenced more than 1000 mutants and have analyzed 41 mutant lines in detail. From this analysis of 41 mutants we identified a total of 2,418 mutations affecting 1,433 unique genes. These mutations include 1,273 single base substitution (SBS), 864 deletions, 145 insertions, 82 inversions, 49 translocations, and 5 tandem duplications Figure 1). One of the mutants, rcs60, shows an altered saccharification profile. An online, user- friendly database, named KitBase (Figure 2) contains sequence and phenotypic data of each mutant has been created. These resources will facilitate studies of grass cell wall biosynthesis and modification.

This work is supported by the Office of Science of the U.S. Department of Energy Contract No. DE-AC02-05CH11231 to the Joint Genome Institute and a National Science Foundation grant (IOS-1237975) to PCR

139. Optimization of Recombinant Cellulase Mixtures for Degradation of [C2C1Im][OAc] Pretreated Biomass

Taya Feldman1,3, Joel Guenther2,3, Vimalier Reyes-Ortiz1, Huu Tran2,3, Anup Singh3, Paul Adams2,4, Blake A. Simmons1,3 and Kenneth L. Sale1,3,*

1Deconstruction and 2Technology Divisions, Joint BioEnergy Institute, Emeryville, CA; 3Sandia National Laboratories, Livermore, CA; 4Lawrence Berkeley National Laboratory, Berkeley, CA

http://www.jbei.org

Project Goals

The overall goal of this project is to optimize the ratios of a minimal set of cellulolytic enzymes that maximize glucose yields from 1-ethyl-3-methylimidazolium acetate ([C2C1Im][OAc]) pretreated biomass at 70 °C and 20% [C2C1Im][OAc]. Achieving this goal requires meeting two specific aims 1) develop a platform for characterizing large numbers of enzymes sampled from available glycosyl hydrolase and auxiliary enzyme families databases (e.g., CAZy) and from newly available studies of extremophilic microbial communities for their ability to remain stable and active at thermophilic temperatures and in the presence of [C2C1Im][OAc] and 2) develop an experimental design approach to optimizing both the set of enzymes in the mixture and their ratios.

Salts with Tm < 100 °C are called ionic liquids (ILs), and the IL 1-ethyl-3- methylimidazolium acetate ([C2C1Im][OAc]) has been shown to effectively pretreat lignocellulosic biomass such that enzymes can readily saccharify the cellulose into glucose at high yields (1, 2). Technoeconomic modeling at JBEI shows that an industrial- scale lignocellulosic biofuel process based on [C2C1Im][OAc] can be cost-competitive with fossil fuels (3). To sustain economic feasibility, the enzymes need to maintain robust activity at ≥70 °C in ≥20% ionic liquid, and the number and quantity of enzymes used must be minimized. Since biomass-degrading microorganisms have not evolved in [C2C1Im][OAc] environments, identifying C2C1Im][OAc] tolerant enzymes is challenging. Our strategy was to develop the JBEI Suite for Automated Lignocellulosic Saccharification (jSALSA), a high-throughput enzyme screening platform in which enzymes are characterized for their temperature, pH and [C2C1Im][OAc] concentration optima on a range of substrates, including [C2C1Im][OAc] pretreated biomass. Enzymes found to be stable and active in a minimum of 20% [C2C1Im][OAc] are then used in a design of experiments for mixtures to determine the ratios of a minimal set of recombinant enzymes that maximize glucose yields from [C2C1Im][OAc] pretreated biomass. Here we demonstrate this process to optimize the ratios of an endoglucanase, cellobiohydrolase and b-glucosidase mixture.

References

1. C. Li et al., Comparison of dilute acid and ionic liquid pretreatment of switchgrass: Biomass recalcitrance, delignification and enzymatic saccharification. Bioresoure Technology 101, 4900–4906 (2010).
2. J. Shi et al., One-pot ionic liquid pretreatment and saccharification of switchgrass. Green Chemistry 15, 2579–2589 (2013).
3. D. Klein Marcuschamer, B. A. Simmons, H. W. Blanch, Technoeconomic analysis of a lignocellulosic ethanol biorefinery with ionic liquid pretreatment. Biofuels, Bioproducts and Biorefining 5, 562–569 (2011).

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 Department of Energy under Contract No. DE-AC02-05CH11231.

140. Stacking of Biomass Traits

Rhea Stoppel1*(rstoppel@lbl.gov), Khanh Vuu1, Bianca Manalansan1, Patrick Shih1, Camille Chalvin1,2, Vibe Gondolf1,3, Berit Ebert1,3, April Liwanag1, Dominique Loqué1, Henrik Scheller1,4

1Feedstocks Division, Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA; 2Ecole Normale Supérieure de Cachan, 94230 Cachan, France; 3University of Copenhagen, Department of Plant and Environmental Sciences, Frederiksberg, DK-1871, Denmark; 4University of California, Department of Plant and Microbial Biology, Berkeley, CA, 94720

http://www.jbei.org/research/divisions/feedstock/cell-wall-biosynthesis/

Project Goals: Our focus is on the development of plants with biomass optimized for downstream processing into biofuels.

Engineering of plants with a composition of lignocellulosic biomass that is more suitable for downstream processing is of high interest for next-generation biofuel production. Lignocellulosic biomass contains a high proportion of pentose residues, which are more difficult to convert into fuels than hexoses. Therefore, increasing the hexose/pentose ratio in biomass is one approach for biomass improvement. A genetic engineering approach was used to investigate whether the amount of pectic galactan can be specifically increased in cell walls of Arabidopsis fiber cells, which in turn could provide a potential source of readily fermentable galactose. Stacking of AtUGE2 and GalS1 genes indicates that their simultaneous overexpression increases the cell wall galactose to much higher levels than can be achieved by overexpressing either one of these proteins alone. Engineering plants with complex metabolic pathways or multiple traits is challenging because of the number of introduced genes that are required to reach the final product. Therefore, there is a great need for synthetic biology tools to express multiple genes (gene stacking) with controllable expression strengths and in specific tissues. We present a strategy utilizing in vivo yeast homologous recombination to assemble multiple functional gene cassettes (promoter::ORF::terminator) together. To facilitate the upstream assembly of functional gene cassettes, we have developed a library of Golden Gate cloning-compatible promoters, ORFs, and terminators. Importantly, the assembly of these functional gene parts is not limited to Golden Gate assembly, providing scientist the flexibility to choose whatever method best suits their needs. Ultimately, functional gene cassettes are stacked using yeast assembly based on overlapping terminator sequences, enabling homologous recombination. Using this newly developed jStack method, we now aim to further improve the galactan engineering system and in addition combine the increase in galactan with a decrease in lignin. Furthermore, we aim to stack genes in order to engineer plants with increased galactan deposition in the roots, thus improving interactions between plants and beneficial microbes.

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.

141. MaxBin: An Automated Metagenomic Binning and Reassembly Algorithm

Steven W. Singer1,2*(SWSinger@lbl.gov), Yu-Wei Wu1,3, Jeffrey A. Kimbrel1,3, Stephanie Eichorst1,3, Yung-Hsu-Tang1,4, Janet Jansson1,2,5, Susannah G Tringe6, Blake A. Simmons1,5

1Joint BioEnergy Institute, Emeryville, CA; 2Earth Sciences and 3Physical Biosciences Divisions, Lawrence Berkeley National Laboratory, Berkeley, CA; 4San Francisco State University, San Francisco, CA; 5Biological Sciences Division, Pacific Northwest National Laboratory; 6Joint Genome Institute, Walnut Creek, CA; 7Biological and Materials Science Center, Sandia National Laboratories, Livermore, CA

http://downloads.jbei.org/data/MaxBin.html.

Project Goals: The Microbial Communities group at the Joint BioEnergy Institute translates activities found in the environment to applications in the biorefinery. The group develops bioinformatic tools to investigate microbial consortia that deconstruct plant biomass. These tools have broad application in understanding the metabolic potential of natural and engineered microbial communities.

Recovering individual genomes from metagenomic datasets by binning assembled sequence provides access to the metabolic potential of uncultivated microbial populations that may have important roles in natural and engineered ecosystems. We have developed a binning algorithm, MaxBin, which automates the binning of assembled metagenomic scaffolds using an expectation-maximization algorithm and is capable of binning individual metagenomes using tetranucleotide frequencies and read coverage of assembled sequences. MaxBin 2.0 was developed to bin genomes from multiple samples using differential read coverage and was compared to CONCOCT and MetaBAT, two recently described metagenomic binning algorithms that recover genomes using tetranucleotide frequencies and differential coverage. MaxBin also incorporates a reassembly function, which automatically retrieves the paired-end reads associated with the assembled sequences of binned genomes and reassembles them. MaxBin has been applied to recover genomes from microbial communities relevant to biomass deconstruction, including thermophilic microbial communities grown on biomass and halophilic communities found along a salinity gradient in San Francisco Bay.

Publications

1. “MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation- maximization algorithm” Wu, Y.W.; Tang, Y.H.; Tringe, S.G.; Simmons, B.A.; Singer, S.W. Microbiome, 2014, 2, 26.

This work was performed 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

142. The Unexpected Importance of Actinobacterial GH12 in Hydrolysis of Crystalline Cellulose

Jennifer Hiras1,2* (jhiras@lbl.gov), Yu-Wei Wu1,2, Susannah G. Tringe3, Carrie Nicora4, Joshua Aldrich4, Errol Robinson4, Jon Jacobs4, Blake A. Simmons1,5 and Steven W. Singer1,6

1Joint BioEnergy Institute, Emeryville, CA; 2Physical Biosciences Divisions, Lawrence Berkeley National Laboratory, Berkeley, CA; 3Joint Genome Institute, Walnut Creek, CA; 4Environmental Molecular Sciences Laboratory, Richland, WA; 5Biological and Materials Science Center, Sandia National Laboratories, Livermore, CA; 6Earth Sciences Divisions, Lawrence Berkeley National Laboratory, Berkeley, CA

Project Goals: The primary research goal of the JBEI Deconstruction Division is to discover and optimize methods by which biomass can be efficiently converted into fermentable sugars and aromatics. The specific goal of this project is to discover and optimize lignocellulolytic enzymes and enzyme mixtures for enhanced hydrolysis of pretreated biomass.

Production of fermentable sugars via enzymatic hydrolysis of complex plant polysaccharides is a subject of intense global interest. Here, we investigated compost enrichments as sources of novel thermophilic enzymatic mixtures, referred to as cocktails, for cellulose hydrolysis. Bacterial consortia were adapted to microcrystalline cellulose (MCC) at 60°C under aerobic conditions using Berkeley Green Waste (BGW) compost as the inoculum. 3,5-dinitrosalysilic acid (DNS) and p-nitrophenol (pNP) enzyme assays indicated enrichment supernatants had high levels of glycoside hydrolase activities. Isolated supernatants were then used to successfully saccharify MCC and 1-ethyl-3-methylimidazolium acetate ([C2C1Im][OAc]) pretreated switchgrass (ILSG). Together, these data demonstrate that BGW consortia are an excellent source of enzymes for development of enzymatic cocktails. In addition, supernatants were tolerant of ionic liquids, emerging chemicals for biomass pretreatment, indicating the enzyme in these supernatants may be useful in industrial-scale biofuel production. Microbial community compositional analysis revealed that Actinobacteria dominated early passages of two enrichment lineages. One lineage sustained Actinobacterial dominance over two-week intervals, while the community profile of a second lineage changed from predominantly Actinobacteria to a mixture of Bacteroidetes and Firmicutes. These data suggest a succession in primary and secondary cellulose degraders in these cultures that may be important in understanding the mechanism of cellulose hydrolysis by thermophilic bacterial populations. The supernatant from the Actinobacterial-dominated lineage was particularly effective at saccharification of crystalline cellulose. Comparative metagenomic and proteomic analyses revealed differential expression of glycoside hydrolases (GHs) from Actinobacteria were responsible for the changes in observed activities on biomass substrates. In particular, the abundance of GH12 from the Actinobacteria correlated with activity of the supernatants on microcrystalline cellulose. This result was unexpected because GH12 is commonly characterized from bacteria and fungi as an endoglucanase with negligible activity on crystalline substrates. Heterologous expression of the Actinobacterial GH12, along with GH6 and GH48 from the same organism (a close relative of Thermobispora bispora) demonstrated that it had high activity on CMC and pNPC substrates. Heterologous expression of these GHs (GH12, GH48, GH6, AA10) has provided purified enzymes to formulate an enzymatic mixture for high-temperature saccharification of biomass substrates with crystalline cellulose.

This work was performed 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

143. Droplet Microfluidics for Synthetic Biology

Philip C. Gach (pcgach@lbl.gov),1* Steve C.C. Shih,1 Jess Sustarich,1 Peter Kim,1 Daniel H. Arlow,1 Garima Goyal,1 Randy Drevland,1 Ken L. Sale,1 Nathan J. Hillson,1 Jay D. Keasling,1 Paul D. Adams1 and Anup K. Singh1

1Joint BioEnergy Institute, Emeryville, California

http://www.jbei.org/research/divisions/technology/microfluidic-assays/

Project Goals: Synthetic biology experiments require optimization of pathways consisting of many genes and other genetic elements and given the large number of alternatives available for each element, optimization of a pathway can require large number of experiments consuming prohibitively-expensive amounts of DNA and enzymes. Droplet microfluidics, because of its ability to process small volumes, presents a cost-effective solution for conducting high-throughput cloning and expression experiments. Unfortunately, these platforms are generally not designed in a robust nature with the feedback and temperature control necessary for screening large combinatorial libraries. Towards these ends, we are developing devices that integrate droplet microfluidics, digital microfluidics and peltier heaters/coolers. These microfluidic systems will be utilized for automating all critical steps of synthetic biology including DNA assembly, transformation, culture and expression. We will utilize these technologies for screening a variety of biological constructs relevant to the mission statement of the Joint BioEnergy Institute.

Rapid and inexpensive synthesis of plasmid-sized DNA would dramatically accelerate synthetic biology research. We are developing an integrated system for rapid DNA synthesis by iterative construction of single DNA molecules in a microfluidic device (Figure 1A). As reagents are dispensed to the reaction chamber in nanoliter quantities, reaction progress is monitored by single-molecule fluorescence on a custom-built total internal reflection fluorescence (TIRF) microscope. Once a single DNA molecule with the desired sequence is constructed, it is amplified for use in synthetic biology applications.

Synthetic biology has become an approach to understand and to manipulate biological systems in bacteria and in yeast for biofuel applications. However, given this extremely challenging goal, the biological design cycle (specify-design-construct- verify) is currently very slow, expensive, and laborious. To expedite this process, significant efforts are needed to develop enabling technologies for rapid biological engineering (e.g. an automated platform for constructing designs and verifying the constructs in a host organism). A microfluidic device has been developed which allows DNA assembly and verification (Figure 1B). This system was utilized to assemble 16 different plasmids using three different DNA assembly methods: Golden Gate, Gibson, and yeast assembly.

Transformation of exogenous DNA into bacterial cells is a powerful technique for genetics studies. A highly integrated device based on digital microfluidics and strategically positioned peltiers allows all transformation procedures to be carried out on a single device (Figure 1C). Three peltiers were positioned below the device to provide the 4oC, 30oC, 37oC and 42oC temperatures employed during chemical heat shock. The digital microfluidic device allows automation of all critical steps of transformation and culture including plasmid addition, transformation by heat-shock, addition of selection medium, culture and protein expression. Flexibility in device operating parameters allows transformation of various DNA plasmids into many different cell types.

A platform that is a combination of classic-droplet- microfluidics and digital microfluidics is being developed (Figure 1D). The grouping of these two techniques allows us to take advantage of both platforms' strengths, negating their pitfalls. To demonstrate the device's ability, we are conducting a study for the directed evolution of enzyme ionic liquid tolerance. The device will be able to combine reagents, incubate, and screen for enzyme activity at high throughput. Only droplets with enzymes that

can efficiently generate glucose in the presence of ionic liquids will be collected for further mutagenesis of the isolated DNA.

We have developed devices that integrate droplet microfluidics, digital microfluidics and peltier heaters/coolers. These microfluidic systems have been utilized for automating all critical steps of synthetic biology including DNA assembly, transformation, culture and expression. Future plans are to expand the applications of this technology to other projects that are in need of a highly-controlled yet high-throughput platform, such as gene design and protein pathway engineering.

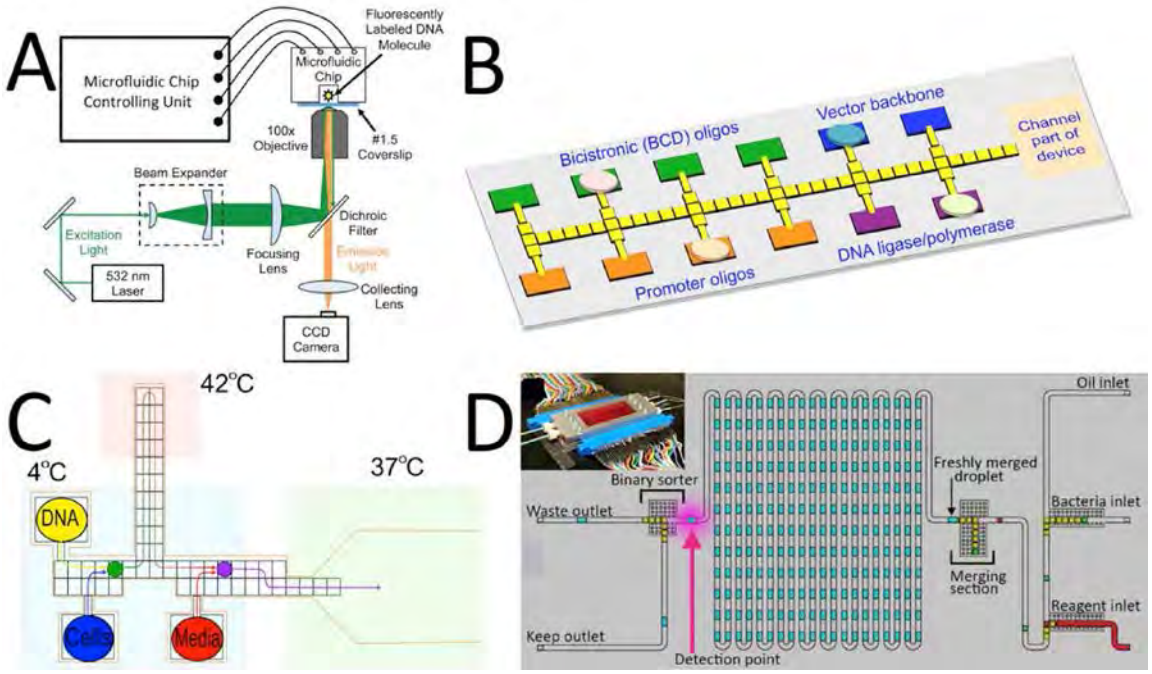


Figure 1. A) Schematic of microfluidic chip based DNA synthesis system on TIRF microscope for fluorescent single molecule detection. B) Schematic of microfluidic device for DNA assembly and verification C) Schematic of a DMF device for cell transformation. D) Photo and schematic of the architecture and function of a chip for directed evolution of ionic liquid tolerance.

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. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04- 94AL85000.

144. Ionic Liquid Pretreatment: State of Technology and Future Perspective

Seema Singh*, Jian Sun, Tanmoy Dutta, Ramakrishnan Parthasarathi, Jian Shi, Blake A. Simmons

Presenter email: ssingh@lbl.gov, seesing@sandia.gov

Abstract: Lignocellulosic biomass is an abundant renewable feedstock with great geological diversity and availability. There is a great opportunity for the production of various commodities such as biofuels, chemicals and biomaterials. ILs have proven the extraordinary potential in facilitating the fractionation and separation of biomass components. The progress made in last few years have demonstrated that ILs can be successfully used as pretreatment or bioprocessing medium in a biorefinery to obtain cellulose, hemicellulose and lignin fractions from a variety of biomass feedstocks or mixtures with purity and efficiency equal or superior to the currently employed pretreatment methods for the second generation biofuels and biochemical production. The chemistry during IL pretreatment largely depends on the selection of ILs, the processing conditions and recovery methods and ties closely to how the biomass being altered during the process and the pretreatment efficacy. The IL pretreatment technology is evolving very fast and this rapid development is augmented by the development of new ILs, process development, and nature of ILs being a designer solvents. Despite the enormous potential and growth, challenges have to be addressed to make IL pretreatment an economically viable technology in the biorefinery context. This presentation will cover the state of IL pretreatment technology and briefly provide outlook on several different scenarios currently available for feedstock to fuels using IL pretreatment technology via integrated efforts and progress made in different divisions at JBEI.

145. Interactions between Ascomycete fungi and Actinomycete bacteria during litter decomposition under an N gradient

John Dunbar1, La Verne Gallegos-Graves1* (Laverne@lanl.gov), Andrea Porras-Alfaro2, Chris Yeager1, Rebecca Mueller1, Cedar Hesse1, Cheryl R Kuske1

1 Los Alamos National Laboratory, Los Alamos, NM
2 Western Illinois University, Macomb, IL

Project Goals: Studies of ecosystem response to N deposition have reported declines in plant litter decomposition rates and declines in the relative abundance of Basidiomycota fungi, which are root-associated plant symbionts and plant biomass decomposers with key roles in cycling N in surface soils and acquiring N for plant growth. When Basidiomycota relative abundance declines, the relative abundance of Ascomycota fungi often increases. Thus, other fungi and bacteria can become more important in plant litter decomposition. Our work aims to determine how N gradients influence the activity and interactions of two broad decomposer groups---the Ascomycete fungi and Actinobacteria-that are known to harbor lignocellulosic decomposition traits and are potential key players in litter decomposition.

Atmospheric N deposition has increased N availability in natural ecosystems by 2-fold, on average, but as much as 100-fold in specific locations. Increased N availability has been widely reported to reduce decomposition of plant litter, at least on short time-scales. In conventional models, increased N availability alters plant-Basidiomycota interactions such that hyphal networks decline and Basidiomycota relative abundance declines. In theory, other decomposer groups can fill the available niche space. Ascomycota fungi and Actinobacteria in particular are expected to play larger roles in litter decomposition when Basidiomycota fungi decline. While increases in relative abundance of these two groups have been documented in some studies, other studies suggest that increased N availability might inhibit Ascomycete activity.

We are monitoring the collective activities of mixed communities of Ascomycota and Actinobacteria in time course experiments, where defined mixtures of five fungal and five bacterial genera, decompose plant litter (arid-land grasses) in sand microcosms under five nitrogen treatments. Given the difficulty of manipulating natural communities, defined mixtures provide the best approach to decipher functional responses, interactions, inherent biological barriers, and relevant mechanistic phenomena. To identify general patterns, instead of the eccentric response of one or two specific mixtures, we are documenting trends across many independent mixtures. Fungal and bacterial isolates for the mixtures were obtained from arid grassland sites and a pine forest field experiment where N application was an experimental variable. Measurements of the defined mixtures include initial biomass, CO2 evolution over a 30- 50 day time-course, initial and final community composition (rRNA surveys of fungal and bacterial composition), and metatranscriptome analyses.

Preliminary results show evidence of composition-dependent behavior. Nonetheless, strong patterns are emerging across fungal-bacterial mixtures that document the general interaction between fungi and bacteria in these communities and the general contribution of N availability to Ascomycete-dominated decomposition of arid-land grass litter.

Funding statement: The information in this poster was supported by the U.S. Department of Energy Biological System Science Division, through a Science Focus Area Grant (2014LANLF260).

146. Metatranscriptomes to explore the effects of nitrogen enrichment on forest soil microbial communities

Cedar Hesse1, Momchilo Vuyisich1, La Verne Gallegos-Graves1, Rebecca Mueller1, Donald R. Zak2, and Cheryl R. Kuske1* (Kuske@lanl.gov)

1 Los Alamos National Laboratory, Los Alamos, NM,
2 University of Michigan, Ann Arbor, MI

Project Goals: A diverse community of bacteria and fungi mediate plant litter decay in forests, within which they regulate the cycling and storage of C and N in soil. Anthropogenic perturbation of nitrogen balance alters the rate of forest litter decomposition and potentially influences overall soil community metabolism. However, the microbial mechanisms mediating a reduction in decay are largely unknown. In this study we are using environmental metatranscriptome techniques to explore the collective metabolism of natural microbial communities in forest floor material and soils exposed to experimental nitrogen enrichment. This poster focuses on the development of metatranscriptome methods that capture both eukaryotic (polyA) and prokaryotic mRNA and emphasizes the comparative analysis of carbohydrate active enzymes (CAZymes) within datasets.

Prior studies using soil metatranscriptomes suffered from low sequencing coverage or incomplete community representation (eukaryotic only). We developed a metatranscriptome approach using high-throughput Illumina sequencing that significantly enriched for mRNA in forest floor and soil RNA pools, and that included both eukaryote (primarily fungal) and bacterial transcripts. Our analyses included functional annotation of sequencing reads to PFAM, CAZymes, KEGG, FOAM, and Gene Ontology databases, allowing us to thoroughly explore the genomic basis of biogeochemical pathways mediated by soil fungi and bacteria. Metatranscriptome data was collected from three distinct biomes in North America: (1) soil from a pine-dominated temperate forest (North Carolina, USA), (2) decomposing leaf litter from a maple-dominated hardwood forest (Michigan, USA), and (3) a Cyanobacteria-dominated biological soil crust from a semi-arid grassland (Utah, USA). The work presented here focuses primarily on the forest biomes. Metatranscriptome expression profiles were correlated with additional biotic and abiotic measurements from the same samples including fungal and bacterial taxonomic abundance (rDNA), total bacterial and fungal abundance (qPCR), and soil chemistry. Results show that our mRNA isolation methods sufficiently deplete the rRNA fraction of total RNA and that subsequent metatranscriptome sequencing captures eukaryotic, bacterial, archaeal, and viral transcripts. KEGG analysis indicates we have covered the core metabolic pathways expected in microbial systems although the largest shifts in community metabolism among treatments are exhibited by the CAZyme analysis. In maple leaf litter, nitrogen enrichment suppresses a large suite of eukaryotic, primarily fungal, CAZymes while enriching unique sets of CAZymes at each site. In the pine forest, community expression was strongly differentiated by soil horizon and less so by nitrogen enrichment. Comparative analyses across all metatranscriptome datasets show distinct partitioning by biome and soil horizon correlating with broad-scale taxonomic shifts in the total microbial community.

Kuske CR, CN Hesse, JF Challacombe, D Cullen, JR Herr, RC Mueller, A Tsang, R Vilgalys (2015) Prospects and challenges for fungal metatranscriptomics of complex communities. Fungal Ecol doi:10.1016/j.funeco.2014.12.005

Funding statement: The information in this poster was supported by the U.S. Department of Energy Biological System Science Division, through a Science Focus Area Grant (2014LANLF260).

147. Microhabitat-dependent responses of bacterial and fungal communities to experimental nitrogen additions in arid ecosystems

Rebecca C Mueller1* (rmueller@lanl.gov), La Verne Gallegos-Graves1, Jayne Belnap2, Sasha C Reed2, Robert Sinsabaugh3, Noelle Martinez3, Cheryl R Kuske1

1 Los Alamos National Laboratory, Los Alamos, NM; 2 United States Geological Survey, Moab, UT; 3 University of New Mexico, Albuquerque, NM

Project Goals: Anthropogenic N deposition is a chronic and increasing condition in temperate regions that may strongly influence C cycling dynamics. One major theme of our Science Focus Area is to determine the influence of chronic N deposition on microbial C cycling processes in two major biomes of Earth's temperate regions, forests and arid grass/shrublands. In both biomes fungal and bacterial biomass is concentrated in shallow surface soil strata where C and N cycling are major processes. The goals of this project are to (a) correlate the resident fungal and bacterial communities, enzyme activities, and local geochemistry across shallow strata of strongly stratified forest soils and biocrust dominated soils of arid grass/shrub lands, (b) determine the impacts of chronic N amendment across multiple forest and arid grass/shrubland biomes using a suite of long-term field experiments, and (c) compare the ability of phylogenetic rRNA gene surveys, soil enzyme assays, and metatranscriptome surveys to detect shifts in community structure and concomitant changes in C cycling processes in response to altered N conditions. Achieving these goals will provide an understanding of the active and responsive components of arid land soils that contribute to carbon cycling, their collective responses to environmental change, and development of efficient molecular tools for broad-scale soil monitoring. This poster focuses on our recent results from two arid land N deposition experiments.

Research on microbial communities, both in terms of surveying resident communities and quantifying response to environmental perturbations, has largely ignored arid ecosystems.

However, drylands cover ~40% of the terrestrial surface of the earth, and are expanding in many regions as a result of desertification. Historical inputs of nitrogen to drylands are generally low, but are increasing due to human activities. The climatic conditions and nutrient cycles in these ecosystems differ dramatically from well-studied mesic systems; they are characterized by highly variable environmental conditions due to sporadic precipitation events, extreme temperature fluctuations and UV radiation stresses, as well as low productivity, patchy distributions of biotic resources (microhabitats) and atypical sources of nutrient inputs. Given these unique characteristics, it is unlikely that responses seen in mesic forests and grasslands will apply in drylands.

We are using high throughput sequencing methods to target the soil bacterial and fungal communities, which together are responsible for the majority of nutrient cycles, including the N and C cycles. These studies were conducted at experimental N deposition field experiments located in an arid shrubland (Nevada) and an arid grassland (Utah) in the Southwestern US. We are applying multivariate statistical approaches to examine shifts in the phylogenetic structure of these communities. In an effort to predict the functional effects of community shifts, we are simultaneously measuring relative microbial biomass, soil chemistry and enzyme assays to link specific environmental shifts to community responses and their functional consequences.

Community surveys indicate that microbial community responses differed between the two sites. In the arid grassland, we found no evidence for community changes in either bacteria or fungi. In the shrubland, however, community composition results found divergent responses of the bacterial and fungal communities, shifts that are dependent upon microhabitat and soil depth. Unlike responses commonly

seen in temperate systems, we found limited response in the fungal community, but strong shifts in richness, community composition and overall relatedness in the bacterial community. We also documented increased fungal and bacterial biomass (qPCR), elevated photosynthetic activity in biocrusts, and shifts toward enzymes linked to phosphorus acquisition in the N plots. There is thus evidence for shifts toward fungal-dominated nutrient cycling in response to nitrogen additions. By combining sequence-based community analyses with soil and enzyme activity measures, we may identify key responsive community members with relevance as indicators of community change and with utility for modeling soil processes.

1. Reed SC, FT Maestre, R Ochoa-Hueso, CR Kuske, M Oliver, A Darrouzet-Nardi, B Darby, et al. (2015) Biocrusts in the context of global change. In: Biological Soil Crusts: Structure, Function and Management, 3rd edition. Ed. J Belnap. (in press)
2. Maier S, L Muggia, CR Kuske, M Grube (2015) Fungi and bacteria in biological soil crusts. In: Biological Soil Crusts: Structure, Function and Management, 3rd edition. Ed. J Belnap. (in press)
3. Steven B, C Hesse, LV Gallegos-Graves, CR Kuske (2015) Fungal diversity in biological soil crusts of the Colorado Plateau. Proc 12th Biennial Conf Science Management Colorado Plateau (in press)
4. Steven B, LV Gallegos-Graves, C Yeager, J Belnap, CR Kuske (2014) Common and distinguishing features of the bacterial and fungal communities in biological soil crusts and shrub root zone soils. Soil Biol Biochem 69:302-312.
5. Sinsabaugh RL, J Belnap, SG Findlay, JJ Follstad Shah, BH Hill, KA Kuehn, CR Kuske, M Litvak, NG Martinez, DL Moorhead, DD Warnock (2014) Extracellular enzyme kinetics scale with resource availability. Biogeochemistry 121(2):287-304. doi 10.1007/s10533-014-0030-y
6. Lipson DA, CR Kuske, LV Gallegos-Graves, WC Oechel (2014) Elevated atmospheric CO2 stimulates soil fungal diversity through increased fine root production in a semiarid shrubland ecosystem. Glob Change Biol 20: 2555–2565 doi:10.1111/gcb.12609.
7. Steven B, LV Gallegos-Graves, J Belnap, CR Kuske (2013) Dryland soil bacterial communities display spatial biogeographic patterns associated with soil depth and soil parent material. FEMS Microbiol Ecol 86:101-113.
8. Porras-Alfaro A, K-L Liu, CR Kuske, G Xie (2013) From genus to phylum: large-subunit and internal transcribed spacer rRNA operon regions show similar classification accuracies influenced by database composition. Appl Environ Microbiol, doi:10.1128/AEM.o2894-13.

Funding statement: The information in this poster was supported by the U.S. Department of Energy Biological System Science Division, through a Science Focus Area Grant (2014LANLF260).

148. The Passalid Beetle (Odontotaenius disjunctus) Microbiome and its Metabolic Potential for Lignocellulose Degradation

Javier A. Ceja-Navarro (JCNavarro@lbl.gov) 1*, Ulas Karaoz1, Zhao Hao1, Markus Bill1, Mark Conrad1 Richard A. White III2, Mary S. Lipton2, Joshua N. Adkins2, Timothy Filley3, Meredith Blackwell4, Jennifer Pett-Ridge5, Eoin L. Brodie1

1Lawrence Berkeley National Laboratory, Berkeley, CA; 2Pacific Northwest National Laboratory, Richland, WA; 3Purdue University, West Lafayette, IN; 4Louisiana State University, Baton Rouge, LA. 5Lawrence Livermore National Laboratory, Livermore, CA.

https://envmicro.wordpress.com/researchprojects/passalid-beetles-natures-efficient-lignocellulosic-biorefineries/

Project Goals: Our research aims to develop an integrated analysis of energy flow in complex microbial communities by combining multi-scale approaches including biogeochemical, stable isotope probing, metagenomic/transcriptomic, proteomic and computational analyses, to understand nutrient cycling and biofuel production in complex microbial communities. A comprehensive understanding of such communities may help in the development of efficient, industrial-scale processes for lignocellulose degradation. Our ultimate goal is the development of multi-scale models that can predict ecological and biochemical relationships within multi- trophic microbial systems.

Insects are the most abundant group of metazoans on Earth and have evolved to exploit a diverse array of environments, an ability that may be facilitated and enhanced by their gut-associated microbes. The wood-feeding beetle Odontoatenius disjunctus displays morphologically differentiated gut regions where stratified microbial communities degrade lignocellulosic materials and fix nitrogen. Our goal is to characterize the genomic potential stored in the microbiome of the passalid beetle for the optimization of lignocellulosic-dependent energy production processes.

Metagenomes were prepared from four gut regions of replicate beetles, sequenced, co- assembled and annotated. The ability to transform lignocellulosic materials by the passalid beetle was tested by Fourier Transform Infrared Spectroscopy and by measuring the fermentation products of plant polymer decomposition (H2 and CH4) with gas chromatography-isotope ratio mass spectrometry (GC-IRMS). The transformation of lignin after its passage through the gut was determined by 13C-labeled tetramethylammonium hydroxide thermochemolysis. Proteomics and metabolomics were applied to detect expressed proteins and produced metabolites in each gut region.

After passage through beetles, oak wood showed cellulose and hemicellulose hydrolysis and lignin side chain oxidation. C and H stable isotope fractionation indicated that hydrogenotrophic methanogenesis dominates methane production the in beetle. These processes are mediated by a diverse range of bacteria, archaea and fungi in the metagenomes. Cellulose, starch, and xylan degradation genes were particularly abundant in the midgut (MG) and posterior hindgut (PHG). Genes involved in hydrogenotrophic production of methane and nitrogen cycling were more abundant in the anterior hindgut, confirming our previous phylogenetic and nitrogen fixing studies of compartmentalization in the passalid beetle gut1. A filtered isolate database and predicted protein sequences from the metagenomes were used to search peptide spectra for proteome reconstruction. Proteomics supported metagenome observations, detecting expression of carbohydrate active enzymes principally in the MG and PHG, and nitrogenases and the enzymes involved in methane production in the AHG. Co-assembly and binning identified linked functions within microbial genomes including organisms with cellulosomes and a combined potential for cellulose, xylan and starch binding and hydrolysis.

Our multi-scale approach demonstrates that the passalid beetle harbors and expresses the functional potential to deconstruct lignocellulosic materials and produce H2, CH4 and potentially other biofuels. By studying such highly spatially and biochemically evolved polymer deconstruction and fermentation systems we hope to identify design criteria for improved lignocellulosic fuel production processes.

References

1. Ceja-Navarro, J. A. et al. Compartmentalized microbial composition, oxygen gradients and nitrogen fixation in the gut of Odontotaenius disjunctus. ISME J. 8, 6–18 (2014).

This research was supported by the LLNL Biofuels Scientific Focus Area, funded by the Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under FWP SCW1039. Work was performed at Lawrence Berkeley National Laboratory under U.S. Department of Energy Contract No. DE-AC02-05CH11231. Part of this work was performed at the Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344 and at Pacific Northwest National Lab supported by the OBER/GSP under the Pan-Omics project. We also acknowledge EMSL support to E. Brodie through proposal #47988.

149. Protection from Oxidative Stress in Algal-Bacterial Co-Cultures

Danielle YoungSmith1, Michelle R. Brann2, Xavier Mayali3, Oana E. Marcu4,5*
(oana.marcu@nasa.gov)

1Barnard College, New York, NY; 2Wellesley College, Wellesley, MA; 3Lawrence Livermore National Laboratory, Livermore CA; 4Carl Sagan Center, SETI Institute, Mountain View, CA; and 5NASA Ames Research Center, Moffett Field, CA

Project Goals: The LLNL Biofuels SFA seeks to support robust and sustainable microalgae fuel production through a systems biology understanding of algal-bacterial interactions. We hypothesize that by understanding the factors that control cellular physiology and biogeochemical fluxes in and out of algal cells, particularly through the phycosphere, we can advance the efficiency and reliability of algal biofuel production. Our research includes studies of probiotic traits of phycosphere- associated bacteria, systems biology studies of model algae, and genome-enabled metabolic modeling to predict the interspecies exchanges that promote algal growth, lipid production and healthy co- cultures. Our overall goal is to develop a comprehensive understanding of complex microbial communities needed to advance the use of biological properties for practical energy production.

As part of a new SFA project focused on algal-bacterial interactions, the LLNL Biofuels SFA has recently initiated collaborations with O. Marcu's group at the SETI Institute. The Marcu lab is focused on laboratory co-cultures of the green alga Chlamydomonas reinhardtii grown in co-culture with bacterial strains that enhance the growth and lipid production of the microalga. The bacteria were isolated from the topsoil of Mojave Desert and are pre-adapted to the oxidative stress imposed by the surface exposure to UV radiation and desiccation. In the presence of these bacteria, the Chlamydomonas cultures show an initial slight increase in the levels of reactive oxygen species (ROS), but increased growth and photosynthetic efficiency. Remarkably, in the presence of a secondary stress induced by heat, the ROS levels in algae decrease as compared to controls, suggesting that the bacteria provide priming and protection against oxidative stress. Bacteria also protect the algae from hydrogen peroxide-induced stress, most likely through the activity of extracellular bacterial catalase, and from the stress induced by the removal of copper. Monitoring of ROS levels in the intracellular compartment versus supernatant suggests a role for the extracellular matrix in buffering their toxicity. Co-cultures that show increased lipid production can be reconstructed in laboratory conditions.

Currently the work focuses on the multiple bacterial consortia in co-cultures and their physical association with the algae, gene expression under stress conditions, the lipid metabolism and metal transport, and the role of the algal extracellular matrix in mediating the association and response to stress, for short- and long-term growth of algal cultures with enhanced productivity.

This material is based upon work supported partly by the National Aeronautics and Space Administration Science Mission Directorate, Astrobiology: Exobiology and Evolutionary Biology Program under Agreement No. NNX10AI34A to O.M, and by NASA OSSI internships to DYS and MRB. Future collaboration is to be supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under the LLNL Biofuels SFA, FWP SCW1039. Work at LLNL is performed under the auspices of the U.S. Department of Energy under Contract DE-AC52-07NA27344.

150. Diel Metagenomics and Metatranscriptomics of Elkhorn Slough Hypersaline Microbial Mat

Jackson Z. Lee1§, Ulas Karaoz2* (ukaraoz@lbl.gov), Angela Detweiler1, Craig Everroad1, Leslie Prufert-Bebout1, Rhona Stuart3, Mary Lipton4, Eoin L. Brodie2, Peter Weber3, Brad Bebout1, Jennifer Pett-Ridge3

1NASA Ames Research Center, Mountain View, CA. 2Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA. 3Lawrence Livermore National Laboratory, Livermore, CA, 4Pacific Northwest National Laboratory, Richland, WA. §Current Affiliation: MIT Lincoln Laboratory, MA.

Project Goals: The LLNL Biofuel SFA investigates systems biology of complex microbial communities relevant to bioenergy production. To understand nutrient cycling and potential biofuel production in complex microbial communities we employ an integrated analysis of energy flow using multi-scale approaches including biogeochemical, stable isotope probing, metagenomic/transcriptomic, proteomic/metabolomic and computational analyses. Our ultimate goal is the development of multi-scale models that can predict ecological and biochemical relationships within multi-trophic microbial systems.

Microbial mats are amongst the most diverse microbial ecosystems on Earth, inhabiting some of the most inclement environments known, including hypersaline, dry, hot, cold, nutrient poor, and high UV environments. Photosynthetic microbial mats found in intertidal environments are stratified microbial communities with anoxic conditions at night, generating significant amounts of H2 and organic acids. The high microbial diversity of microbial mats makes for a highly complex series of ecological interactions. To address this challenge, we are using a combination of metagenomics, metatranscriptomics, metaproteomics, iTags and culture-based simplified microbial mats to study biogeochemical cycling (H2 production, N2 fixation, and fermentation) in mats collected from Elkhorn Slough, Monterey Bay, California.

To understand the variation in gene expression associated with the daytime oxygenic phototrophic and nighttime fermentation regimes in hypersaline microbial mats, a contiguous mat piece was sampled at regular intervals over a 24-hour diel period. Additionally, to understand the impact of sulfate reduction on biohydrogen consumption, molybdate was added as an inhibitor to a parallel experiment. Four metagenome and 12 metatranscriptome Illumina HiSeq lanes were completed for samples collected from day / night, and control / molybdate experiments.

Our preliminary examination of gene expression in midday versus midnight samples (mapped using bowtie2 to reference genomes) has revealed several notable features, particularly relevant to the dominant mat-building cyanobacterium Microcoleus chthonoplastes.

M. chthonoplastes expresses several pathways for nitrogen scavenging, including nitrogen fixation. Reads mapped to M. chthonoplastes indicate expression of two starch storage and utilization pathways, a starch-trehalose-maltose-glucose pathway, and a UDP-glucose- cellulose-β-1,4 glucan-glucose pathway. The overall trend of gene expression was primarily light driven up-regulation followed by down-regulation in the dark; much of the remaining expression profile appears to be constitutive. Metaproteome analyses, conducted in collaboration with PNNL's Pan-Omics project (mapped using co-assembled metagenome), indicate upregulation of Chloroflexi-assigned proteins in the dark and upregulation Cyanobacteria-assigned proteins in the light

Co-assembly of quality-controlled reads from 4 metagenomes was performed using Ray Meta with progressively smaller K-mer sizes, with bins identified and filtered using principal component analysis of

coverages from all libraries and a %GC filter, followed by reassembly of the remaining co-assembly reads and binned reads. A total of 20 near-complete (>80%) and an additional 50 minor genomic bins have been identified. Despite having relatively similar abundance profiles in each metagenome, this binning approach was able to distinctly resolve bins from dominant taxa, as well as sulfate reducing bacteria that are critical to our understanding of molybdate inhibition effects. Bins generated from this iterative assembly process are being used for downstream mapping of transcriptomic reads as well as isolation efforts for Cyanobacteria- associated bacteria.

Publications

1. Lee, J.Z., L. Burow, D. Woebken, R.C. Everroad, M.D. Kubo, A. Spormann, P.K. Weber, J. Pett-Ridge, B.M. Bebout, T.M. Hoehler. (2014) Fermentation couples Chloroflexi and sulfate- reducing bacteria to Cyanobacteria in hypersaline microbial mats. Frontiers in Microbial Physiology and Metabolism. 5:61.

This research was supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under the LLNL Biofuels SFA, FWP SCW1039. Work at LLNL was performed under the auspices of the U.S. Department of Energy under Contract DE-AC52-07NA27344 and at Pacific Northwest National Lab supported by the OBER/GSP under the Pan-Omics project. Work at LBNL was performed under the auspices of the U.S. Department of Energy under Contract DE-AC02-05CH11231.

151. Identifying Microbial Resource Use with Chip-SIP

Erin Nuccio1* (nuccio1@llnl.gov), Xavier Mayali,1 Steven Blazewicz,1 Peter Weber,1 Javier Ceja-Navarro,2 Mary K. Firestone3, Bruce Hungate,4 Eoin Brodie,2 Jennifer Pett-Ridge1

1Lawrence Livermore National Laboratory, Livermore, CA; 2Lawrence Berkeley National Laboratory, Berkeley, CA, 3University of California, Berkeley, CA, 4 Northern Arizona University, Flagstaff, AZ

Project Goals: The LLNL Biofuel SFA investigates systems biology of complex microbial communities relevant to bioenergy production. To understand nutrient cycling and potential biofuel production in complex microbial communities we employ an integrated analysis of energy flow using multi-scale approaches including biogeochemical, stable isotope probing, metagenomic/transcriptomic, proteomic/metabolomic and computational analyses. Our ultimate goal is the development of multi-scale models that can predict ecological and biochemical relationships within multi-trophic microbial systems.

One of the key challenges in microbial ecology is determining the ecophysiology of microorganisms in in situ complex communities. Stable isotope probing is a technique that identifies taxa that consume specific substrates labeled with rare isotopes in complex assemblages. As part of the LLNL Biofuel SFA, our group developed Chip-SIP, a method which combines NanoSIMS isotope imaging with phylogenetic microarrays to determine the isotopic enrichment of specific populations in the microbial community1. Recently, we have made a number of recent advances to facilitate the application of Chip-SIP by the greater scientific community. First, we have developed an automated probe design pipeline to accelerate the design of custom Chip-SIP arrays from next-generation sequencing datasets. This process allows researchers to target populations that are specific to their studies. The pipeline clusters taxa into OTUs, generates OTU-specific probes, and then validates the specificity of those probes in-silico using the SILVA database. Second, we have expanded the repertoire of isotopes that can be detected simultaneously by Chip-SIP to include 18O, in addition to 13C and 15N. Isotopically-labeled water (H218O) is an ideal substrate to identify active populations in many different environments, as all known organisms consume water. Third, we've developed a tool to rapidly detect isotopic enrichment (13C, 15N, 18O) in picogram quantities of RNA using NanoSIMS, which can quickly identify samples that are suitable for further analysis. Finally, we will report on several recent applications of Chip-SIP: to detect substrate preference in the rhizosphere, and cellulose-degrading communities in the beetle hindgut. These advances highlight the utility of this method to identify specific metabolisms within complex microbial communities, and should facilitate use of the Chip-SIP approach by other research groups working independently or in collaboration with the LLNL Biofuels SFA.

Publications

1. Mayali, X., P.K. Weber, E.L. Brodie, S. Mabery, P. D. Hoeprich, J. Pett-Ridge. (2012). High-throughput isotopic analysis of RNA microarrays to quantify microbial resource use. ISME Journal 6: 1210-1221.

This research was supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under the LLNL Biofuels SFA, FWP SCW1039 and collaborative research projects on microbial carbon cycling (with B. Hungate---FWP SCW 1424 and M.K. Firestone---FWP SCW 1421). Work at LLNL was performed under the auspices of the U.S. Department of Energy under Contract DE-AC52-07NA27344. Work at LBNL was performed under the auspices of the U.S. Department of Energy under Contract DE-AC02-05CH11231.

152. NanoSIMS Isotope Imaging to Investigate Algal-Bacterial Interactions in Biofuel-Producing Communities

Peter K. Weber1* (weber21@llnl.gov), Luz E. de-Bashan2,3, Yoav Bashan2,3, Nestor Arandia4, Brad Bebout5, Jennifer Pett-Ridge1, and Xavier Mayali1

1Lawrence Livermore National Laboratory, Livermore CA; 2Auburn University, Auburn AL; 3The Northwestern Center for Biological Research, Mexico, 4University of Gijon, Spain; and 5NASA Ames Research Center, Mountain View, CA;

Project Goals: The LLNL Biofuels SFA seeks to support robust and sustainable microalgae fuel production through a systems biology understanding of algal- bacterial interactions. We hypothesize that by understanding the factors that control cellular physiology and biogeochemical fluxes in and out of algal cells, particularly through the phycosphere, we can advance the efficiency and reliability of algal biofuel production. Our research includes studies of probiotic traits of phycosphere-associated bacteria, systems biology studies of model algae, and genome-enabled metabolic modeling to predict the interspecies exchanges that promote algal growth, lipid production and healthy co-cultures. Our overall goal is to develop a comprehensive understanding of complex microbial communities needed to advance the use of biological properties for practical energy production.

To better understand the mechanisms by which probiotic bacteria promote growth, health, and/or lipid-producing capabilities of unicellular microalgae, we are currently investigating algal-bacterial interactions in two experimental systems: 1) simple co-cultures incubated in the laboratory consisting of one bacterium and one algal species, and 2) complex, natural communities (man-made algal ponds or natural aquatic ecosystems) with a diverse array of bacteria, algae, and other organisms. One of the key approaches we have used, NanoSIP (stable isotope probing followed by NanoSIMS isotope imaging analysis), allows us to monitor the transfer of metabolites specific to one partner in algal-bacterial consortia.

To examine exchange of metabolites in a highly controlled algal-bacterial system, we conducted a co-culture experiment using the microalga Chlorella vulgaris and the plant growth-promoting bacterium Azospirillum brasilense 1 and investigated C and N transfers between them using NanoSIP. The two isolates were grown separately, labeled with 13C and/or 15N, and then co-incubated inside alginate beads for four days. We used the NanoSIMS 50 at LLNL to examine the transfer of metabolites from one partner to the other. Labeling of bacteria with heavy isotopes provided the unequivocal identification of bacteria attached to algal cells, the former being often hidden due to their small sizes and altered shapes when attached. Using this approach, we have demonstrated the reciprocal transfer of carbon and nitrogen in this partnership. Our results show that physical attachment between algae and bacteria resulted in higher levels of transfer in most cases, but attachment was not necessary for transfer to occur.

In our first investigations of algal-bacteria symbiosis in a natural system, we have examined the cell-specific metabolism of microalgae and bacteria collected from a natural springtime bloom in the coastal Eastern Atlantic. Mixed communities were incubated for 12 hours in the presence of 13C-bicarbonate and 15N leucine, the former serving as a marker for photosynthetic carbon fixation and the latter for heterotrophic bacterial carbon production. Again we used NanoSIP, this time to measure autotrophic and heterotrophic activity. This approach allows us to compare cell-specific metabolism of algal and bacterial populations, and also enabled us to identify algal and bacterial cells attached to one another, which appears to be a relatively common phenomenon. NanoSIP data for over 4,000 cells indicate that mixotrophy was uncommon in the bloom samples analyzed. NanoSIP was able to detect increased bacterial metabolism in response to increases in temperature. NanoSIP enables cell-specific

measurements of activity for both algae and bacteria and allows the quantification of C and N transfer between cells, providing valuable insight into symbiotic relationships that occur in constructed as well as natural ecosystems.

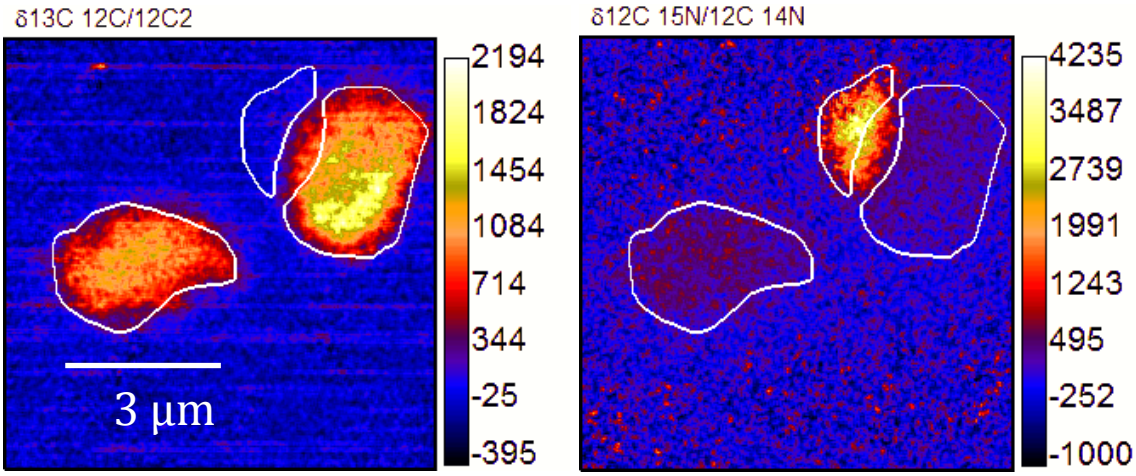


Figure 1. NanoSIP images showing autotrophy and heterotrophy in a mixed assemblage. The d13C image (left) shows 2 algal cells that took up the 13CO2 tracer, indicating autotrophy. The d15N image (right) shows a bacterium (attached to one of the algal cells) that incorporated the 15N-leucine tracer, indicating heterotrophy.

Publications

1. Gonzalez, L. E. & Bashan, Y. Increased Growth of the Microalga Chlorella vulgaris when Coimmobilized and Cocultured in Alginate Beads with the Plant-Growth- Promoting Bacterium Azospirillum brasilense. Appl. Environ. Microbiol. 66, 1527-1531 (2000).

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.

153. Systems Biology of Autotrophic-Heterotrophic Symbionts for Bioenergy

Tyler Huelsman1, Cristal Zuniga1* (crzuniga@eng.ucsd.edu), Jennifer Levering1, Michael Guarnieri2, George Oyler3, Michael Betenbaugh3, and Karsten Zengler1

1University of California, San Diego; 2National Renewable Energy Laboratory; 3Johns Hopkins University

Project Goals: The goal of this proposal is to combine autotrophs and heterotrophs as a novel sustainable symbiotic platform for the production of biofuels. Genome-scale reconstructions of autotrophic-heterotrophic co-cultures will be deployed to decipher and predict metabolic interactions and furthermore to optimize system performance for production of desired biofuels products.

Reconstructions are biochemically, genetically and genomically structured knowledge-bases that contain information of reaction stoichiometry, reaction reversibility, and the association between genes, proteins and reactions. Currently genome-scale metabolic reconstructions are available for a number of heterotrophs appropriate for biofuels. In this project, a reconstruction for the autotroph, Chlorella vulgaris has been generated in this framework to include the DOOR operon database, genome

The first step in the genome-scale metabolic network reconstruction process involves the generation of a draft reconstruction and the genome annotation. Different approaches to annotate the transcripts were used; BLASTp against SwissProt enzyme protein sequences, InterProScan and PRIAM, the 85% of the Chlorella's predicted genes were annotated.

The draft reconstruction was done based on homology to Chlamydomonas rheinhardtii. The draft reconstruction accounts for 621 genes associated with 1,108 reactions and 1,249 metabolites distributed in six compartments: cytoplasm, extracellular space, chloroplast, mitochondria, thylakoid, and glyoxysome. In the second step the draft reconstruction is manually refined following a protocol for generating a high-quality genome-scale metabolic reconstruction, during the manual curation the biomass function (objective function) is defined, it contains the cell composition, minerals and energetic requirements to generate biomass; for this step literature, annotation of other models and databases were used in order to improve and complete the reconstruction.

The current reconstruction for C. vulgaris contains over 900 genes for more than 1750 reactions, almost 1600 metabolites, and includes 250 transport reactions. Lipid metabolism represents 35% of the model; other main pathways such as photosynthetic pigments and carbohydrate storage were also included. Over 200 references have been added to the model. Finally, using the COBRA Toolbox the manually curated reconstruction is analyzed for connectivity, mass and charge balance, and converted into a functional mathematical model. The functions in the automated reconstruction will be evaluated using physico-chemical constraints and tested against well-known metabolic capabilities of C. vulgaris such as growth rate, by-products and secretion.

The final C. vulgaris reconstruction is currently validated using experimental data such as cellular biochemical composition obtained under different growth conditions. Given the importance of a comprehensive organization of the available data and information to be used during the model curation, a specific bibliome database for C. vulgaris has been created.

Supported by DE-FOA-0001060, Biological Systems Science Division and EFRI grant.

154. Engineering Robust Hosts for Microbial Biofuel Production

Tim Tomko *1(ttomko@uvm.edu), Jessica Lindle1, Will Turner1, Mary J. Dunlop1

1 University of Vermont, School of Engineering, Burlington, VT 05405

Project Goals: The overall goal of this project is to enhance microbial synthesis of next- generation biofuels by developing tools for improving microbial tolerance of biofuel production conditions. Research is organized around three objectives: (1) Identify novel biofuel tolerance mechanisms from microorganisms that naturally thrive in hydrocarbon- rich environments. (2) Engineer a synthetic feedback loop that responds to biofuel production. To optimize biofuel production yields, cells must balance several competing sources of stress. We are designing and constructing a novel feedback loop that senses biofuel production and turns on export pumps in response. (3) Integrate multiple tolerance strategies in a biofuel production strain. In addition to having the potential to greatly enhance biofuel yields, this work advances understanding of how multiple tolerance mechanisms interact within a cell.

Abstract:
A major challenge when using microorganisms to produce bulk chemicals like biofuels is that the production targets are often toxic to cells. Biofuel-like compounds are known to reduce cell viability through damage to the cell membrane and interference with essential physiological processes. Thus, cells must trade off biofuel production and survival, reducing potential yields. Recent studies have shown that strains engineered to increase tolerance can improve biofuel production yields.

Microorganisms that survive in oil-rich environments are a valuable source of tolerance mechanisms. Using a transgenic screening approach, we are building fosmid-based libraries using genomic DNA from microbes that have been isolated from hydrocarbons-rich environments. Here, we present data from a transgenic screen using a library from Marinobacter aquaeolei, a hydrocarbon-degrading microorganism isolated from an offshore oil rig. The transgenic library was screened for biofuel tolerance in Escherichia coli using bio jet fuels limonene and pinene. After 4 days of exposure to the biofuels with serial dilutions every 12 hours, we saw marked improvements in tolerance relative to the control strain without the library. We then isolated the fosmids and identified a single 40kb sequence responsible for improving tolerance. Using the fosmid DNA, we constructed a smaller plasmid library and used this to isolate the specific gene responsible for tolerance.

In addition to identifying novel tolerance mechanics, we are designing control systems for efflux pumps known to export biofuel. Pump overexpression inhibits cell growth, suggesting a trade-off between biofuel and pump toxicity. To counter this, we are using the protein MexR, native to Pseudomonas aeruginosa, as a biosensor because it detects oxidative stress such as that caused by the introduction of biofuels. In the feedback loop design, MexR represses the expression of an efflux pump derived from M. aquaeolei by binding to a synthetic promoter region. We developed a library of synthetic promoters, which vary the number and location of MexR binding sites, and screened these for tolerance to pinene, a known pump substrate. The screen tests both constant and dynamic biofuel environments. The dynamic environment is important for selecting a sensor that performs well in both the presence and absence of biofuel. Our experimental findings are further supported by a mathematical model describing the dynamic sensor selection.

We also studied whether expressing multiple pumps in combination could further increase biofuel tolerance. With multiple pumps, the combined impact of pump toxicity and benefits from increased tolerance are unclear. To address this, we measured tolerance of E. coli to pinene in one-pump and two-pump strains. To support our experiments, we developed a mathematical model describing toxicity due to

biofuel and overexpression of pumps. We found that data from one-pump strains can accurately predict the performance of two-pump strains. This result suggests that it may be possible to dramatically reduce the number of experiments required for characterizing the effects of combined biofuel tolerance mechanisms.

Publications

1. W. J. Turner, M. J. Dunlop. "Trade-offs in Improving Biofuel Tolerance Using Combinations of Efflux Pumps." ACS Synthetic Biology, DOI: 10.1021/sb500307w, 2014.
2. M. Frederix, K. Hutter, J. Leu, T. S. Batth, W. J. Turner, T. L. Ruegg, H. Blanch, B. A. Simmons, P. D. Adams, J. D. Keasling, M. P. Thelen, M. J. Dunlop, C. J. Petzold, A. Mukhopadhyay. "Development of a native Escherichia coli induction system for ionic liquid tolerance." PLoS One, 2014.
3. M. E. Harrison, M. J. Dunlop. "Synthetic Feedback Loop Model for Increasing Microbial Biofuel Production Using a Biosensor." Frontiers in Microbiology 3:360, 2012.
4. M. J. Dunlop. "Engineering Microbes for Tolerance to Next-Generation Biofuels. "Biotechnology for Biofuels 4:32, 2011.

This research was supported by the University of Vermont and the Office of Science (Biological and Environmental Research), US Department of Energy.

155. Design Principles Controlling Hydrogen Metabolism in Phototrophic Organisms

Maria L. Ghirardi 1*(maria.ghirardi@nrel.gov) and Matthew Wecker2 (matt.wecker@nrel.gov)

1National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401
2Genebiologics, LLC, Boulder, CO 80303
*Presenting author

Project Goals: To obtain a systems-levelunderstanding of the biological barriers that control hydrogen metabolism and prevent sustained H2 photoproduction in the green alga Chlamydomonas reinhardtii.

Photobiological H2 production from water is a clean, non-polluting and renewable technology. Although the potential light conversion efficiency to H2 by biological organisms is theoretically high (about 10%), the system is currently limited by biochemical and engineering constraints. The specific objectives of this research are covered by two Tasks: (1) development, testing, validation and utilization of novel high-throughput assays to identify photosynthetic organisms with altered H2-producing activities, thus leading to the discovery of novel strategies to circumvent known biochemical limitations; and (2) deconvolution of the network of metabolic pathways centered on six ferrredoxin homologs found in Chlamydomonas, aimed at understanding reductant flux in photobiological hydrogen production, and identifying targets for future metabolic pathway engineering strategies to reduce flux to non-productive pathways.

In Task 1, we developed, tested and validated anovel high-throughput assay to identify high H2-producing strains from an insertional mutagenesis library of C. reinhardtii. The assay uses the H2-sensing system of Rhodobactercapsulatus that responds to H2-production by algal colonies through activation of a GFP signal (Weckeret al., 2011). We validated this assay with well-characterized mutants that are eitherlow of high H2-producers (Weckerand Ghirardi, 2014). Finally, using this H2-sensing system, we have isolated fourinsertional mutant strains of C. reinhardtii that exhibit high-light H2 production and have shown that these strains show up to 100-fold increased H2 production levels compared to their wild type strains when grown at elevated light levels. We are currently identifying the site of insertion and further characterizing these strains to understand which genes are responsible forthese high-light H2-production phenotypes.

Concomitantly, we inserted aheterologous hydrogenase from Clostridium acetobutylicum into our R. capsulatus sensorstrain. In doing so, we find that (i) hydrogen is produced by the heterologous hydrogenase; (ii) hydrogen production is detected by the H2-sensorof the organism; and, remarkably, (iii) the H2 produced is derived both from fermentative and photosyntheticprocesses. A manuscript is in preparation. We therefore have created anovel selectivemeans of testing and developing hydrogenases and this system is amenable to directed evolution studies. The heterologous hydrogenase shows some uptake hydrogenaseactivity as well, and we are currently working to understand if this uptake activity is sufficient to drive photoautotrophic (H2 and CO2) or chemoautotrophic (H2, CO2, and O2) growth of the organism. If so, we may be able to use growth on H2 as an additional selection tool for hydrogenase development.

Publications:
1. Wecker, M.S. and M.L. Ghirardi, High-throughput biosensor discriminates between different algal H2-photoproducing strains. Biotechnology and Bioengineering, 2014. 111(7): p. 1332-1340.

This project was supported by the Office of Science (BER) under FWP #ERWER38.

156. Design Principles Controlling Hydrogen Metabolism in Phototrophic Organisms

Maria L. Ghirardi (maria.ghirardi@nrel.gov) and Alexandra Dubini* (Alexandra.dubini@nrel.gov)

National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401
*Presenting author

Project Goals: To obtain a systems-level understanding of the biological barriers that control hydrogen metabolism and prevent sustained H_2 photoproduction in the green alga *Chlamydomonas reinhardtii*. Photobiological H_2 production from water is a clean, non-polluting and renewable technology. Although the potential light conversion efficiency to H_2 by biological organisms is theoretically high (about 10%), the system is currently limited by biochemical and engineering constraints. The specific objectives of this research are covered by two Tasks: (1) development, testing, validation and utilization of novel high-throughput assays to identify photosynthetic organisms with altered H_2-producing activities, thus leading to the discovery of novel strategies to circumvent known biochemical limitations; and (2) deconvolution of the network of metabolic pathways centered on six ferrredoxin homologs found in Chlamydomonas, aimed at understanding reductant flux in photobiological hydrogen production, and identifying targets for future metabolic pathway engineering strategies to reduce flux to non-productive pathways.

Our previous work in Task 2 demonstrated that both FDX1 and FDX2 interact with FNR and HYDA1 and are able to reduce NADP+ and protons in vitro, although with different kinetic properties. We are currently assessing whether the differences are due to their dissimilar redox potentials or to altered binding to their interacting partners. In an effort to continue characterizing FDX2 function, we generated and successfully identified a FDX2 knock out strain. Our data show that the FDX2 mutant strain evolves around 15% less hydrogen than the wild-type, confirming that it does have a role, although secondary to FDX1 in hydrogen metabolism. On the other hand, preliminary results suggest that this minor ferredoxin is probably more closely associated with cyclic electron flux (CEF), diverting photosynthetic reductant away from hydrogen production. FDX2 was shown to interact with PGRL1, an essential component of the PSI supercomplex that meditates CEF in Chlamydomonas. Interestingly, Western blot analysis indicates also that the levels of other minor FDXs and hydrogenases in the FDX2 knock-out mutant vary, depending on the conditions tested, providing an additional clue that FDX2 is an important player on cell adaptation and electron flux modulation. Future work will include metabolomics and fluxomics analysis to understand the specific vs. complementary roles of the different FDXs in algal metabolism under stress conditions.

Publications:

1. Boehm, M, Alahuhta, M, Long, H, Old, WM, Peden, EA, Mulder, D, Brunecky, R, Lunin, V, King, P, Ghirardi, ML and Dubini, A. Two amino acid residues contribute to functional differences between the *Chlamydomonas reinhardtii* root-type Ferredoxin 2 and the leaf-type Ferredoxin 1. *In preparation.*

This project was supported by the Office of Science (BER) under FWP #ERWER38.

157. Ensemble modeling for increasing lipid production in Yarrowia lipolytica

Jimmy G Lafontaine Rivera1*, Yun Lee1, Matthew Theisen1,2, James Liao1,2

1University of California, Los Angeles, Department of Chemical and Biomolecular Engineering, Los Angeles, California; 2University of California, Los Angeles, Department of Bioengineering, Los Angeles, California.

http://www.seas.ucla.edu/~liaoj/

Project Goals: Use novel simulation methods not requiring a priori knowledge of enzyme parameters to identify possible in vivo genetic manipulations that will increase lipid production.
In metabolic engineering, the number of possible targets for genetic manipulation is prohibitively high for unguided experimental efforts. Thus, metabolic modeling is an important tool for the identification of the most promising targets for enzyme manipulation. Typical kinetic modeling is often an involved process requiring empirical determination of kinetic parameters. Ensemble modeling (EM) exploits network information, like network stoichiometry and reference steady state, to eliminate the need for determining kinetic parameters. Using a Monte Carlo approach, sufficient numbers of random parameter sets are chosen which satisfy these requirements. In EM, parameter sets aren't empirically determined. However, the network information constrains the parameter space to realistic behavior. Additionally, other data such as metabolomics or production yields can be used to further refine ensembles. Model construction for a large-scale model using EM is demonstrated here in Yarrowia lipolytica. Following model construction, identification of potential targets to increase fatty acid production is a straight-forward automated process. The methods demonstrated here are readily generalizable to other organisms with minimal information required.

This work is funded by Office of Biological and Environmental Research in the DOE Office of Science.

158. Development of A Regulated Model for Clostridium acetobutylicum

Satyakam Dash1 (sud25@psu.edu), Thomas J. Mueller1, Keerthi P. Venkataramanan2, Eleftherios Papoutsakis2, and Costas D. Maranas1

1Department of Chemical Engineering, Pennsylvania State University, University Park, Pennsylvania, USA; 2Department of Chemical Engineering and Delaware Biotechnology Institute, University of Delaware, Newark, Delaware, USA

http://maranas.che.psu.edu

Project Goals: This project aims to understand and model the stress response of Clostridium acetobutylicum ATCC 824 to two important toxic metabolites: butanol and butyrate, using a regulated genome scale model.

Clostridia are anaerobic Gram-positive Firmicutes with the ability to use varied substrates to produce a range of industrial compounds. In particular, Clostridium acetobutylicum has been used to produce butanol on an industrial scale through acetone-butanol-ethanol (ABE) fermentation. A genome-scale metabolic (GSM) model is a powerful tool to understand the metabolic capacities of an organism and develop metabolic engineering strategies for strain development. The GSM model can be integrated with stress-related specific transcriptomics information to elucidate the nexus points of regulation, which underlie cellular response to the stressors.

We describe here the construction and validation of a GSM model for C. acetobutylicum ATCC 824, iCac802. iCac802 spans 802 genes and includes 1,137 metabolites and 1,462 reactions, along with gene-protein-reaction associations. Both 13C-MFA and gene deletion data in the ABE fermentation pathway were used to test the predicted flux ranges allowed by the model. We also describe the CoreReg method to integrate transcriptomic data and identify core sets of reactions that, when their flux was selectively restricted, reproduced flux and biomass-formation ranges seen under all regulatory constraints. CoreReg was used in response to butanol and butyrate stress to tighten bounds for 50 reactions within the iCac802 model. The model, incorporating the regulatory restrictions from CoreReg under chemical stress, exhibited an approximate 70% reduction in biomass yield for most stress conditions.

CoreReg regulated the model for the two stresses and identified differences in their respective responses, including distinct core sets and the restriction of biomass production similar to experimental observations. Given the core sets predicted by the CoreReg method, remedial actions can be taken to counteract the effect of stress on metabolism. In the case of lesser-known systems, countermeasures such as plausible regulatory loops can be suggested around the affected metabolic reactions, and the hypotheses can be tested experimentally.

The work was supported by the genomic science grant from Department of Energy, USA (grant # DE-SC0007092).

159. Ensemble cell-wide kinetic modeling of anaerobic organisms to support fuels and chemicals production

Satyakam Dash1*(sud25@psu.edu), Ali Khodayari1*(auk241@psu.edu), M. Ahsanul Islam2, Yuting Zheng2, Paul Lin3, James C. Liao3, Gregory Stephanopoulos2, and Costas D. Maranas1

1The Pennsylvania State University, University Park; 2Massachusetts Institute of Technology; 3University of California, Los Angeles.

http://maranas.che.psu.edu/ http://bamel.scripts.mit.edu/gns/ http://www.seas.ucla.edu/~liaoj/

Project Goals: The goal of the project is to systematically construct dynamic models of two anaerobic organisms, Clostridium thermocellum and Moorella thermoacetica by making use of Ensemble Modeling (EM) paradigm through integration of multiple omic information (transcriptomic, proteomic, metabolomic & fluxomic). These models will be instrumental in exploring genetic interventions for overproduction of biofuel products.

Thermophilic microorganisms have garnered the interest of the bioprocess industry due to their high temperature optimal growth conditions. In particular, two phylogenetically close organisms Clostridium thermocellum and Moorella thermoacetica have been focused on in the recent years. While C. thermocellum can metabolize cellulose into biofuels such as ethanol, M. thermoacetica can metabolize syn gas using the unique Wood-Ljungdahl pathway. Despite their increasing role as bio-production platforms, they remain poorly characterized with significant uncertainty in their metabolic repertoire. To this end, we develop cell-wide dynamic models of transcription and metabolism of these two organisms using the concept of Ensemble Modeling (EM) which requires curated genome-scale metabolic (GSM) models of the organisms as its foundation.

The second generation GSM for C. thermocellum (iCth446) has been developed, which contains 446 genes and includes 598 metabolites and 637 reactions, along with gene-protein-reaction associations. The GSM is devoid of thermodynamically infeasible cycles and contains elementally and charge-balanced reactions. The GSM was simulated with down-regulation of phosphoenolpyruvate carboxykinase, or down-regulation of malic enzyme and malate dehydrogenase knocked out along with exogenous pyruvate kinase knocked in and lactate dehydrogenase knocked out. The simulations showed higher yield of ethanol production compared to wild-type conditions as observed experimentally [1]. Likewise, the GSM results showed that only lactate hydrogenase knock out did not have any effect on growth rate as observed experimentally [2].

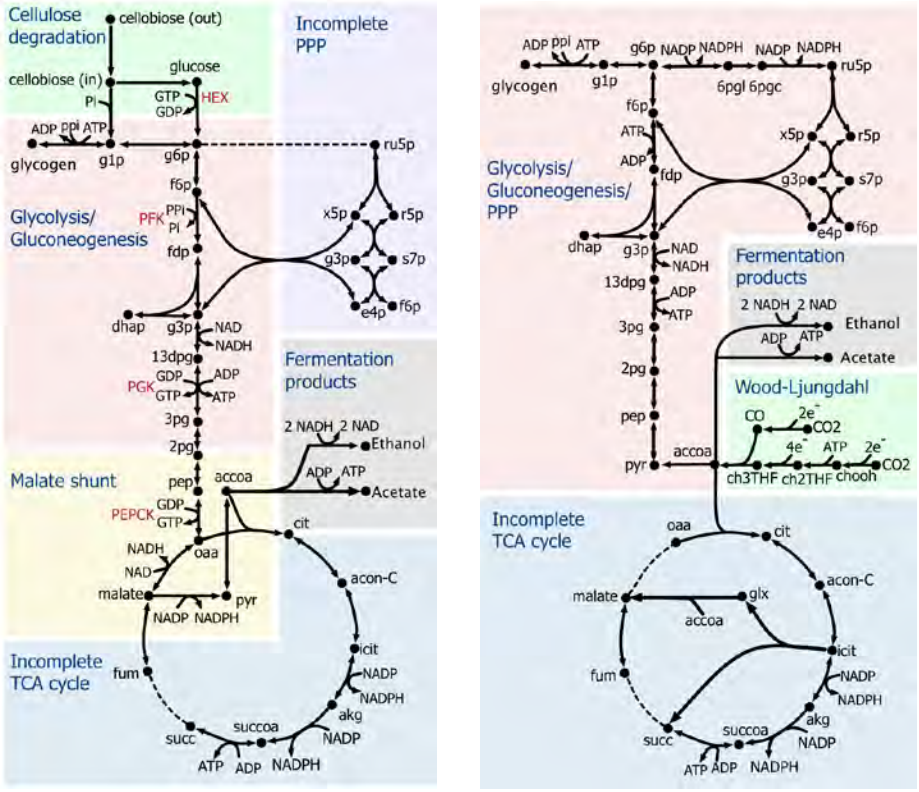
The GSM model for M. thermoacetica has been constructed using a semi-automatic pipeline in the SEED database [3]. This model contains 517 genes and 812 reactions, supplemented with 57 exchange reactions including biomass demand reaction. The GSM will be curated, gap-filled, and incorporated with the estimated growth and non-growth associated ATP (GAM & NGAM) requirements. Steady-state 13C-labelling experiments will be used to validate flux distribution predicted by the GSM, followed by instationary 13C-labeling experiments to identify the robustness of the GSM under various growth conditions with different substrates in distinct growth phases. Ultimately, these experiments will resolve the long-standing debate surrounding the hypothetical incomplete TCA cycle of M. thermoacetica.

The constructed stoichiometry representations will subsequently serve as the scaffold for building kinetic model using the EM approach for C. thermocellum and M. thermoacetica. The EM procedure will allow us to integrate substrate-level as well as transcriptional level regulatory interactions into the framework. The constructed kinetic models will be ultimately used to identify the effect of transcription factors as

well as enzyme level manipulations on metabolic fluxes leading to explore key metabolic derivers that underpin various biofuels production.

The work was supported by the genomic science grant from Department of Energy, USA (grant # DE-FOA-0001060).

1. Deng Y, Olson DG, Zhou JL, Herring CD, Shaw AJ, Lynd LR: Redirecting carbon flux through exogenous pyruvate kinase to achieve high ethanol yields in Clostridium thermocellum. Metab Eng 2013, 15:151-158.
2. van der Veen D, Lo J, Brown SD, Johnson CM, Tschaplinski TJ, Martin M, Engle NL, van den Berg RA, Argyros AD, Caiazza NC et al: Characterization of Clostridium thermocellum strains with disrupted fermentation end-product pathways. J Ind Microbiol Biotechnol 2013, 40(7):725-734.
3. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, et al: The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res 2005, 33:5691- 5702.



Central metabolism of Clostridium thermocellum (left) and Moorella thermoacetica (right)

160. Improving computational strain design strategies by incorporation of kinetic information

Ali Khodayari (auk241@psu.edu), Anupam Chowdhury, and Costas D. Maranas

Department of Chemical Engineering, Pennsylvania State University, University Park, Pennsylvania, USA.

http://maranas.che.psu.edu

Project Goals: The goal of this effort is to assess the scope of using a hybrid kinetic and stoichiometric representation of metabolism for computational strain design. In particular, we compared the performance of the k-OptForce protocol with stoichiometry-only procedures under the same condition for which the kinetic model was parameterized as well as a different environmental condition.

A key methodological impediment of existing computational strain-design approaches is the stoichiometry-only representation of metabolism and the on-off representation of regulation. This may lead to a metabolite concentration, enzymatic activity and metabolic regulation-agnostic intervention strategies. Therefore, identified flux redirection predictions (especially up/down flux modulation) are sometimes difficult to translate into actionable genetic interventions. Strain design prediction accuracy has been the focus for many recent efforts through the selective integration of kinetic information into metabolic models. The recently developed k-OptForce procedure [1] extended the previously developed strain-design OptForce algorithm [2] by integrating all available mechanistic detail afforded by kinetic models within a constraint-based optimization framework tractable even for genome-scale models.

Application of k-OptForce in case studies revealed that, in general, the kinetic model prediction quality is determined by the range and scope of genetic and/or environmental perturbations used during its parameterization. In our most recent work [3], we applied the k-OptForce procedure on a kinetic model of E. coli core metabolism constructed using the Ensemble Modeling (EM) method and parameterized using seven mutant strains flux data under aerobic respiration with glucose as the carbon source [4]. The kinetic model includes 138 reactions, 93 metabolites and 60 substrate-level regulatory interactions. Minimal interventions are identified that improve succinate yield under both aerobic and anaerobic conditions to test the fidelity of model predictions under both genetic and environmental perturbations. Under aerobic condition, k-OptForce identifies interventions that match existing experimental strategies pointing at a number of unexplored flux redirections such as routing glyoxylate flux through the glycerate metabolism to improve succinate yield. Many of the identified interventions rely on the kinetic descriptions and would not be discoverable by a purely stoichiometric description. In contrast, under fermentative (anaerobic) conditions, k-OptForce fails to identify key interventions including up-regulation of anaplerotic reactions and elimination of competitive fermentative products. This is due to the fact that the pathways activated under anaerobic conditions were not properly parameterized as only aerobic flux data were used in the model construction.

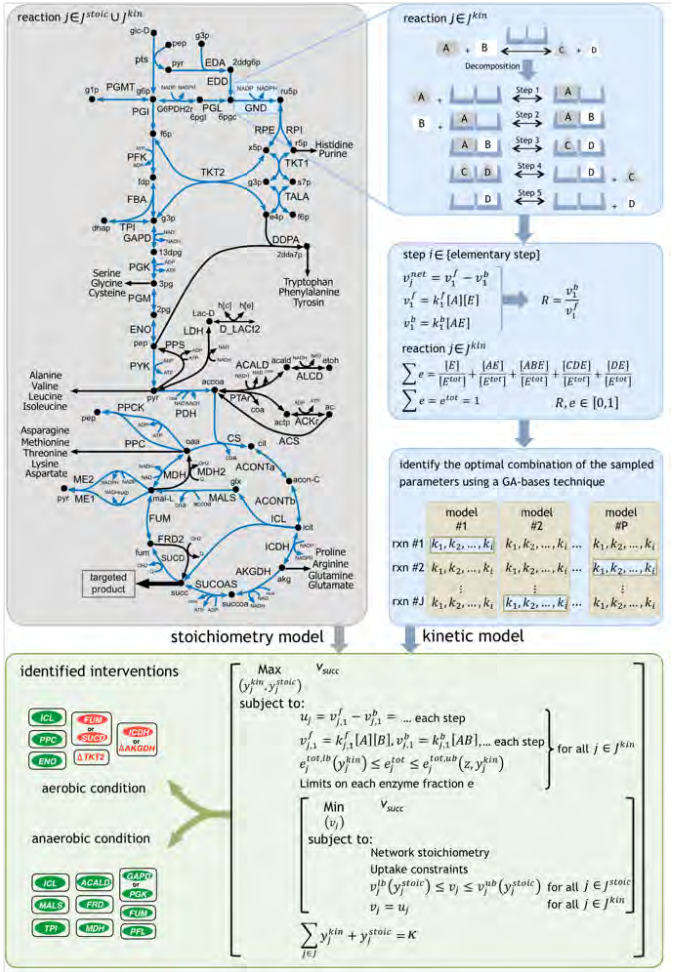
Inspired by these observations, we extend the range and scope of the current kinetic model with inclusion of all relevant reactions from the genome-scale iAF1260 model. Parameterization is carried out using steady-state flux data for 25 mutant strains under both aerobic as well as anaerobic conditions with different carbon sources (glucose, acetate, and pyruvate). The model consists of 457 reactions, 337 metabolites and 295 substrate-level regulatory interactions accounting for glycolysis/gluconeogenesis, Pentose Phosphate (PP) pathway, TCA cycle, anaplerotic reactions, amino-acid synthesis/degradation, fatty acid oxidation/synthesis, proline synthesis and a number of reactions in other parts of the metabolism. The developed kinetic description will be ultimately integrated with computational strain

design protocols to improve the accuracy of the identified strategies for overproduction of chemicals of interest.

This study reveled the importance of condition-specific model parameterization and provides guidelines on how to augment kinetic models so as to correctly respond to genetic as well as environmental perturbations.

The work was supported by the genomic science grant from Department of Energy, USA (grant# DE-ER65254).

1. Chowdhury A, Zomorrodi AR, Maranas CD: k-OptForce: Integrating kinetics with flux balance analysis for strain design. PLoS Comput Biol 2014.
2. Ranganathan S, Suthers PF, Maranas CD: OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions. PLoS Comput Biol 2010, 6(4):e1000744.
3. Khodayari A, Chowdhury A, Maranas CD: Succinate Overproduction: A Case Study of Computational Strain Design Using a Comprehensive Escherichia coli Kinetic Model. Frontiers in bioengineering and biotechnology 2014, 2:76.
4. Khodayari A, Zomorrodi AR, Liao JC, Maranas CD: A kinetic model of Escherichia coli core metabolism satisfying multiple sets of mutant flux data. Metabolic engineering 2014, 25C:50-62.



A schematic representation of the procedure [3]

161. Metabolic modeling of multi-tissue and multi-organism systems

Margaret Simons1*(mns157@psu.edu), Rajib Saha2, Mohammad Mazharul Islam1, Ali R. Zomorrodi3, Bertrand Hirel4, and Costas D. Maranas1

1Department of Chemical Engineering, The Pennsylvania State University, University Park, PA; 2Department of Energy, Environmental, and Chemical Engineering, Washington University, St. Louis, MO; 3Bioinformatics Program & Biomedical Engineering Department, Boston University, Boston, MA; 4Institut Jean-Pierre Bourgin, Institut National de la Recherche Agronomique, Centre de Versailles-Grignon, UR 511, Route de St Cyr, F-78026 Versailles Cedex, France

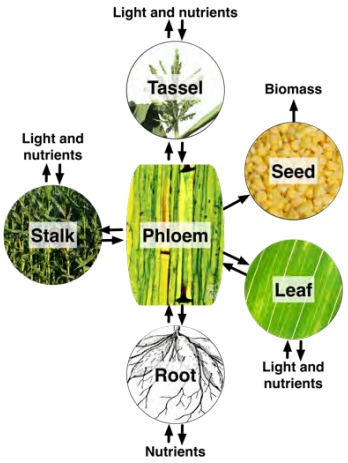
http://maranas.che.psu.edu/

Project Goals: The project aims to utilize both multiple tissue types in plants and multiple species in microbial communities to improve the understanding of metabolic interactions. The goals of this project are: 1) to determine bottlenecks in nitrogen metabolism, suggest genetic manipulations to improve nitrogen use efficiency, and enhance the understanding of nitrogen flow through maize, and 2) to analyze the physiological responses and interactions within microbial communities. A genome-scale multi-tissue maize model will be constructed to analyze nitrogen metabolism in the plant. By developing a dynamic modeling framework that utilizes constraint-based multi-level optimization, we can study the metabolic trade-offs within natural and bioengineered microbial communities, capture the temporal changes of the community, and incorporate substrate uptake kinetics.

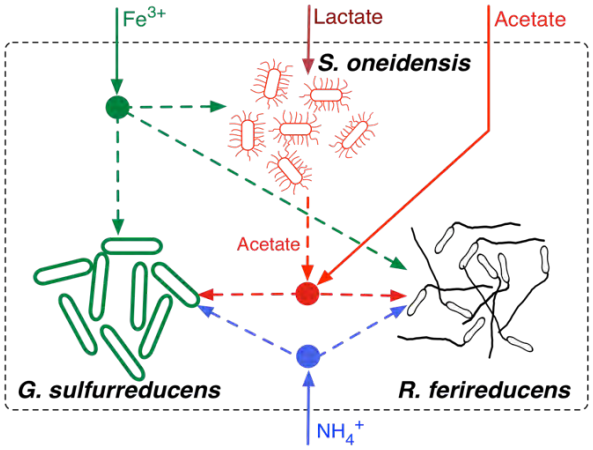
Flux balance analysis of genome-scale metabolic models is a platform used to investigate the metabolic behaviors of complex organisms or microbial communities. By reconstructing multi-tissue and multi-organism models, we can determine the interactions between different cell/tissue-types or organisms, resolve bottlenecks in limiting pathways, and study the metabolic trade-offs between species-level and community-level fitness functions.

A second-generation metabolic genome-scale model of Zea Mays L. has been constructed to investigate nitrogen assimilation and to capture C4 carbon fixation. The maize model simulates the interactions between the bundle sheath and mesophyll cells in the leaf tissue. Incorporation of the previously published iRS1563 model, the MetaCrop database, and the MaizeCyc database using Metrxn yielded a model spanning 5,824 genes and 8,525 reactions. Biomass proportions were experimentally measured under excess nitrogen, limited nitrogen, and two glutamine synthetase mutants (i.e. gln1-3 and gln1-4) yielding condition-specific biomass equations. By applying the condition-specific biomass equations and condition-specific regulatory constraints based on transcriptomic and proteomic data, four nitrogen conditions were simulated. Simulated results achieved 90% accuracy when comparing the wild-type in a nitrogen complete condition with the nitrogen deficient condition. The flux through chlorophyll biosynthesis decreases in the limited nitrogen wild-type, gln1-3 mutant, and gln1-4 mutant conditions compared to the nitrogen complete wild-type condition confirming the important association between nitrogen metabolism and chlorophyll synthesis. Regulatory constraints based on the transcriptomic and proteomic data for 19 maize lines were applied to the model to determine the metabolic differences between maize lines. As expected, when comparing two lines within the same geographical region the metabolism was similar for over 80% of the reactions. However, when comparing the B73 inbred line to the C105 northern flint line, approximately 42% of the reactions had overlapping flux ranges at maximum biomass. Ultimately, the goal is to reconstruct a multi-tissue model of all major tissue- types in maize (i.e. the root, stalk, leaf, tassel, and seed) using the phloem as a metabolite transporter. This model will be applied to analyze the flow of nitrogen from the plant root to the other tissues, suggest genetic interventions to improve nitrogen use, and study the effect of nitrogen on sugar storage in the seed.

Towards the second aim, we have developed efficient computational tools for the metabolic modeling and analysis of multi-species microbial systems involving the unidirectional or bidirectional exchange of biochemical cues. Microbial communities are known to exhibit dynamic shifts in their metabolism and in their inter-species interactions following perturbations in environmental conditions to support co-growth, survival, and stability. In order to capture the temporal dynamics of microbial communities, we developed a modeling framework called d-OptCom, which incorporates the kinetic uptake of shared metabolites and, analogous to OptCom, integrates species- and community-level fitness functions. d-OptCom was used to assess the dynamics within a uranium-reducing community comprised of Geobacter sulfurreducens, Rhodoferax ferireducens and Shewanella oneidensis. By applying d-OptCom, we suggested that the injection of lactate to the community may be a more effective bioremediation strategy than the currently employed acetate injections. The study highlights the importance of simultaneously accounting for both species- and community-level fitness functions and demonstrates that uptake kinetics substantially restricts the feasible space of inter-species flux trafficking. Overall, this work paves promising frontiers for the dynamic multi- objective analysis of complex microbial ecosystems.



Multi---Tissue Maize Model



Multi---Organism Uranium---Reducing Community

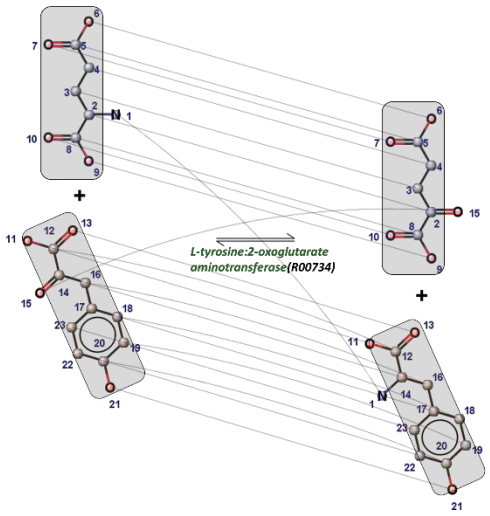
The work was supported by the genomic science grant from Department of Energy, USA (grant# DEFG02-05ER25684).

162. MetRxn2.0: Integrating atom mapping information for pathway comparisons and metabolic flux elucidation through MFA

Akhil Kumar2* (azk172@psu.edu), Saratram Gopalakrishnan1* (sxg375@psu.edu) and Costas D. Maranas1

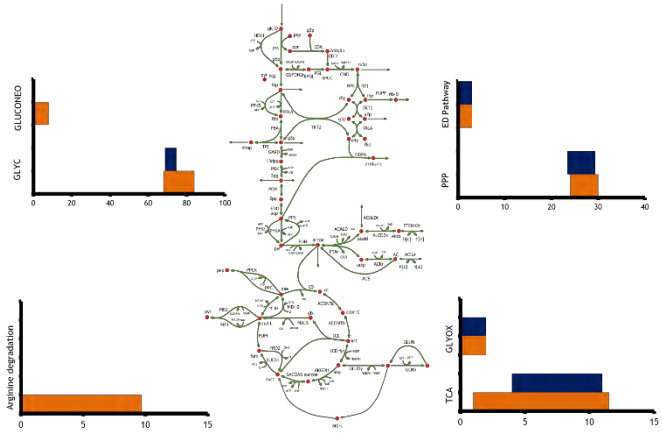
1Department of Chemical Engineering, Pennsylvania State University, University Park, PA; 2 The Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, PA;

Project Goals: This project aims to organize and disseminate standardized metabolite and reaction information to improve metabolic modeling by accurately describing reaction stoichiometry, directionality, atom mapping from reactants to products, and gene to protein to reaction relations. The generated mapping model will then be used to elucidate metabolic fluxes using 13-C Metabolic Flux Analysis (MFA) at the genome-scale to provide insights into the role of model scale-up and biomass composition on metabolic flux resolution.



MetRxn (http://www.metrxn.che.psu.edu/) is a freely available searchable web-resource that integrates heterogeneous information from 8 metabolic databases and 112 metabolic models. The MetRxn project aims to organize and disseminate standardized metabolite and reaction information to improve metabolic modeling by accurately describing reaction stoichiometry, directionality, atom mapping from reactants to products, and gene to protein to reaction relations. A host of standardization algorithms on the integrated dataset is applied to automatically curate information by removing incompatibilities in content representation, fixing stoichiometric errors such as elemental or charge imbalances and resolving incomplete atomistic details. For each reaction, metabolite stoichiometry, atom transition and metabolite compartment information is stored. The reaction and metabolite information is downloadable in SBML 3.0 and in a tabular format. The number of distinct reactions that have been mapped is greater than 27,000 and MetRxn contains tools that allow users to download atom mapping data for each reaction. In addition, all charge and mass balanced reactions within the database are processed by our novel algorithm; Canonical Labelling for Clique Approximation (CLCA). CLCA leverages prime factorization to quickly generate unique molecular graphs, detect symmetries for all metabolites and resolve atom/bond transition maps of reactions. The atom transition information is then utilized for the construction of genome-scale mapping models to address the impact of model scale-up on prediction fidelity of metabolic fluxes using 13C Metabolic Flux Analysis (MFA).

Metabolic models used in 13-C metabolic flux analysis generally include a limited number of reactions from the central metabolic network while omitting degradation pathways, complete cofactor balances, or atom transition contributions for reactions outside central metabolism. The base mapping model employed in this study accounts for (75 reactions and 65 metabolites) primarily from central metabolism. The genome-scale mapping model (GSMM) (697 reaction and 595 metabolites) is constructed using as a basis the iAF1260 model upon eliminating reactions guaranteed not to carry flux based on growth and fermentation data for a minimal glucose growth medium. Metabolic fluxes and confidence intervals are estimated, for both base and genome-scale mapping models, by minimizing the sum of square of differences between predicted and experimentally measured labeling patterns using the EMU decomposition algorithm.



Overall, we find that both the topology and estimated values of the metabolic fluxes remain largely consistent between the base and GSMM. Stepping up to a genome-scale mapping model leads to wider flux inference ranges for 20 key reactions in the base model. The glycolysis flux range doubled due to the possibility of active gluconeogenesis, the TCA flux range expanded by 80% due to the availability of a bypass through arginine consistent with labeling data, and the transhydrogenase reaction flux was essentially unresolved due to the presence of as many as 5 routes for the inter-conversion of NADPH to NADH afforded by the genome-scale model. Owing to higher growth-associated ATP demands, the available free ATP decreased drastically with the lower bound corresponding to maintenance ATP requirement. A non-zero flux for the arginine degradation pathway was identified for meeting biomass precursor demands as detailed in the iAF1260 model. Inferred ranges for 81% of the reactions in the GSM model varied less than one-tenth of the basis glucose uptake rate (95% confidence test). This is because as many as 521 reactions in the GSM are growth coupled meaning that the single measurement of biomass formation rate locks the reaction flux values. This implies that accurate biomass formation rate and composition are critical for resolving metabolic fluxes away from central metabolism and suggests the importance of biomass composition (re)assessment under different genetic and environmental backgrounds.

Supported by funding from the U.S. Department of Energy to Dr. Costas D. Maranas grant DE-FG02-05ER25684.

163. Characterization of an Obligately Syntrophic H2-producing Bacterial Coculture

Breah LaSarre, Alexandra L. McCully, Timothy A. Kremer, Amanda L. Posto, and James B. McKinlay* (jmckinla@indiana.edu)

Indiana University, Bloomington

Project Goals: The goals of this project are to (i) develop a stable hydrogen gas-producing coculture between Rhodopseudomonas palustris and Escherichia coli, (ii) use genetic, biochemical, evolutionary, and systems biology approaches to characterize and manipulate microbial interactions and H2 production, and (iii) establish stable cocultures between R. palustris and other fermentative microbes.

Synthetic microbial communities can be a valuable experimental system to understand microbial interactions and coevolution. Synthetic communities also offer opportunities to combine complementary metabolic traits to convert renewable resources into fuels and other useful chemicals. However, the utility of such systems often hinges on the ability on maintain a stable productive relationship between the species in the synthetic community.

Our work focuses on a synthetic anaerobic community composed of fermentative Escherichia coli and photoheterotrophic Rhodopseudomonas palustris. The coculture converts carbohydrates into H2 gas. E. coli produces H2 gas from carbohydrates but at a low yield due to the obligate production of organic acids and alcohols. R. palustris consumes fermentation products and use some of the electrons to produce H2 gas via nitrogenase. It has long been realized that combining these two lifestyles results in higher H2 yields from carbohydrates (1). However, progress has been impeded by the challenge of maintaining stable relationships. Through defined mutations and environmental conditions we developed a stable coculture of E. coli and R. palustris. As in previous cocultures, E. coli ferments carbohydrates and excretes essential carbon for R. palustris. Our system is stabilized by requiring that R. palustris fix N2 gas and excrete essential nitrogen for E. coli. One species cannot survive without the other.

We are examining the environmental, metabolic, and evolutionary factors that influence coculture productivity and nutrient exchange. Use of N2 versus NH4+ as the sole nitrogen source has profound effects on the species ratio, H2 productivity, and coculture stability. Depending on whether the cocultures are shaken or left static also has profound effects on the H2 yield. Static cocultures are expected to limit N2 diffusion into the medium and thereby induce N2 starvation. However, even under static conditions, cocultures remain viable and give reproducible results through serial transfers.

We have also begun to explore whether other industrially-relevant fermentative microbes can be substituted for E. coli in the coculture. When we attempted to coculture R. palustris with the ethanol-producing bacterium, Zymomonas mobilis, our negative controls lacking R. palustris also grew. This led to the discovery that Z. mobilis has the native ability to fix N2 (2). Remarkably, fixing N2 did not detract from the ethanol yield. Rather, the ethanol yield remained near the theoretical maximum during growth with N2. Growth with N2 also resulted in a higher specific rate of ethanol production and less residual biomass, compared with growth with ammonium.

N2-fixing Z. mobilis is potentially well-suited for cellulosic ethanol production as it does not require traditional nitrogen supplements needed to make up for the low-nitrogen contents of cellulosic feedstocks.

References:

1. Odom JM, Wall JD. 1983. Photoproduction of H2 from cellulose by an anaerobic bacterial coculture. Appl Environ Microbiol 45:1300–1305.

2. Kremer TA, LaSarre B, Posto AL, McKinlay JB. 2015. N2 gas is an effective fertilizer for bioethanol production by Zymomonas mobilis. Proc Natl Acad Sci USA doi: 10.1073/pnas.1420663112.

This work was supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under award number DE-SC0008131.

164. Systems Level Study of a Novel Fast-Growing Cyanobacterial Strain for Next-Generation Biofuel Production

Jingjie Yu1, Bertram M. Berla1, Rajib Saha1, Thomas J Mueller2, Akhil Kumar2, Maggie Simons2, Ali Khodayari2, Whitney Hollinshead3, Mary Abernathy3, Lian He3, Yinjie J. Tang3, Costas D. Maranas2, Himadri B. Pakrasi1,3,*

1Department of Biology, Washington University, St. Louis, MO 63130; 2Department of Chemical Engineering, Pennsylvania State University, University Park, PA 16802; 3Department of Energy, Environmental and Chemical Engineering, Washington University, St. Louis, MO 63130.

http://pages.wustl.edu/photo.synth.bio; http://maranas.che.psu.edu/; http://tang.eece.wustl.edu/

Project Goals: The overall objective of this project is to develop a fast growing cyanobacterial strain as a platform organism for photobiological production of advanced biofuels and other chemicals. Cyanobacteria are oxygenic photosynthetic microbes that use light and CO2 as feedstocks. A key issue in cyanobacterial biofuel production is the growth rates of these microbes. Compared to other photosynthetic organisms, many cyanobacterial strains have superior growth properties. However, they grow significantly slower than heterotrophic microbes such as E. coli and yeast that are commonly used in biofuel research. We have identified Synechococcus elongatus UTEX 2973, a unicellular cyanobacterium that has the ability of rapid photoautotrophic growth comparable to that of Saccharomyces cerevisiae. In this project, we are pursuing a systems approach to develop Synechococcus 2973 as a versatile and useful model cyanobacterial host to be used by the bioenergy research community during the coming era.

Identification of Synechococcus 2973 as a cyanobacterial strain with rapid growth properties (Pakrasi group): Synechococcus 2973 exhibited a 1.9 hours doubling time at 41°C under 3% CO2 and 500 μmole photons m-2s-1. This growth rate was more than 2 times faster than that of its closest related strain, Synechococcus elongatus PCC 7942, a widely studied model cyanobacterial strain1. Interestingly, Synechococcus 2973 was isolated from Austin, Texas, whereas Synechococcus 7942 was isolated from Oakland, California. Between the genomes of these two organisms, there were a total of 55 SNPs and indels. Among these 55 differences, 39 were in chromosomes and 16 were in plasmids. Among the 39 differences in chromosomes, 28 were in protein coding regions, plus a 5 kb deletion and a 188 kb inversion. Among the 16 differences in the plasmids, 7 were in protein coding regions.

Synechococcus 2973 can be readily transformed by conjugation, and fully segregated genetic mutants can be acquired more quickly than with Synechocystis sp. PCC 6803, another model workhorse cyanobacterial strain that we have studied for a long time. During this project period, we are generating a genome wide single gene knockout library of mutants in Synechococcus 2973. Such mutant strains are then being deployed for advanced metabolite analysis to aid development and refinement of a genome scale model (see below).

We have also developed new insights into the physiological function of a promising candidate biofuel molecule in cyanobacteria. All cyanobacterial membranes contain diesel- range C15-C19 hydrocarbons at concentrations similar to chlorophyll. Recently, two universal but mutually exclusive hydrocarbon production pathways in cyanobacteria were discovered. We engineered a mutant strain of Synechocystis 6803 with its alkane production pathway inactivated. This mutant grew poorly at low temperatures. We analyzed this growth defect in the mutant by assessing the redox kinetics of Photosystem I, a central hub in reductant production by the photosynthetic electron transfer pathway. The mutant exhibited enhanced cyclic electron flow (CEF), especially at low temperatures. CEF in photosynthesis raises the ATP:NADPH ratio produced and allows photoautotrophs to balance the reductant requirements of

biosynthesis with maintenance of the redox poise of the electron transport chain. An in silico flux balance analysis showed that growth rate reached a distinct maximum for an intermediate value of CEF equivalent to recycling of 1 electron in 4 from PSI to the plastoquinone pool. Based on this analysis, we conclude that the lack of membrane alkanes necessitates higher CEF for maintenance of redox poise, perhaps because of reduced membrane fluidity and electron carrier mobility. In turn, increased CEF reduces growth by forcing the cell to use less energy- efficient pathways, and in effect lowering the quantum efficiency of photosynthesis. This study highlights the unique and universal role of medium-chain hydrocarbons in cyanobacterial thylakoid membranes: they regulate redox balance and reductant partitioning in these oxygenic photosynthetic cells under stress conditions.

Genome Scale Modeling (Maranas Group): The genome-scale metabolic model for Synechococcus 2973 is currently under development based on (1) an updated version of the iSyn731 model of Synechocystis 6803 containing 109 additional reactions, (2) 106 annotations with associated EC numbers for Synechococcus 2973, and (3) gene annotations from the closely related Synechococcus 7942. Using the GSM model as a starting point, reaction atom mapping information is generated using the MetRxn database and our novel atom mapping algorithm CLCA. For approximately 55% of the iSyn731 reactions, detailed atom mapping information is already in the MetRxn database. CLCA leverages number theory (i.e., prime factorization) to identify a one-to-one mapping between the vertices of reactant and product molecular graphs. Reaction atom mapping information will be used in conjunction with 13C tracer data for metabolic flux elucidation for the wild type and also a comprehensive set of single gene mutants. Metabolic flux information for a number of genetic perturbations will then be used to estimate kinetic parameters of the reactions in central metabolism using an Ensemble Modeling paradigm. This modeling infrastructure has recently been leveraged by the C.D. Maranas group to construct a large-scale kinetic model of E. coli core metabolism spanning 138 reactions, 93 metabolites and 60 substrate-level regulatory interactions. Model parameterization was carried out using flux data for seven separate mutants.

Development of fluxomics tools (Tang group): We are developing new non-stationary MFA protocols for the analysis of metabolism in Synechococcus 2973 wild type and mutant strains. Nonstationary 13C-MFA during photoautotrophic growth of cyanobacteria has been pioneered by Jamie Young's group. We have made several modifications in their cultivation and sampling protocols. First, to avoid metabolic disturbances (light/carbon) caused by culture volume changes during continuous sampling from photobioreactors, we are using an alternative approach. We prepare a mother culture, and then divide it into smaller shake flasks (~25 mL).

Each flask is then pulsed with 1 mL ~40 g/L $NaH^{13}CO_3$, and sampled as one time point (seconds to hours). We use liquid nitrogen to chill the quenching solution (10 mL methanol) and quickly add the shake flask culture to it. The total sampling time can be as fast as 10 s. Secondly, our experiments are also modified using an inverse labeling approach to probe cell metabolism using gaseous CO_2. The mother culture is grown in 13C-bicarbonate, and then the fully labeled culture is pulsed with $^{12}CO_2$ for differential labeling. Thirdly, we use methanol-chloroform method for metabolite extraction. We are employing both traditional GC-MS and advanced LC-MS to analyze metabolite labeling. Once fully standardized, this method will be used for rapid analysis of a large set of Synechococcus 2973 mutant strains generated during this project period (see above).

1. Yu, J., Liberton, M., Cliften, P. F., Head, R. D., Jacobs, J. M., Smith, R. D., Koppenaal, D. W., Brand, J. J. and Pakrasi, H. B. (2015) Synechococcus elongatus UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO_2. Nature Scientific Rep., 5: 8132.

This project is being supported by the Office of Biological and Environmental Research in the DOE Office of Science.

165. Intracellular Iron Distribution in Chlamydomonas in Response to Fe Nutrition Status

Marcus Miethke1* (miethke@chem.ucla.edu), Jefferson Chan,2 Peter K. Weber,3 Christopher J. Chang,2 Joseph A. Loo,1,4 Jennifer Pett-Ridge,3 and Sabeeha S. Merchant1,4

1Department of Chemistry and Biochemistry, University of California, Los Angeles, CA; 2Department of Chemistry and Howard Hughes Medical Institute, University of California, Berkeley, CA; 3Lawrence Livermore National Laboratory (LLNL), Livermore, CA; 4Institute of Genomics and Proteomics, University of California, Los Angeles, CA

http://www.chem.ucla.edu/dept/Faculty/merchant/#research

Project Goals: The LLNL Biofuels SFA seeks to support robust and sustainable microalgae fuel production through a systems biology understanding of algal-bacterial interactions. We hypothesize that by understanding the factors that control cellular physiology and biogeochemical fluxes in and out of algal cells, particularly through the phycosphere, we can advance the efficiency and reliability of algal biofuel production. Our research includes studies of probiotic traits of phycosphere-associated bacteria, systems biology studies of model algae, and genome-enabled metabolic modeling to predict the interspecies exchanges that promote algal growth, lipid production and healthy co-cultures. Our overall goal is to develop the comprehensive understanding of complex microbial communities needed to advance the use of biological properties for practical energy production.

Abstract: The green alga Chlamydomonas requires metal cofactors to sustain its photosynthetic and respiratory capacities, and of these Fe is the most dominant. The demand for Fe is reflected by the intracellular quota (~2 x 108 atoms/cell), which in Chlamydomonas is strongly dependent on the available external concentration. Fe can accumulate beyond the quota when excess Fe is supplied in the external growth milieu. To better understand acclimations resulting in Fe homeostasis in a phototrophic organism while it is facing changes in iron bioavailability, we have used a suite of complementary methods including (LC)-ICP-MS, confocal microscopy (Fe(II)-specific fluorescent sensors), NanoSIMS isotopic imaging and quantitative proteomics. Our results identify the pathways and compartments involved in Fe metabolism in the algal reference organism, Chlamydomonas. We grew Chlamydomonas in media where Fe concentrations ranged from 0.1 µM Fe (severe limitation) to 200 µM Fe (excess), and found the intracellular Fe content increases more than 20-fold. Strikingly, Chlamydomonas expresses one of its major iron storage proteins (ferritin 1, Fer1) under low Fe conditions, while its expression under high Fe is rather low. We separated the soluble Fer1 fraction from cultures grown under different Fe concentrations by size exclusion chromatography and quantified both protein and Fe contents by LC-MSE and ICP-MS analysis, respectively (Fig. 1). Between low and high Fe conditions, the total Fer1 pool decreases ~12-fold, however, the average iron content per Fer1 oligomer increases ~150-fold. Further, the total cellular Fe content during growth under high iron is ~5-fold higher than the Fer1-associated Fe pool.

To understand how accumulated Fe is stored at the individual cell level, we stained cells with a Fe(II)-specific fluorescent dye, and in the high iron treatment, found it appears to be sequestered in distinct subcellular compartments. Complementary NanoSIMS studies revealed a partial co-localization of Fe, Ca and P in these sites, suggesting that excess intracellular Fe is stored in lysosome-associated organelles known as acidocalcisomes1. Chlamydomonas mutant strains that are not capable of acidocalcisome formation did not show bulk scale Fe accumulation under high Fe conditions, nor did they form obvious sites of Fe sequestration at the subcellular scale. Using 57Fe stable isotope tracing followed by LC-ICP-MS and NanoSIMS isotopic imaging, we studied how the kinetics of Fe enrichment shifted from low to high Fe conditions. When cells were shifted from a 56Fe to a 57Fe enriched media, we found the isotope

label accumulated in the ferritin compartment during the first 12 hours of incubation. While the ferritin-associated Fe pool subsequently decreased; the 57Fe enrichment of lysosome compartments (acidocalcisomes) continued to increase up until 24 h (Fig. 2). These studies are allowing us to dissect the intracellular Fe distribution pathways both kinetically and spatially during the initial stages of cell growth under Fe xcess; in our upcoming work we plan a step-by-step characterization of the molecular components that constitute these pathways.

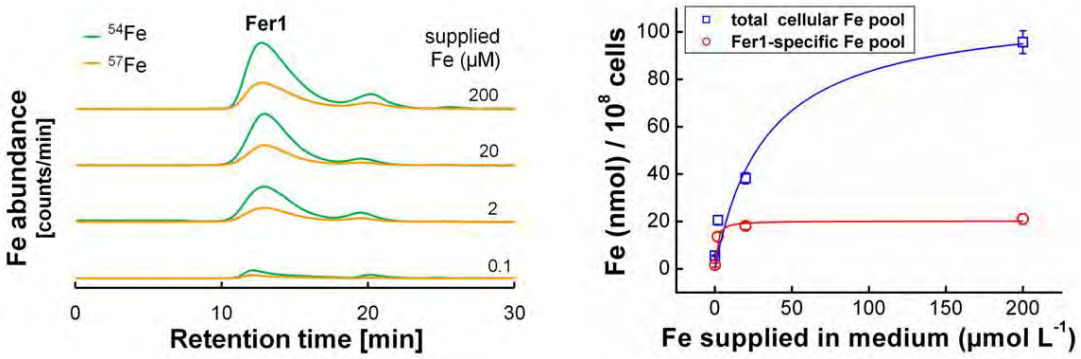


Figure 1

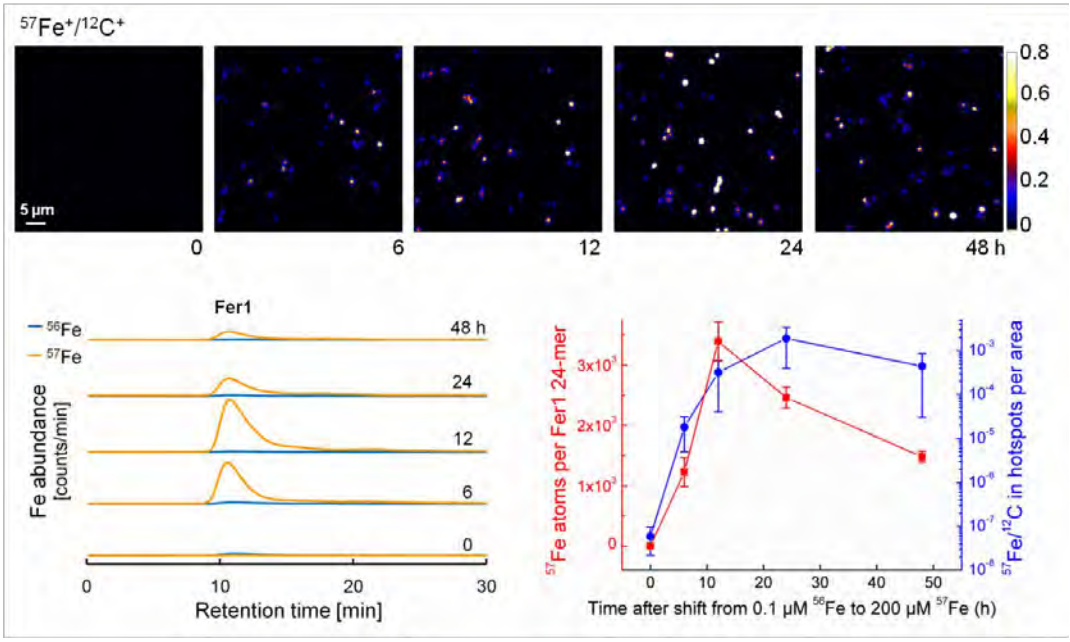


Figure 2

Publications

1. Hong-Hermesdorf, A.,* M. Miethke,* S.D. Gallaher, J. Kropat, S.C. Dodani, D. Barupala, J. Chan,
2. D. W. Domaille, D.I. Shirasaki, J.A. Loo, P.K. Weber, J. Pett-Ridge, T.L. Stemmler, C.J. Chang & S.S. Merchant. Selective sub-cellular visualization of trace metals identifies dynamic sites of Cu accumulation in Chlamydomonas. Nature Chemical Biology 10: 1034–1042 *equal contrib.

This work was supported in part by the U.S. Department of Energy (project grant DE-FD02- 04ER15529 and DE-FC02-02ER63421). Work conducted 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 supported by the Genome Sciences Program of the Office of Biological and Environmental Research under the LLNL Biofuels SFA, FWP SCW1039.

166. Towards Engineering mRNA Stability: Global Analysis of mRNA Turnover in Cyanobacteria and E. coli

Jeffrey C. Cameron1, Gina C. Gordon1,2, and Brian F. Pfleger1,2,*

1Department of Chemical and Biological Engineering , University of Wisconsin- Madison, 1415 Engineering Drive, Madison, WI 53705
2Microbiology Doctoral Training Program, University of Wisconsin-Madison, Madison, WI

Messenger RNA (mRNA) are labile intermediates that play critical roles in determining the translation rate and steady state protein concentrations in the cell. The majority of studies on mRNA turnover have focused on the model heterotrophic bacteria Escherichia coli and Bacillus subtilis. In contrast, much less is known about mRNA turnover in photosynthetic microbes. We generated a compendium of the RNases and provide an in-depth analysis of RNase III-like enzymes in commonly studied and diverse cyanobacteria. Furthermore, using targeted gene deletion, we genetically dissected the RNases in Synechococcus sp. PCC 7002, one of the fastest growing and industrially attractive cyanobacterial strains. We found that all three cyanobacterial homologs of RNase III and a member of the RNase II/R family are not essential under standard laboratory conditions, while homologs of RNase E/G, RNAse J1/J2, PNPase, and a different member of the RNase II/R family appear to be essential for growth. These and other strains are being used to identify specific ribonuclease targets with base-pair resolution and determine genome-wide rates of mRNA turnover in Synechococcus sp. PCC 7002 and E. coli using Next-Gen sequencing. The principles of mRNA decay uncovered in this analysis are being used to construct designer transcripts with predictable lifetimes to enable precise control of gene expression for metabolic engineering.

167. An integrated 'omics approach to large-scale quantitative analysis of cellular metabolic regulation

Sean R. Hackett, Vito Zanotelli, Wenxin Xu, Jonathan Goya, David H. Perlman, Joshua D. Rabinowitz

Department of Chemistry and Lewis-Sigler Institute for Integrative Genomics, Princeton University

Project Goals: Quantitative and comprehensive understanding of natural metabolic regulatory mechanisms should facilitate metabolic engineering efforts. The overarching goal of this project is to quantitatively integrate systems-level data on metabolite concentrations, enzyme concentrations, and metabolic flux to elucidate physiologically relevant metabolic regulation.

We present a strategy for discovering metabolic regulation through analysis of steady-state integrative 'omics data. Previously, we and others have attempted to elucidate regulation through measurement of time-dependent changes in metabolite concentrations (e.g., in response to acute nutrient perturbation) and dynamical modeling of the resulting data. This approach has had successes, especially for small networks. Larger dynamical models of nonlinear systems such as metabolism, however, are often intractable. To avoid complications associated with dynamic interplay between reactions, the present approach examines steady-state flux control on a reaction-by-reaction basis. Building from systems-level data on metabolite concentrations, enzyme concentrations, and fluxes across many different steady- state conditions, we assess whether the observed fluxes can be accounted for by a Michaelis-Menten relationship between enzyme, substrate and product concentrations or whether further regulation is necessary to explain flux.

As an initial test case, we used Saccharomyces cerevisiae, itself an important biofuel producer. Cells were grown in chemostats at 25 different steady states. Concentrations of metabolites were measured by LC-MS-based metabolomics and of metabolic enzymes by LC-MS/MS-based proteomics. To infer fluxes, uptake and excretion rates of the diversity of metabolites were measured, as was detailed biomass composition; together these measurements were sufficient to constrain a genome-scale flux-balanced metabolic model, resulting in reliable determination of many core metabolic fluxes. Full information (flux, enzyme concentration, and all relevant metabolite concentrations) was obtained across all 25 conditions for 55 enzymes. For about 25 of these 55 enzymes, the concentrations of the enzyme, substrates, and products alone suffice to explain the observed fluxes. For another 20 enzymes, the observed fluxes can be explained by including a single allosteric effector (e.g., fructose-1, 6- bisphosphate activation of pyruvate kinase). The approach was also able to identify an example of unexpected enzyme regulation that was biochemically verified.

A key output of the method is a quantitative assessment of which factors, overall, most strongly control metabolic flux under physiological circumstances. For reversible reactions, we find that most flux control occurs via changes in substrate and product concentrations. In contrast, irreversible reactions involve important contributions also from enzyme concentrations and allostery. Overall, the fraction of flux control residing in enzyme concentrations was modest, hinting at the difficulty of re---wiring core metabolism solely through controlling enzyme levels. Due to its reaction-by-reaction approach, the present method does not require extensive prior knowledge and thus appears well-suited to deciphering regulation also in less well studied biofuel producers.

Funding: DOE DE-SC0012461

168. Conservation of absolute cellular metabolite concentrations

Junyoung O. Park1,2* (jopark@princeton.edu), Sara A. Rubin,3 Yi-Fan Xu,1,3 Daniel Amador-Noguez,1 Jing Fan,1,3 and Joshua D. Rabinowitz1,3

1Lewis-Sigler Institute for Integrative Genomics; 2Department of Chemical and Biological Engineering; and 3Department of Chemistry, Princeton University, Princeton, N.J.

Project Goals: Absolute metabolite concentrations control enzyme binding site occupancies and the thermodynamics of metabolic reactions. Accordingly, they are critical to both metabolic modeling and metabolic engineering. Here we aim to apply a combination of direct mass spectrometry measurements and thermodynamic constraints to determine absolute metabolite concentrations across E. coli, yeast, and mammalian cells. We also aim to elucidate design principles underlying the observed abundances. To quantitate metabolite abundances, we grew organisms on isotope-labeled nutrients to label cellular metabolites, which were then extracted in the presence of unlabeled metabolite standards of known concentrations. Subsequent liquid chromatography-mass spectrometry (LC-MS) analysis yielded the ratio of labeled endogenous metabolites to unlabeled standards, which we in turn converted (after correcting for any unlabeled endogenous compound, which we measured independently) into absolute cellular metabolite concentrations. This approach, which we have described previously,[1] works well for abundant and stable metabolites with commercially available standards. Some key intermediates cannot, however, be measured directly. To evaluate concentrations of these intermediates, we used isotope tracers to determine the relative forward and backward flux through reversible reactions such as triose-phosphate isomerase. The ratio of forward to backward flux determines the reaction free energy $\Delta G'$ in the cell. Given direct measurement of all but one of the substrates/products, we can then determine the missing concentration based on the fundamental equation: $\Delta G' = \Delta G^{\circ\prime} + RT \ln Q$. We directly validated that the associated reaction free energy $\Delta G'$ approaches zero with increasing enzyme expression. Through this approach, we were able to measure 10 previously unknown absolute concentrations of central carbon metabolites and also to refine the concentrations of many others.

Here we report the absolute concentrations of 117, 82, and 93 metabolites in E. coli, yeast, and the cultured mammalian cell line iBMK. We compare the resulting metabolite concentrations to substrate and effector binding affinities in the BRENDA[2] database. We find that, as previously observed for E. coli,[1] most substrate concentrations exceed the associated enzyme binding site Km, i.e., enzyme active sites are saturated. In contrast, we find that inhibitor concentrations are typically not saturating. These trends hold true across all three organisms. Perhaps most surprisingly, despite marked differences in central metabolic fluxes, absolute metabolite concentrations were substantially conserved across the three organisms. A potential explanation is that only a finite set of metabolite concentrations simultaneously meet global thermodynamic requirements while producing sufficiently high metabolite levels to result in enzyme active site saturation and thus efficient enzyme utilization. Apparently, due to such design constraints, a wide variety of cell types from prokaryotes to mammals operate with absolute metabolite concentrations within a similar range.

References

1. Bennett, B. D. et al. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. Nat Chem Biol 5, 593-599, doi:Doi 10.1038/Nchembio.186 (2009).
2. Schomburg, I. et al. BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. Nucleic Acids Res 41, D764-D772, doi:Doi 10.1093/Nar/Gks1049 (2013).

Funding: DOE DE-SC0012461

169. The Systems Biology of Protein Acetylation in Fuel-Producing Microorganisms

Birgit Schilling1, Kori D. Dunn2, David Christensen3, Bozena Zemaitaitis3, Alexandria K. Sahu1, Bradford W. Gibson1, Christopher V. Rao*2, and Alan J. Wolfe3

1 Buck Institute for Research on Aging, Novato, CA; 2 University of Illinois at Urbana-Champaign, Urbana IL; 3 Loyola University Chicago, Maywood, IL

* Email: chris@scs.uiuc.edu

Project Goals: The goal of this project is to determine how protein acetylation affects metabolism in engineered microorganisms. Lysine acetylation is a common post-translational modification that eukaryotes, archaea, and bacteria employ to regulate protein activity. Multiple studies have recently shown that lysine acetylation predominantly targets metabolic enzymes – in fact, most metabolic enzymes are subject to lysine acetylation. Still, we know very little about the consequences of lysine acetylation, particularly in the case of bacteria. We hypothesize that bacteria employ lysine acetylation as a global mechanism to regulate metabolism in response to their energy and redox status. Our previous work suggests that lysine acetylation may be an attractive and innovative target for metabolic engineering. We are investigating how lysine acetylation affects fuel production in engineered microorganisms. The significance of this work is that it will address a fundamental gap in our understanding of bacterial metabolism and identify new approaches for overcoming the problems associated with the production of advanced biofuels.

We hypothesized that supplementation of growth media with glucose would significantly increase E. coli protein acetylation by elevating the levels of acetyl-phosphate (acP). To test this hypothesis, wild-type E. coli cells were grown in buffered tryptone broth (TB7) supplemented with or without 0.4% glucose, and temporal changes in the acetylation status of the bacterial proteome were observed. Anti-acetyllysine Western analysis revealed a proteome that became more acetylated the longer cells were exposed to glucose. To obtain a more precise understanding of glucose-induced acetylation, we affinity-enriched acetyllysine-containing peptide fractions and analyzed by mass spectrometry using an AB SCIEX TripleTOF 6600 and Skyline MS1 Filtering. Across all growth conditions and time points, we confidently identified 2813 unique lysine acetylation sites on 780 unique acetylated E. coli proteins. During growth in TB7 with glucose, the number of identified unique acetylated lysines (Kac) and proteins increased steadily over time, with 1068, 1658, 2226 and 2338 Kac sites observed at 2, 5, 8 and 12 h, respectively. To quantify changes at specific lysine acetylation sites over time in glucose, we used Skyline MS1 Filtering analysis. The acetylation status of 1091 lysines on 420 proteins showed statistically significant increases (>2 fold), with median changes of 2.7-fold (5h/2h), 5.7-fold (8h/2h), and 6.6-fold (12h/2h). However, the rate of change in acetylation varied greatly between individual lysines from the same proteins, suggesting a high degree of specificity. Network mapping of these acetylated proteins showed involvement of a diverse set of cellular processes; however, acetylations were particularly prominent in central metabolic pathways, such as the TCA cycle and glycolysis (the Embden-Meyerhoff-Parnas and pentose phosphate pathways). To represent the dynamic nature of acetylation during the investigated glucose time course, we created pictorial representations in which each enzyme of a specific pathway (i.e., TCA cycle) is displayed as a color-coded pie chart that indicates both specific lysine sites and their fold-changes. Such analysis shows that glucose-regulated sites overlap extensively with acP-regulated sites (as previously determined by comparing an ackA mutant and its WT parent). We propose that acP-dependent protein acetylation is a response to carbon flux that has the potential to regulate central metabolism.

To explore the significance of this regulation, we measured polyhydroxybutyrate (PHB) production – as a

proxy for acetyl-CoA pools – in engineered strains of E. coli. Our data demonstrate that PHB production is significantly delayed and somewhat reduced in mutants that have altered acetylation ability (yfiQ and pta/ack), suggesting that lysine acetylation plays a key role in regulating acetyl-CoA pools. We are currently assessing these results in light of our MS data.

This study was supported by the U.S. Department of Energy ER213630.

170. Estimating Intracellular Fluxes Using Relative Expression and Phenotype Datasets (REPPS)

Matthew Long1* (mrlong2@wisc.edu) and Jennifer L. Reed1

1Dept. of Chemical and Biological Engineering, Univ. of Wisconsin-Madison

Project Goals: Cyanobacteria offer a promising route for directly converting solar energy and CO2 into biofuels. The objectives of this research are to integrate modeling and experimental approaches to guide development of a butanol producing cyanobacterium, Synechococcus sp. PCC 7002. New computational approaches will be developed to facilitate these efforts which will (1) design experiments and analyze their results, and (2) identify genetic engineering strategies for improving butanol production in S. 7002. Experiments will subsequently be performed to construct and analyze Synechococcus 7002 strains engineered for butanol production. The developed approaches will be systematically applied to suggest genetic engineering strategies for improving production of a variety of biofuels in five other microorganisms. This research will support the U.S. Department of Energy's mission for developing renewable ways of producing advanced biofuels.

Renewable sources of transportation fuels are needed to reduce the amount of oil used to satisfy transportation energy needs in the U.S. and to alleviate our dependence on foreign sources of oil. Microbes can be used to produce a wide variety of liquid biofuels including: ethanol, butanol, isobutanol, isoprene, hydrogen, and alkanes. Cyanobacteria offer an alternative route for converting solar energy and CO2 into biofuels, without the need for using lignocellulosic biomass as an intermediate. The biofuel production capabilities of microbes can be improved through metabolic engineering, where metabolic and regulatory processes are adjusted using targeted genetic manipulations. Traditionally, metabolic engineering strategies are found through manual inspection of metabolic pathways, where enzymes involved in biosynthesis are overexpressed or added, competing pathways are eliminated, and the performance of resulting strains are evaluated. However, such approaches cannot predict the effects that these changes will have on other parts of metabolism, and generally will not suggest alterations to more distant pathways.

Computational models of cellular metabolic and regulatory networks can be used to guide and accelerate these metabolic engineering efforts by integrating and analyzing experimental data, and identifying genetic manipulations that would increase product yields. In the process of developing metabolically engineered strains, genetic manipulations proposed by computational strain design algorithms depend on the metabolic state of parental strains. One such algorithm is RELATCH (for relative change), which has been shown to accurately predict the effects of gene deletions and environmental shifts on metabolic fluxes [1]. However, RELATCH requires knowledge of both gene expression and fluxes in the parental strain, including intracellular flux measurements (e.g., 13C metabolic flux analysis), to predict fluxes in knockout mutants. While gene expression is easily measured, intracellular flux measurements are harder to generate and are not widely available, particularly for cyanobacteria. As such, alternative methods for obtaining knowledge of fluxes through metabolism are needed to evaluate and improve engineered strains.

A number of experimental measurements can be made to evaluate the metabolic state of a cell, such as enzyme activity, gene expression, metabolite concentrations, protein concentrations, and cellular uptake and secretion rates. The integrated analysis of these various datasets can be used to help estimate metabolic fluxes in cells and identify potential bottlenecks in biofuel production. Here we have developed a novel constraint-based modeling method for calculating the flux distribution and enzyme contributions in a parental strain using experimental data from multiple gene deletion strains. This method, Relative

Expression and Phenotypes for Parental Strain estimation (REPPS), incorporates multiple modules which can utilize growth rates, extracellular fluxes, and gene expression data from multiple knockout reference strains and the parental strain to predict intracellular fluxes. We have further evaluated the importance of both the abundance and type of data used by REPPS on the resulting intracellular flux estimates for both Escherichia coli and Saccharomyces cerevisiae. By integrating multiple datasets, we are able to more accurately estimate the parental strain flux distribution, yielding as much as a ~44% improvement compared to existing approaches (e.g., pFBA). The improved parental strain flux prediction from REPPS has then been used with RELATCH to accurately predict fluxes in new mutant strains with greater accuracy. Future work will be to validate and apply these methods to the identification and alteration of fluxes in cyanobacterial strains (e.g., Synechococcus 7002 and Synechocystis sp. PCC 6803) engineered for biofuel production.

References:

1. Kim, J. and J.L. Reed, RELATCH: relative optimality in metabolic networks explains robust metabolic and regulatory responses to perturbations. Genome Biol, 2012. 13(9): p. R78.

This work was supported by the Office of Science (BER), U.S. Department of Energy (DE- SC0008103).

171. Spatial connectomics to identify agents relevant to lignocellulose deconstruction in fungi

Melania Figueroa1, Kenneth E. Hammel2, Christopher J. Hunt2, Ellen A. Panisko3, Gerald N. Presley1, Jiwei Zhang1, and Jonathan S. Schilling1

1University of Minnesota, Saint Paul, 2U.S. Forest Products Laboratory (FPL); 3Pacific Northwest National Laboratory

http://schillinglab.cfans.umn.edu

Project Goals: Our overarching goal is to discover which genes are differentially up-regulated across the mycelia of brown rot type wood-degrading fungi in planta, particularly at the leading edge of wood decomposition. The relevance of this goal to the mission at BER and to the funding call, specifically, resides in the relevance of the fungal mechanism to lignocellulose bioprocessing. These unique fungi accomplish what we have difficulty achieving – energy from plant biomass. They do this without mutualists or in consortia, with an abbreviated CBH-free cellulase system in most cases, and often with enhanced decay rate with lignin present rather than removal as a prerequisite for access to carbohydrates. Furthermore, and of relevance to this spatially-motivated project, they apparently couple an oxidative pretreatment step with enzymatic saccharification in discrete space, via partitioning of reactions. We have previously shown evidence that these are governed by differential expression, but with genes putative, transport out of hyphae unclear, and the secretome poorly studied, historically, a comprehensive approach is needed that can also limit ('winnow') data sets from powerful global analytical tools to focus on the genes and pathways that matter.

To do this, we are comparing global expression profiles among mycelial regions and matching what we see to what they do, in terms of changing wood physiochemistry. To map a 'connectome' in a wood-fungal interaction, specifically, we must also match gene expression patterns with the extracellular secretome and with physiochemical wood modifications. Given this potential for substrate-fungus feedback, we are trying to cross-check genes using separate clade representatives for brown rot fungi alongside their white rot ancestors, harnessing the JGI MycoCosm portal and key expertise at the USDA Forest Products Laboratory and the Pacific Northwest National Laboratory.

Objective 1 – Zone localization: Use the wood wafer design to resolve a depolymerization zone in P. placenta near the mycelial leading edge, and optimize RNA extraction for thin-sectioning.
Objective 2 - Fungal connectomics: Co-localize gene expression with the secretome and with relevant physiochemical modifications made to the wood, e.g. hemicellulose loss, porosity changes.
Objective 3 - Clade comparisons: Compare key zones among brown rot clades, in context of white rot same-clade ancestors, to target universal 'brown rot' genes and candidates for bioprocessing.

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. A similar approach, 'connectomics,' has been used to map the human nervous system, and its application here is timely. First, wood-degrading fungal genomes are an emerging resource, particularly brown rot functional types. Second, 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.

References

1. Schilling, J.S., Duncan, S.M., Presley, G.N., Filley, T.R., Jurgens, J.A, and Blanchette, R.A. (2013) Colocalizing incipient reactions in wood degraded by the brown rot fungus Postia placenta.
2. International Biodeterioration and Biodegradation 83: 56-62.

Funding statement: This work is supported by the Office of Biological and Environmental Research in the DOE Office of Science.

172. Spatially co-localizing the incompatible oxidative and enzymatic steps during fungal brown rot wood degradation - Early Career Program

Robert A. Blanchette,1 Shona M. Duncan,1 Timothy Filley,2 Joel Jurgens,1 Justin Kaffenberger1, Jon Menke1* (menk0001@umn.edu), Gerald N. Presley, Jiwei Zhang, and Jonathan S. Schilling1

1University of Minnesota, Saint Paul; 2Purdue University, West Lafayette, IN

http://schillinglab.cfans.umn.edu

Project Goals: Wood-degrading fungal genomes are increasingly sequenced and annotated, including brown rot functional types, initiated with Postia placenta in 2009. Wood-degrading fungi metabolize wood either by removing lignin to access carbohydrates (white rot), or they extract carbohydrates using a mechanism to circumvent lignin, largely (brown rot). The brown rot fungi evolved more recently and at least seven different times from ancestral white rot lineages, suggesting an advantage in efficiency. As the first annotated genome representative of 'typical' brown rot fungi, P. placenta is capable of an oxidative hydroxyl radical pretreatment that occurs concurrently with enzymatic saccharification of woody carbohydrates. This consolidation of otherwise incompatible reactions is fundamentally interesting and has great implication on the potential to consolidate harsh pretreatments with saccharification in a single processing step. Therefore, our research goals have been as follows:

1) Physically sample wood degraded by the brown rot fungus P. placenta in order to map coincident pretreatment and saccharification reactions and to correlate relevant lignocellulose chemistry.

2) Use microscopy to image the fungus-plant interface along a continuum, and layer data with information about structural carbohydrate degradation and other wood modifications.

3) Map, along the active hyphal front, the co-occurring expression of oxidative genes associated with pretreatment and of glycosyl hydrolases (GHs) used in saccharification.

Abstract: Enzymatic bioconversion of lignocellulose plant tissues generally requires an initial pretreatment step, followed by saccharification and then fermentation or other downstream processing approaches. Consolidated bioprocessing (CBP) of lignocellulose combines enzymatic sugar release (saccharification) with fermentation, but pretreatments typically remain separate and costly. In nature, lignocellulose-degrading brown rot fungi consolidate pretreatment and saccharification, likely using spatial gradients to partition these incompatible reactions. Our goal is to characterize how this is achieved, in order to better understand the fungus and to potentially apply this approach in a mimicked consolidated approach.

The goal of this research is characterizing spatially the co-occurrence of relevant reactions in this biological system of high relevance for bioenergy objectives. We are doing this with a novel directional growth strategy on wood wafers, using P. placenta as our model system and a variety of physiochemical and molecular tools to overlay spatial patterns in place of a 'noisier' temporal sequence from whole wood blocks. This project is nearly complete, with one key publication close to submission. From the physical sampling efforts, we learned that there is an area of depolymerization apparently beyond those areas of lignin demethylation and glycosyl hydrolase activity, occurring at the hyphal front colonizing wood. We have matched this observation to gene expression patterns, as well, using fluorescence in situ hybridization (FISH) and confocal microscopy but without the signal strength or the chemical sensitivity with wood physiochemical analyses to connect potential with outcomes. Instead, our final efforts have coupled qPCR with the coarse sampling information generated in Objective 1 to colocalize expression of

target genes. These results are showing conclusively three novel results. 1) This fungus (P. placenta), assumed in many literature reports to have constitutive expression of cellulases indeed does not, showing inducible patterns that lag behind the hyphal front. 2) Most of the genes we tested that are putatively associated with an oxidative brown rot 'pretreatment' are, in fact, being expressed ahead of cellulases with highest transcript counts at the hyphal front and repression at the induction point for cellulases. 3) These patterns are matched by carbon release induction/repression patterns as the wood is disassembled. This collectively is in line with the theory of brown rot as a two-step pretreatment/saccharification mechanism, consolidated over small spatial distances in wood. This offers great promise for consolidating bioprocessing steps and for understanding a vital carbon cycling pathway via wood decomposition, but other aspects such as putative gene functions and intra vs. extracellular fate of proteins remain unknown, particularly given the likelihood that hyphal fronts normally confront other colonizers in nature. Addressing these gaps are critical next steps in the fundamental side of this work, and this 'wafer' system of directional growth is proving highly valuable in segregating and targeting key reactions for deeper omics analyses.

References

1. Schilling, J.S., Duncan, S.M., Presley, G.N., Filley, T.R., Jurgens, J.A, and Blanchette, R.A. (2013) Colocalizing incipient reactions in wood degraded by the brown rot fungus Postia placenta.
2. International Biodeterioration and Biodegradation 83: 56-62.

Funding statement: This work is supported by the Early Career Program and within the Office of Biological and Environmental Research in the DOE Office of Science.

173. Mapping Inter-species Interactions and Metabolic Synergy for Next-generation Biofuel Production

Arion Stettner1*(arions@bu.edu), Meghan Thommes1*(mthommes@bu.edu), Andrea Lubbe2* (alubbe@lbl.gov), Ilija Dukovski3, Brian Granger3, Jessica Lee4,5, Christopher Marx4,5, Trent Northen2 and Daniel Segre1,3,6

1Dept. of Biomedical Engineering, Boston University; 2Lawrence Berkeley National Laboratory; 3Bioinformatics Program, Boston University; 4Dept. of Biological Sciences, and 5Institute for Bioinformatics and Evolutionary Studies, University of Idaho; 6Dept. of Biology, Boston University

Project Goals: The goal of our research plan is to implement a high-throughput pipeline for the systematic, computationally-driven study and optimization of microbial interactions and their effect on lignocellulose degradation and biofuel production. We combine multiple approaches, including computer modeling of ecosystem-level microbial metabolism, mass spectrometry of metabolites, and experimental evolution.

The biologically mediated production of energy-dense fuels from the degradation of lignocellulosic material is a complex, multistep metabolic process. Harnessing microbial communities for biofuel production has the potential to enable efficient, robust, and adaptive approaches to consolidated bioprocessing. We are addressing this challenge through an integrated effort to engineer a consortium for biofuel production. In our strategy, we seek to take into account (i) the specific complex nature of the molecular substrates and intermediates consumed and produced by different organisms alone or in combinations; (ii) the natural tendency of microbes to operate at metabolic regimes that reflect their own evolutionary history and individual benefit rather than the laboratory environment and objectives, and finally, (iii) the great opportunity offered by recent developments in systems biology, in particular through predictive genome-scale models of metabolism. The set of organisms we have chosen for this project is driven by the criteria of immediate feasibility, as well as by the desire to develop a community that could serve as a prototype for follow-up practical developments. In particular, our approach is to use four different functional groups of microbes; a lignin degrader, a (hemi)cellulose degrader, a strain to remove and detoxify methoxy groups on lignin, and a biofuel producing yeast. Rather than the specific optimization of selected strains, our primary goal is a flexible, predictive strategy to maximize community output given the members present.

Towards the development of a prototype community, we have started characterizing in detail the metabolic inputs and outputs of the different members of our synthetic consortium using mass spectrometry (MS) to measure the use and production of plant hydrolysate and the microbial uptake and release of metabolites. These data will be used to test and improve genome scale metabolic network models of individual species. By extending these analyses to all pairs of organisms, we will comprehensively detect possible synergies and partner-induced changes in uptake/excretion. In parallel to these experimentally driven efforts, we are testing and refining stoichiometric models for the relevant organisms, starting from published reconstructions, and from gap-filled draft models obtained through the KBase pipeline. Simulations of the metabolic activities of these organisms can be performed through COMETS (Computation Of Microbial Ecosystems in Time and Space), our platform for dynamic flux balance modeling. Several ongoing COMETS improvements will help perform more accurate and comprehensive simulations, including a 3D convection-diffusion implementation of the growth process, and a streamlined pipeline for visualizing COMETS results through detailed multilayer networks that can highlight metabolic competition and cross-feeding interactions.

This work was supported by the Office of Science, Office of Biological and Environmental Research, of the

U. S. Department of Energy, Award No. DE-SC0012627, and the Office of Science Early Career Research Program, Office of Biological and Environmental Research, of the U. S. Department of Energy under Contract No. DE-AC02- 05CH11231.

174. Illumination of Photosynthetic Trans-Kingdom Consortia

Stephanie Hays,1* Daniel Ducat,2 & Pam Silver1

1Harvard University Department of Systems Biology, Boston MA; 2 Michigan Statue University Department of Biochemistry and Molecular Biology, East Lansing MI

http://openwetware.org/wiki/Silver_Lab http://www.prl.msu.edu/faculty/ducat

Project Goals: We are interested in engineering and studying photosynthetic consortia. Monocultures provide insights into microbiology and have been compelled to make many a product, but they are inherently limited in both capacity to manufacture goods and how honestly such growth represents the natural world. Out in the wildernesses of the world (be they in far off countries or just on the surface of your skin) microbes are fighting and struggling, helping or hindering, and simply coexisting with a slew of partners. In an effort to better understand these complex populations, we seek to engineer interactions between microbes such that they are more involved but not as convoluted as natural systems. Focusing on photosynthetic co-cultures allows us to develop knowledge of these fundamental interactions and demonstrate proof-of-principle examples for downstream biotechnology applications, while probing the limits of photosynthetic efficiencies.

The components of the co-cultures are Synechococcus elongatus, Saccharomyces cerevisiae, Bacillus subtilis, and Escherichia coli and their interactions range from natural to engineered. Growth is demonstrated while culture dynamics illuminate complicated relationships.

We would like to thank the collaborators that are helping us delve deeper into the interesting questions raised by this research. The following principal investigators and the scientists in their labs further our scientific progress.

Michael J. Betenbaugh, Johns Hopkins University
Michael Guarnieri, National Renewable Energy Laboratory George Oyler, Johns Hopkins University
Karsten Zengler, University of California, San Diego

This project was funded by the Systems Biology National Institutes of Health Grant GM080177, F32 Grant Number 10392604 from the National Center for General Medical Sciences, Army Research Office Award No. W911NF-09-1-0226, the Wyss Institute for Biologically Inspired Engineering, and special the Systems Biology of Autotrophic-Heterotrophic Symbionts for Bioenergy Grant DE SC0012658 from the DOE.

175. Development of the Oleaginous Yeast Rhodosporidium toruloides as a New Model Organism for Systems-Level Analysis of Lipid Productivity

Dominic Pinel1,2, Gina Geiselman3, Samuel T. Coradetti3,4, Adam P. Arkin2, Rachel B. Brem*3,4 (rbrem@buckinstitute.org) and Jeffrey M. Skerker*1,3(skerker@berkeley.edu)

1Energy Biosciences Institute, Berkeley, CA; 2Lawrence Berkeley National Laboratory, Berkeley, CA; 3University of California, Berkeley; 4Buck Institute for Research on Aging, Novato, CA

http://www.energybiosciencesinstitute.org/program_project/development-oleaginous-yeast-model-system

Project Goals: The most urgent need in bioenergy research is for commercial-scale production of advanced biofuels that can serve as "drop-in" replacements for petroleum-derived fuels. Toward this end, oleaginous yeasts hold great promise owing to their naturally evolved pathways for high fatty acid flux. We are carrying out massively parallel mapping of genotype to phenotype in the oleaginous yeast Rhodosporidium toruloides. Our focus is on the genes that underlie lipid production, plant feedstock hydrolysate tolerance, low-oxygen metabolism, and co-utilization of sugars in plant material. Objective 1 is to survey these attributes across wild R. toruloides strains. Objective 2 is to identify genes that underlie biofuel-relevant traits using insertional mutagenesis. Objective 3 is to map genes that underlie differences between wild R. toruloides isolates. Together, our findings will advance the understanding of lipid metabolism in R. toruloides, and enable engineering of an optimized biofuel production host.

The R. toruloides system. Upon nitrogen starvation, oleaginous yeast species accumulate large amounts of lipids, primarily triacylglycerols (TAGs), in an organelle called the lipid body; lipid content can accumulate in excess of 70% of total dry biomass. Harnessing TAG production directly for transesterification into fatty acid methyl esters (i.e., biodiesel), or ultimately, diversion of fatty acid flux in these species to synthesis of fatty acid esters and alkanes, is thus a promising strategy for large-scale production of biofuels. However, without a systems-level understanding of lipid metabolism in these systems, their use in the biofuel industry has been limited. We are focusing on Rhodosporidium toruloides, an oleaginous yeast that can natively metabolize sugars that are abundant in plant hydrolysates (glucose, xylose, arabinose, and cellobiose); has high de novo lipid productivity from glucose; grows and produces lipids in sugar concentrations equivalent to those in cellulosic plant hydrolysates; and does not itself consume alkanes, which are target biofuels. Although a full genome and annotation of one R. toruloides strain was recently reported, in only a handful of studies have R. toruloides genes been functionally characterized. To meet the challenges of strain engineering and development in industrial settings, we are carrying out a systems-level effort to identify the genes that govern R. toruloides growth and metabolic behaviors.

Lipid production and natural variation in R. toruloides. We collected six haploid R. toruloides isolates (three of each mating type) from diverse geographic locations and environments to sample genotypic and phenotypic diversity within this species. Sole-carbon-source tests revealed that all strains were able to metabolize glucose, arabinose, and cellobiose, and likewise all isolates except one grew on xylose. Thus, metabolism of sugars present in plant hydrolysates is thus far the rule rather than the exception among wild R. toruloides isolates, in striking contrast to Saccharomycetes. We used staining with the lipid dye BODIPY to qualitatively assess lipid production in these R. toruloides strains. After culturing in nitrogen starvation conditions, an environmental treatment known to induce TAG biosynthesis, each of our six wild R. toruloides isolates produced large amounts of lipid de novo from glucose. For one isolate, CBS 6570, we also confirmed lipid production after culturing on arabinose, xylose, and cellobiose. Detection of BODIPY staining with both a plate reader and a flow cytometer revealed a significant difference in

lipid accumulation between the R. toruloides strains CBS 349 and CBS 14. Thus, our data have already established that genetic background influences lipid productivity in R. toruloides.

Functional genomics in R. toruloides. To identify genes underlying lipid productivity and growth traits in R. toruloides, we have pioneered tools for transposon mutagenesis followed by competition of transposon mutants and sequencing (Tn-seq) in this system. We first established a protocol for Agrobacterium tumefaciens mediated transformation, to mutagenize the genomes of R. toruloides strain CBS 14 and CBS 349 with a T-DNA plasmid that confers resistance to the drug nourseothricin. We co-cultured A. tumefaciens strain EHA105 containing the T-DNA plasmid with R. toruloides, and selected random genomic insertions on 100 μg/mL nourseothricin and 300 ug/mL cefotaxime. We have achieved scale-up of this procedure to amass ~40,000 colonies from a given mutagenesis.

Next we established methods to map the genomic location of T-DNA insertions via Tn-seq. Here genomic DNA of a mixed pool of mutagenized R. toruloides is isolated and fragmented; adapters are ligated onto fragment ends; and T-DNA inserts are amplified with a primer recognizing the T-DNA sequence at one end and a primer recognizing the adapter on the other. Each amplicon thus contains a partial fragment of the T- DNA and a fragment of the genomic region into which it has inserted in a given clone, to be sequenced on the Illumina platform (150-bp single-end reads). Proof of concept sequencing experiments have yielded, from 1 million reads per library of DNA fragments from a pool of ~16,000 mutants cultured in rich medium, ~3200 genes in which we ascertained >20 Tn-seq reads per insertion site, i.e. 37% of the ~8600-gene R. toruloides genome. Read counts per insertion site agreed between replicates at $R2 > 0.95$.

Future directions. In the coming year we will work to scale up construction and sequencing of large pools of barcoded R. toruloides mutants. We will develop selection schemes for oil production and plant hydrolysate tolerance, in which high-fitness mutants rise to high frequency in a mixed pool while low-fitness alleles drop out, with fitnesses measured by mutant abundance in barcoded Tn-seq data. As these experiments identify genes that modulate biofuel-relevant traits, we will develop a pipeline to validate their role in focused gene disruption experiments.

This work was supported by The Office of Biological and Environmental Research in the DOE Office of Science, and by the Energy Biosciences Institute at the University of California, Berkeley.

176. Predicting Clostridium ljungdahlii Cellular Phenotypes Through a Metabolic and Gene-Expression Model

J. Liu1* (jkl055@ucsd.edu), A. Ebrahim1, C. Lloyd1, K. Zengler1

1University of California, San Diego, 9500 Gilman Drive 0419 La Jolla, CA 92093

http://sbrg.ucsd.edu

Project Goals: This project will substantially enhance our knowledge about chemolithoautotrophs and their potential for advanced biocommodity production by detailing the interconnectivity of metabolism, energy conservation, and regulation of Clostridium ljungdahlii through sequencing and modelling. As a result, we will have next-gen modeling capability as well as methods to utilize next-gen models (also known as metabolic and gene-expression models, abbreviated ME-models) for the design of tunable systems that produce biocommodities from inexpensive sources.

Clostridium ljungdahlii is emerging as an acetogenic model organism to be developed into a new chassis for strain designed chemical production as well as a platform for gaining in-depth knowledge about acetogens. There are many attractive features that endorse C. ljungdahlii for strain design, including its ability to grow either heterotrophically on a variety of sugars or autotrophically on carbon monoxide (CO), or carbon dioxide (CO2) and hydrogen (H2), as well as mixtures of those gases (i.e. syngas). Metabolism of syngas by acetogenic microorganisms produces multi-carbon organics, an ability that may be engineered to produce biocommodities.

C. ljungdahlii's known metabolic and energy pathways were reconstructed in the form of a constraint-based metabolic model (M-model). Furthermore, we have also reconstructed the macromolecular synthesis machinery (E-matrix) through sequence homology and literature curation of over 100 genes. Integration of the E-matrix into the M-model produced a predictive genome-scale ME-model for C. ljungdahlii. This ME-model serves as a theoretical baseline to understand and predict the metabolic, transcriptomic, and proteomic responses to environmental or genetic perturbations. For example, the ME-model predicts dramatic phenotypic differences in autotrophic and heterotrophic growth conditions. Unlike traditional approaches, the ME-model also accounts for critical factors and properties of candidate pathways (e.g., cost of enzymes, protein complex stoichiometry, codon usage, cofactor dependency and prosthetic group-usage), which will be of utmost importance for tuning expression of non-native pathways for biocommodity production.

Supported by DOE-DE-SC0012586, Next-Gen3: Sequencing, Modeling and Advanced Biofuels

177. Regulation of Neutral Lipid Compartmentalization in Vegetative Plant Tissues

Kent D. Chapman1 * (PI, chapman@unt.edu), John M. Dyer2 (Co-PI, john.dyer@ars.usda.gov), Robert T. Mullen3(Co-PI, rtmullen@uoguelph.ca)

1University of North Texas, Center for Plant Lipid Research, Denton, TX; 2USDA-ARS, US Arid-Land Agricultural Research Center, Maricopa, AZ; 3University of Guelph, Department of Molecular and Cellular Biology, Guelph, Ontario, Canada.

Project Goal:
The overall goal of our research program is to identify, characterize and manipulate the cellular machinery that influences the accumulation and compartmentalization of neutral lipids in vegetative tissues of plants.

Our research is focused on identifying the subcellular machinery that plants use to compartmentalize neutral lipids in cells of vegetative tissues. Virtually all cells in plants (and other organisms) synthesize triacylglycerols (TAGs) and deposit them in cytosolic lipid droplets (LDs). While seed tissues accumulate large amounts of TAGs in specialized, oleosin protein- coated LDs, most non-seed tissues do not. TAGs normally accumulate to a very small percentage in vegetative tissues of plants such as in leaves and stems. Recent work by our lab and others has demonstrated that it is possible to substantially increase the percentage of TAG in leaf tissues, raising the prospects for enhancing the bioenergy content in crop biomass. Still, there is little detailed mechanistic information available regarding the cellular machinery involved in the formation, packaging and turnover LDs, especially in cells of non-seed tissues.

In our work supported by BER, we have taken a combination of approaches to discover previously underappreciated proteins that participate in LD formation and turnover. First, proteins known to be involved in LD biogenesis in mammalian cells have been tested for their ability to modulate LD formation in plant cells, including several endogenous Arabidopsis homologues of mammalian proteins (e.g., CGI-58, seipins), or other genes encoding foreign proteins that are not found in plants (e.g., the mammalian fat-inducing transmembrane protein 2, FIT2). Further, our recent proteomics analysis of non-seed LDs derived from avocado tissues, revealed a new class of LD-associated proteins (LDAPs) that are highly conserved in plants, and we are currently investigating the ability of these proteins to influence the formation and/or regulation of LDs in leaves of tobacco and Arabidopsis. We are also stacking together various genes and/or gene knockouts to determine which combinations are most effective for elevating oil content in plant leaves. Collectively, our research has important bioenergy applications, but also will help to unravel the complex machinery in eukaryotes that is deployed for the metabolism and maintenance of neutral lipids in LDs. Our results are also expected to stimulate new ideas about the dynamic interplay between lipid storage, mobilization and signaling in eukaryotic systems -- an understanding that will be important to achieving the broader BER goal of "sustainable and affordable production of renewable biofuels in an environmentally conscientious manner".

Publications supported by award (2013-14):
1. Chapman KD, Dyer JM, Mullen RT. (2013) Commentary: why don't plant leaves get fat? Plant Science 207:128-34. doi: 10.1016/j.plantsci.2013.03.003.
2. Park S, Gidda SK, James CN, Horn PJ, Khuu N, Seay DC, Keereetaweep J, Chapman KD, Mullen RT, Dyer JM. (2013) The α/β hydrolase CGI-58 and peroxisomal transport protein PXA1 coregulate lipid homeostasis and signaling in Arabidopsis. Plant Cell 25(5):1726-39. (Faculty of 1,000 recommended)
3. Horn PJ, Silva JE, Anderson D, Fuchs J, Borisjuk L, Nazarenus TJ, Shulaev V, Cahoon EB, Chapman KD (2013) Imaging heterogeneity of membrane and storage lipids in transgenic Camelina sativa

seeds with altered fatty acid profiles. Plant Journal, 76(1):138-50. doi: 10.1111/tpj.12278. Epub 2013 Aug 5.
4. Horn PJ, James CN, Gidda SK, Kilaru A, Dyer JM, Mullen RT, Ohlrogge JB, Chapman KD. (2013) Identification of a new class of lipid droplet-associated proteins in plants. Plant Physiology, 162(4):1926-36. doi: 10.1104/pp.113.222455.
5. Gidda SK, Watt S, Collins-Silva J, Kilaru A, Arondel V, Yurchenko O, Horn PJ, James CN, Shintani D, Ohlrogge JB, Chapman KD, Mullen RT, Dyer JM (2013) Lipid droplet-associated proteins (LDAPs) are involved in the compartmentalization of lipophilic compounds in plant cells. Plant Signaling and Behavior, 8(11). pii: e27141.
6. Chapman KD, Dyer JM, Mullen RT (2013). Deciphering the role of CGI-58 in lipid regulation: more than one way to trim the fat? ASBMB Today 12(11):26-27.
7. Park S, Keereetaweep J, James CN, Gidda S, Chapman KD, Mullen RT, Dyer JM (2014) CGI-58, a key regulator of lipid homeostasis and signaling in plants, also regulates polyamine metabolism. Plant Signaling and Behavior, 9(2):e27723.
8. Vanhercke T, El Tahchy A, Liu Q, Zhou XR, Shrestha P, Divi UK, Ral JP, Mansour MP, Nichols PD, James CN, Horn PJ, Chapman KD, Beaudoin F, Ruiz- López N, Larkin PJ, de Feyter RC, Singh SP, Petrie JR (2014) Metabolic engineering of biomass for high energy density: oilseed-like triacylglycerol yields from plant leaves. Plant Biotechnology Journal 12(2):231-9. doi: 10.1111/pbi.12131.
9. Yurchenko OP, Park S, Ilut DC, Inmon JJ, Millhollon JC, Liechty Z, Page JT, Jenks MA, Chapman KD, Udall JA, Gore MA, Dyer JM (2014) Genome-wide analysis of the omega-3 fatty acid desaturase gene family in Gossypium. BMC Plant Biology. 18;14(1):312.
10. Horn PJ, Chapman KD (2014) Lipidomics in situ: insights into plant lipid metabolism from high resolution spatial maps of metabolites. Progress in Lipid Research, 54:32-52.
11. Horn PJ, Sturtevant D, Chapman KD (2014) Modified oleic cottonseeds show altered content, composition and tissue-specific distribution of triacylglycerol molecular species. Biochimie, 96:28-36.

This research is supported by the U.S. Department of Energy, Office of Science, Division of Biological and Environmental Research (BER; DE-SC0000797).

178. Nutrient Cycling for Biomass: A multi-omics approach to model ectomycorrhizal regulatory networks

Geetika Trivedi1*(gt0001@uah.edu), Avinash Sreedasyam,2 Bich Tran,1 Stefanie Lasota,3 Alistair Rogers,3 Landon S. Wilson,4 Helen Kim,4 Stephen Barnes,4 Sarah Zerbs,5 Peter E. Larsen,5 Frank R. Collart,5 and Leland J. Cseke1

1The University of Alabama in Huntsville, Huntsville, AL; 2HudsonAlpha Institute for Biotechnology, Huntsville, AL; 3Brookhaven National Laboratory, Upton, NY; 4The University of Alabama at Birmingham, Birmingham, AL; 5Argonne National Laboratory, Lemont, IL

Project Goals: This project is focused on developing system-scale models at the symbiotic interface between ectomycorrhizal fungi (such as Laccaria bicolor) and tree species (such as poplar) in response to environmental nutrient availability/biochemistry. A multiple "omics" approach is implemented that integrates next generation sequencing transcriptomics, proteomics, biochemical analyses and ChIP-Seq analyses to construct ectomycorrhizal regulatory networks and computational modeling approaches to predict how atmospheric carbon is sequestered as plant and/or subsurface fungal biomass.

Poster Abstract: Recent advancement in high-throughput ‘omics’ approaches has led to the precise quantification of all classes of biological molecules, from genes through mRNA to proteins and metabolites. Though each class of ‘omics’ data provides important information, no single analytical approach provides comprehensive insight into the molecular mechanisms of biological phenomena. Therefore, multi-‘omics’ approaches (i.e. integration of multiple layers of information) allow a better understanding of the complex molecular interactions that orchestrate biological processes.
Our lab has been using such approaches to build a predictive understanding of the molecular mechanisms underlying symbiotic plant-fungal associations called ectomycorrhiza (ECM). Such interactions play an essential part in the regulation of soil nutrient cycling and subsequent carbon management in a variety of tree species. However, the complexity and variation of such systems in natural environments has hindered the accurate assessment of the factors that signal, establish and maintain these interactions. Consequently, our project employs a controlled laboratory Populus tremuloides X Laccaria bicolor in vitro ECM model system to explore nitrogen utilization and carbon management under nutrient limiting conditions.

RNA sequencing (RNA-Seq) and ChIP sequencing (ChIP-Seq) were employed to produce global transcript abundance data along with increased resolution on transcriptional regulation. Our developed methods are among the first to successfully use such techniques for an ectomycorrhizal system. The transcriptomics data generated a mycorrhizal metabolome model that predicts the synthesis of specific nitrogen rich compounds by Laccaria in response to nitrogen limitation. Such compounds or their metabolites may in turn be used for the growth of tremuloides in exchange for photosynthetically derived sugars. The data also identified regulated metabolic pathways, transporter proteins and transcription factors (TFs) that are implicated in nutrient assimilation and carbon management. Sixteen TFs having potential roles in mycorrhiza-specific carbon and nitrogen flux metabolism were selected for ChIP-Seq. The results from ChIP-Seq identified several enriched binding motifs that may be indicative of new classes of the transcription factor target genes. For example, 961 loci were identified as potential targets of the selected MADS-Box regulators, and the expected CC[A/T]6GG motif (termed the CArG-box) was present in the promoters of many of these target loci, thus validating the ChIP- Seq pipeline in P. tremuloides. However, other equally well-represented motifs were also identified in the promoters of other possible target genes, including some genes that are involved in biochemical and metabolic pathways associated with cellulose and lignin biosynthesis during plant growth and development.
As a means to further interrogate the sequencing data, biochemical analyses were performed to quantify key C, N and P metabolites in the leaves, stems and roots of P. tremuloides subjected to nutrient limiting

conditions during interaction with Laccaria. Metabolite levels were then correlated with the transcriptomics data to delineate ~200 biochemical pathways that play roles in symbiosis-linked nutrient uptake and resulting biomass development. The pathways involved in starch and sucrose biosynthesis as well as nitrate reduction were especially sensitive to mycorrhizal interaction under varying nutrient stresses. Here, optimum nitrogen availability produced trees that directed more carbohydrates in form of soluble sugars towards the mycorrhizal roots, leading to healthier colonization and higher tree biomass. However during nitrogen limiting conditions, the carbohydrate supply to the roots was reduced as the trees put more sugar resources into the production of sucrose and starch within the shoot tissues. Further correlations of these biochemical pathways with the ChIP-seq data and quantitative proteomics are currently underway.

As transcript expression levels do not always correlate well with protein expression levels, our project is using nanoLC-ESI-SWATH-MS to fill the gap in our understanding of protein activities during mycorrhizal interaction under nutrient limiting conditions. Sequential Windowed Acquisition of all THeoretical MS (SWATH-MS) is a quantitative proteomics approach that detects large numbers of peptides with high sensitivity, quantitative accuracy and reproducibility. In order to implement this technology for P. tremuloides, we developed a P. tremuloides protein database with 35,328 unique peptide sequences using our transcriptomics data (>780 million RNA-Seq short reads), significantly improving the protein identifications for this tree species by 48% over the P. trichocarpa database alone. This subsequently allowed the identification and quantification of proteins within crude protein extracts from the various tissues of P. tremuloides. For example, SWATH-MS analysis of the leaf proteome successfully detected and quantified 6,250 peptides, representing the abundance levels of ~2,000 proteins under each fungal and nutrient condition. Our analysis revealed that 626 of these proteins have differential abundance in response to different nitrogen availability, whereas only 38 proteins show differential abundance in response to mycorrhizal interaction. This highlights the importance of nitrogen availability on the overall system. In addition, using NonLinear Dynamics software, 2D gel electrophoresis (IEF/SDS-PAGE) detected both post-translational modifications and/or changes in protein abundance for a total of 62 proteins that were significantly different in each of the experimental groups. Thus, the new technologies and continuing computational approaches that have been implemented during the course of this project have not only identified novel biochemical and regulatory pathways but also implicate specific protein expression and post-translational modifications as playing important roles in the molecular mechanisms of how plants, including trees, respond to changes in environmental nutrient conditions.

179. New tools for studying water deficit responses in roots

Jose Sebastian1, Muh-Ching Yee1, Ruben Rellan-Alvarez1, Rita Nieu2, Amanda Shrager3, John Vogel2, Julin Maloof3, Thomas P. Brutnell4 and José R. Dinneny1 (jdinneny@carnegiescience.edu)

1Carnegie Institution for Science, Department of Plant Biology, Stanford, CA, 2US Department of Energy, Joint-Genome Institute, Walnut Creek, California, 3University of California, Davis, CA, 4Donald Danforth Plant Science Center, St. Louis, Missouri

http://sviridis.org/, http://foxmillet.org/

Project Goals: "A Systems-Level Analysis of Drought and Density Response in the Model C4 Grass Setaria viridis". The specific aims of the proposal are to: 1) Identify QTL for the effect of drought and density on biomass and seed yield components of Setaria. 2) Conduct in–depth physiological profiles in roots and leaves of a subset of selected lines 3) Integrate datasets and develop metabolic and gene networks for Setaria 4) Develop transformation technologies for Setaria viridis 5) Functionally examine the role of candidate genes deduced by network models; and 6) Develop protocols and best practices for monitoring gene flow in transgenic Setaria. To achieve these aims we will produce one of the most extensive molecular characterizations of plant growth in the field to date, generating several million data points that will be collected from physiological and molecular genetic studies. We will develop novel informatics models and network tools that will guide future molecular characterization in S. viridis and guide breeding efforts in major feedstock targets.

Droughts are the consequence of common weather patterns that strike important agricultural areas around the world and affect food security. Water deficit and elevated temperatures are common stresses that plants experience during drought and have been extensively studied in field, green house and laboratory settings. Water is not uniformly distributed in soil and plants that develop deeper root systems are known to be able to better withstand lack of water by avoiding dry soil.

Experimental conditions that mimic water deficit in the lab often use gel-based media and dissolved osmolytes that reduce water potential. These conditions usually elicit a reduction in root growth, which allows the plant to economize resources. Such tolerance mechanism may be less important to field-grown plants where sustaining biomass and productivity are important targets. A new integrated plant growth and imaging system termed GLO-Roots (Growth and Luminescence Observatory for Roots) has recently been developed that allows root systems to be characterized in soil grown plants. Physiologically relevant water-deficit conditions have been designed that cause drying of the soil column from the top down, which causes roots to grow faster and deeper. Such changes in root architecture likely represent adaptive avoidance responses that facilitate water uptake from deep in the soil column. Use of dual-color luciferase reporter systems allow root structure and gene expression to be studied simultaneously and provide tools for characterizing the molecular-genetic pathways controlling changes in growth and gravitropism observed during water deficit. We have adapted GLO-Roots successfully in Arabidopsis, Brachypodium and Tomato and will soon begin studies in Setaria viridis.

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

180. Repurposing the Yeast Peroxisome to Compartmentalize Engineered Metabolic Pathways

William C. DeLoache,1 Zachary N. Russ,1* (zruss@berkeley.edu), John E. Dueber1,2

1University of California, Berkeley; 2Lawrence Berkeley National Labs, Berkeley, California

Project Goals: Engineered metabolic pathways often suffer from undesirable interactions with the production host's native cellular processes. Evolution has solved the problem of metabolic crosstalk by segregating distinct cellular functions into membrane---bound organelles. The goal of this project is to repurpose one of these organelles – specifically the peroxisome of Saccharomyces cerevisiae – for compartmentalizing heterologous metabolic pathways. Towards this goal, we are working to 1) improve the targeting efficiency of non---native enzyme cargo to the peroxisome, 2) determine the natural chemical composition of the peroxisomal lumen, 3) establish methods for altering this chemical composition to suit the needs of new enzymatic pathways, and 4) demonstrate successful compartmentalization of a model pathway. Ultimately, this work will contribute to the development of a synthetic organelle that can limit metabolic crosstalk and improve production efficiency for a variety of engineered pathways.

We selected the peroxisome as the basis for a synthetic organelle for a few key reasons. First, and most importantly, loss of the peroxisome does not adversely affect S. cerevisiae growth provided that long chain fatty acids are not used as the sole carbon source. Second, while peroxisomes occupy only a small fraction of the total cell volume under normal growth conditions, they are capable of expanding dramatically when induced. In some methanol utilizing yeasts such as Pichia pastoris and Hansenula polymorpha, methanol induction generates massive peroxisomes that often take up more than 70 percent of the total cell volume—an appealing prospect for high---flux engineered pathways. Third, unlike other organelles, the peroxisome's cargo protein is imported directly from the cytosol in the folded state, reducing the probability of protein misfolding for non---native cargo. Finally, many organisms have already repurposed the peroxisome to perform specialized functions—evidence of the organelle's inherent plasticity.

We have focused on working with and improving peroxisomal targeting tags by importing the bacterial enzyme vioE, which produces the easily detectable green pigment prodeoxyviolacein (PDV). When vioE is expressed in the cytosol of S. cerevisiae, it generates large amounts of PDV, resulting in green colonies. However, when sequestered in the peroxisome, vioE is unable to access its substrate, and no PDV is produced. Using this assay, we found that C---terminal PTS1 peroxisomal import tag is many times more efficient than N---terminal PTS2 peroxisomal import tag. In fact, the C---terminal PTS1 tag proved to be more efficient at sequestering vioE than canonical targeting tags for the vacuole, mitochondrion, or even extracellular secretion. Even with a PTS1 tag, however, we still observed detectable levels of cytosolic vioE when the enzyme was expressed at the very high levels utilized by most metabolic engineers. To further improve the import efficiency of PTS1---tagged cargo, we constructed a randomized library of six amino acids preceding the native PTS1 tag and screened for colonies that showed reduced levels of vioE in the cytosol. The output of this screen showed a pronounced trend for improved import when basic residues were preceding the PTS1 tag. Based on these results, we now have a sequence---optimized targeting tag that is capable of importing protein to the peroxisome extremely efficiently. This tag is also modular—we have shown that it maintains its efficiency when fused to a variety of cargo proteins.

To further characterize this improved tag, we used it to evaluate the rate and capacity of protein import of fluorescent proteins and vioE using the improved tag. We found that peroxisomal import can be used to

clear a cell of cytosolic fluorescent proteins in an hour, and that 3---10% of a cell's total protein can be encapsulated in unmodified peroxisomes. The improved tag also made possible the encapsulation of a two---enzyme vioB/vioE pathway. Finally, we demonstrated use of this modular tag on a number of essential proteins and found the tag enabled inducible growth arrest even when the naïve tag failed. With this enhanced tag, we have achieved efficient, timely targeting of protein cargo. However, our strategy also demands that we have control over the metabolite pool within the peroxisome.

To address this problem, we are attempting to determine which metabolites are naturally present in the peroxisome and to understand how they get there. Previous studies on the chemical composition of the peroxisome have employed either fluorescent biosensors or purified peroxisomes. We have opted to instead utilize an enzyme---based, in vivo approach that is more sensitive than fluorescent biosensors and more physiologically relevant than purified systems. Our assay uses our optimized PTS1 tag to target enzymes of interest to the peroxisome. By comparing enzymatic activity with and without peroxisomal localization, we can determine whether the substrate for each enzyme is present in the peroxisome. Thus far, we have found that a surprisingly high number of small metabolites are present in the lumen of the peroxisome, and have conducted experiments to give an approximate measure of peroxisomal permeability. We believe these metabolites get through the peroxisomal membrane by freely diffusing through a non---specific pore protein— something that has been hypothesized in the literature but has yet to be identified or confirmed in vivo. Our current efforts are focused on identifying this pore protein so that we can knock it out in our engineered system.

181. New Capabilities at the Center for Structural Molecular Biology at ORNL

Paul Langan,1 Volker Urban,1 Sai Venkatesh Pingali,1 Shuo Qian,1 Qiu Zhang,1 Kevin Weiss,1 Hugh O'Neill1* (oneillhm@ornl.gov)

1Oak Ridge National Laboratory, Oak Ridge, Tennessee

http://www.csmb.ornl.gov/

Project Goals: The Center for Structural Molecular Biology (CSMB) at Oak Ridge National Laboratory (ORNL) operates a small angle neutron scattering (SANS) instrument, called Bio-SANS, which is located at the High Flux Isotope Reactor. Bio-SANS is a 35m medium resolution SANS instrument that was specifically designed for biological studies and serves a vibrant user community of academic, industrial, and government users. Biological data from the SANS facility has the highest signal-to-noise ratio of any SANS instrument in this country. The Bio-Deuteration Laboratory (BDL) has significant strategic impact on neutron scattering at ORNL, providing benefits not just in higher quality and throughput of experiments but in extending the range, scale and complexity of biological problems that can be addressed. The CSMB complements other DOE OBER facilities for structural biology and supports studies of complex biological systems including biomolecular complexes, biomaterials, and biomass structure as part of the DOE Genomic Science Program.

The Bio-SANS instrument is ideally suited for studies of biomacromolecules including proteins, DNA/RNA, lipid membranes and other hierarchical complexes.1 The recently upgraded Bio-SANS detector has count rate capabilities (>106 Hz) that enable utilization of the full potential of the high neutron flux from the cold source. In addition, we have developed a series of new sample environment capabilities that open untapped opportunities for the studies of biological systems using neutrons. These include a pressure cell to monitor chemical reactions in situ such as biomass pretreatment studies, a multi-position sample holder with rotational (tumbling) capability especially useful for studying suspensions, a humidity-controlled chamber critical for membrane studies and a flow cell for systems that partition to multiple phases (e.g., microemulsions) with additional capability of flowing one or two phases during measurement. Furthermore, we now can perform grazing-incident SANS in conjunction with a humidity chamber for studies of biomembranes and substrate-supported biosensors. The implementation of additional data reduction software, open source MantidPlot and streamlined operations together enhance the efficiency of the neutron scattering experiments for users. Recent upgrades in the BDL include a parallel bioreactor system that allows high-density cell growth with precise control and monitoring of dissolved oxygen, pH, agitation, and feeding rates and a second preparative scale bioreactor system for large-scale fermentations. Other new laboratory capabilities include a Rigaku single crystal diffractometer, two liquid handling robots for preparing crystallization screens, and a suite of incubators for temperature controlled protein crystallization.

Reference

1. W. T. Heller, V. S. Urban, G. W. Lynn, K. L. Weiss, H. M. O'Neill, S. V. Pingali, S. Qian, K. C. Littrell,
2. Y. B. Melnichenko, M. V. Buchanan, D. L. Selby, G. D. Wignall, P. D. Butler and D. A. Myles "The Bio-SANS Instrument at the High Flux Isotope Reactor of Oak Ridge National Laboratory" J. Appl. Cryst. 47, 1238-1246 (2014).

The Center for Structural Molecular Biology (CSMB) is supported by the Office of Biological and Environmental Research Project ERKP291, using facilities supported by the U. S. Department of Energy (DOE), managed by UT-Battelle, LLC under contract No. DE-AC05-00OR22725. Research at Oak Ridge

National Laboratory's High Flux Isotope Reactor and the Spallation Neutron Source is sponsored by the Scientific User Facilities Division, Office of Basic Energy Sciences, U.S. DOE.

182. Risk and Escape Policies, Perspectives, and Practices: Issues and Implications for Biosystems Design R&D on Microbes, Algae, and Plants

Amy K. Wolfe1* (wolfeak@ornl.gov) and Savannah C. Stelling2

1Oak Ridge National Laboratory, Oak Ridge, Tennessee, 2 Oak Ridge National Laboratory, Oak Ridge, Tennessee via Oak Ridge Associated Universities

Project Goals. This project aims to identify circumstances that influence the human health and environmental risks associated with biosystems design research and development (R&D). Inquiries focus on research practices that help assure that new biological entities are safe to humans and the environment. We investigate the array of known and potentially unknown risks, and the potential for unintended adverse consequences, from two main categories of perspectives. One category is the set of perspectives of key players in biosystems design R&D, particularly scientists engaged in that R&D and biosafety professionals. The other category is the varied physical and institutional contexts in which biosystems design R&D takes place. Through analyses of research practices associated with both the "doing" of R&D as well as the development of safe biological organisms and systems, we seek to identify opportunities for avoiding or managing health and environmental risks.

The DOE Office of Science rationale for its investment in biosystems design (synthetic biology) R&D includes the following statement. "The merging of biology, chemistry, physics, and engineering has the potential to transform fundamental and applied science by shedding light on the basic principles of biological system organization and evolution…." (http://genomicscience.energy.gov/biosystemsdesign/). However, "the potential to transform" also carries the potential to create human health and environmental risks. The manifestations of such risks, of course, are conditioned by circumstances that affect their extent and impact—circumstances like those embedded in varying R&D contexts. Our research looks at linkages between R&D contexts and risks from a social and institutional perspective. That is, we investigate how components of R&D context (research setting, research approach, and organism studied, etc.) shape real- world research practices in ways that may contribute to or reduce the environmental and health risks associated with biosystems design R&D. We place particular emphasis on circumstances that inadvertently amplify or create risks during the course of R&D and with regard to the development of new biological organisms or systems.

To achieve our goals, it is essential to understand concepts of "risk," "containment," and "escape." There are ambiguities and differences associated with each of these terms when they are applied to biosystems design R&D. This kind of analysis helps to illuminate and understand connections between research practices and risks—whether during the conduct of R&D or with regard to the biological entity being designed. For example, we sort our information by the three broad categories of organisms that are the subject of DOE-funded research: microbes, plants, and algae. It may seem obvious that both human health and environmental risks associated with biosystems design R&D vary according to the type of organism at issue. However, a surprisingly large subset of the risk-related literature either fails to distinguish among organisms or implicitly seems to refer only to one category of organism (e.g., microbe). "Containment" translates into divergent terms when thinking about microbes versus plants versus algae; open- versus enclosed research settings; and a tool of research (e.g., bioreactor) versus a strategy for risk management. "Escape" becomes a muddied topic when comparing different circumstances. Take, for example, horizontal gene transfer (HGT), which sometimes is undesirable form of escape but sometimes is a desired research outcome.

We have begun to "unpack" these terms by analyzing them from multiple standpoints. Our data sources are published scientific literature, regulatory and biosafety guidelines, and interviews we conduct. This

information guides our inquiries about research practices and helps us analyze where gaps may exist.

This research is funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research.

183. Elucidating Cyanobacterial Recycling of Microbial Mat Extracellular Matrix

Rhona K.Stuart1* (stuart25@llnl.gov), Mary Lipton2, Xavier Mayali1, Jackson Z. Lee3,4, Craig Everroad4, Brad M. Bebout3, Jennifer Pett-Ridge1, Peter K. Weber1, Michael Thelen1.

1Lawrence Livermore National Laboratory, Livermore, CA, 2Pacific Northwest National Laboratory, Richland, WA, 3Exobiology Branch, NASA Ames Research Center, Moffett Field, CA, 4Bay Area Environmental Research Institute, Sonoma, CA.

Project Goals: The LLNL Biofuel SFA investigates systems biology of complex microbial communities relevant to bioenergy production. To understand nutrient cycling and potential biofuel production in complex microbial communities we employ an integrated analysis of energy flow using multi-scale approaches including biogeochemical, stable isotope probing, metagenomic/transcriptomic, proteomic/metabolomic and computational analyses. Our ultimate goal is the development of multi-scale models that can predict ecological and biochemical relationships within multi-trophic microbial systems.

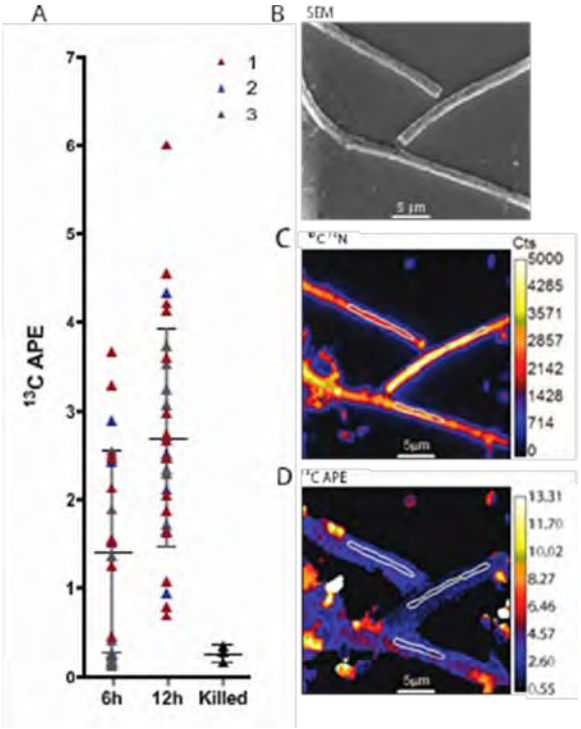


Figure 1: Cyanobacterial trichomes incubated with ^{13}C-labeled EPS become enriched in ^{13}C. NanoSIMS analysis of ESFC-1 trichomes incubated with ^{13}C EPS. (A) Points represent ^{13}C APE (atom percent excess, % over background) values for cyanobacterial trichomes from three biological replicates (1,2 or 3) at 2 time points following addition of ^{13}C label. "Killed" represents killed control cells that were fixed before incubation. (B) Scanning electron microscopy image of representative ESFC-1 trichomes after 12 hours incubation with ^{13}C EPS. (C) ^{12}C ^{14}N NanoSIMS image of same trichomes and (D) ^{13}C APE NanoSIMS image.

Cyanobacterial carbon excretion is crucial to carbon cycling in many microbial communities, but the nature and bioavailability of the carbon excreted is dependent on its physiological function, which is often unknown. Hypersaline laminated photosynthetic mats are an excellent model system for the study of carbon flow in a complex community because they are sustained primarily by photosynthesis in a relatively small- scale closed system. These mats have a large reservoir of carbon in the extracellular matrix, but how cyanobacterial matrix production is regulated and who consumes it is poorly understood. To better understand cyanobacterial carbon excretion, we examined the macromolecular composition of the extracellular matrix of microbial mats from Elkhorn Slough in Monterey Bay, CA USA. In collaboration with PNNL's Pan-Omics group we characterized the mat exoproteome and discovered predominantly cyanobacterial proteins, many with predicted roles in breakdown of organic matter, and detected enzymatic activities which indicate capacity for cyanobacterial degradation of matrix organic matter. To further explore the regulation of these breakdown abilities, we characterized a biofilm-forming cyanobacterial isolate from Elkhorn Slough, ESFC-1, that has a similar extracellular composition to that of our field mats, providing us with a model culture. Using this culture, we identified exoproteins that change in abundance over a diel cycle, and exoproteins that were dark stress induced, suggesting light- dependent regulation of

matrix material breakdown. We then used high resolution imaging mass-spectrometry (NanoSIMS) to characterize EPS-carbon re-uptake in ESFC-

Our results demonstrated light-dependent, rapid uptake of EPS-associated carbon by ESFC-1 in culture. Based on these findings, we propose that mat Cyanobacteria store and recycle their organic carbon from the mat extracellular matrix. Cyanobacteria are such a large percentage of the biomass in the upper phototrophic layer of the microbial mats, that their re-uptake of organic carbon has the potential to re-define carbon availability and turnover in these systems.

Funding was provided by the U.S. Department of Energy (DOE) Genomic Science Program under the LLNL Biofuels SFA, FWP SCW1039. 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 at Pacific Northwest National Lab supported by the OBER/GSP under the Pan-Omics project.

184. Hydration Control of Cellulose Surface Structure and Dynamics

Loukas Petridis1* (petridisl@ornl.gov), Sai Venkatesh Pingali,1 Daisuke Sawada,1 Volker Urban,1 Hugh M. O'Neill,1 Arthur J. Ragauskas,1,2 Barbara R. Evans,1 Jeremy C. Smith,1,3 Paul A. Langan,1 and Brian H. Davison1

1Oak Ridge National Laboratory, Oak Ridge, Tennessee; 2University of Tennessee, Department of Chemical and Biomolecular Engineering and the Department of Forestry, Wildlife, and Fisheries, Knoxville; 3University of Tennessee, Department of Biochemistry and Cellular and Molecular Biology, Knoxville

http://www.ornl.gov/science-discovery/clean-energy/research-areas/systems-biology/bioenergy/dynamic-visualization-of-lignocellulose

Project Goals: Lignocellulosic biomass comprises the vast majority of biomass on Earth and has the potential to play a major role in generation of renewable biofuels if cost-effective conversion can be achieved. Largely composed of plant cell walls, it is a complex biological composite material that is recalcitrant to the structural deconstruction and enzymatic hydrolysis into sugars that is necessary for fermentation to bioethanol. The Scientific Focus Area in Biofuels is developing "Dynamic Visualization of Lignocellulose Degradation by Integration of Neutron Scattering Imaging and Computer Simulation" for multiple-length scale, real-time imaging of biomass during pretreatment and enzymatic hydrolysis. This is providing fundamental information about the structure and deconstruction of plant cell walls that is needed to drive improvements in the conversion of renewable lignocellulosic biomass to biofuels.

Cellulose recalcitrance to enzymatic hydrolysis is determined in part by the structural features of cellulose surface, which influence the binding of hydrolytic enzymes, and the mechanical work needed to pull apart cellulose strands for catalytic cleavage. Cellulose in biomass is found in a hydrated state, and it is therefore important to understand how hydration influences its surface structure and dynamics. By combining neutron scattering experiments and molecular dynamics simulations we characterized the nanosecond motions in cellulose, their dependence on temperature and hydration, and how they affect the surface order of the microfibril.1 The experiments reveal that samples hydrated to 20% w/w exhibit a higher average mean-square displacement above ~240K than dry samples do. The molecular dynamics (MD) simulation reveals hydroxymethyl groups on the surface of the fibril have the highest fluctuations, and these increase significantly when cellulose is hydrated due to faster relaxation of the hydroxymethyl/water hydrogen bond network. Although in the MD simulations the hydroxymethyl groups in the cellulose core are always found in the crystallographically determined trans-gauge conformations, hydration leads to increased disorder in the hydroxymethyl conformation at the cellulose surface. The detailed characterization obtained describes how hydration-dependent increased fluctuations are connected with hydroxymethyl disorder that disrupts the cellulose hydrogen- bond network and makes cellulose more susceptible to enzymatic attack.

Reference

1. L. Petridis, H. M. O'Neill, M. Johnsen, B. Fan, R. Schulz, E. Mamontov, J. Maranas, P. Langan, and J. C. Smith, "Hydration Control of the Mechanical and Dynamical Properties of Cellulose," Biomacromolecules 15(11), 4152-4159 (2014).

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 Office of Biological and Environmental Research in the DOE Office of Science.

185. Investigation of Water Dynamics in a Model Cellulosic System

Hugh O'Neill,1 Loukas Petridis,1 Sai Venkatesh Pingali,1 Barbara R. Evans,1 Junhong He,1 Eugene Mamontov,1 Brian H. Davison1* (davisonbh@ornl.gov), Volker Urban,1 Jeremy Smith,1,3 Paul Langan,1 and Arthur J. Ragauskas1,2

1Oak Ridge National Laboratory, Oak Ridge, Tennessee; 2University of Tennessee, Department of Chemical and Biomolecular Engineering and the Department of Forestry, Wildlife, and Fisheries, Knoxville; 3University of Tennessee, Department of Biochemistry and Cellular and Molecular Biology, Knoxville
http://www.ornl.gov/science-discovery/clean-energy/research-areas/systems-biology/bioenergy/dynamic-visualization-of-lignocellulose

Project Goals: Lignocellulosic biomass comprises the vast majority of biomass on Earth and has the potential to play a major role in generation of renewable biofuels if cost-effective conversion can be achieved. Largely composed of plant cell walls, it is a complex biological composite material that is recalcitrant to the structural deconstruction and enzymatic hydrolysis into sugars that is necessary for fermentation to bioethanol. The Scientific Focus Area in Biofuels is developing "Dynamic Visualization of Lignocellulose Degradation by Integration of Neutron Scattering Imaging and Computer Simulation" for multiple-length scale, real-time imaging of biomass during pretreatment and enzymatic hydrolysis. This is providing fundamental information about the structure and deconstruction of plant cell walls that is needed to drive improvements in the conversion of renewable lignocellulosic biomass to biofuels.

Knowledge of water dynamics and structure at cellulose surfaces is important to understand how pretreatment regimes alter cellulose structure and how cellulases bind to cellulose interrupting hydrogen bonding between cellulose chains as a first step in cellulose hydrolysis. Quasi-elastic neutron scattering measures the local motions of H atoms on a pico- to nanosecond timescale and is used for characterizing water dynamics on surfaces. We report on measurements performed using deuterated bacterial cellulose equilibrated with H2O. Deuterium-labeling of cellulose strongly attenuates its scattering signal revealing the scattering contribution of water associated with the fibers providing insight into the dynamics of water surrounding cellulose fibers with unprecedented detail. Our results show that at lower temperatures the surface water gradually becomes glass-like, shows progressively slowing down dynamics, which could no longer be observed with the resolution of our experiment below about ~220K. The temperature dependent elastic intensity scans of deuterated cellulose hydrated in H2O reveal at least two populations of water present near the surface of the hydrated cellulose; "surface water" confined on the surface of the cellulose fibrils that becomes appreciably mobile at ~220K and "intefibrilar water" present in nanoscale water pockets between cellulose fibrils that becomes mobile at ~250K. Based on our analysis of the elastic scan data, three temperature points, 230, 250K and 265K, were chosen for a more detailed study of the dynamics of the system. The 230K data set probes the dynamics of the surface water, which at this temperature does not display long-range translational diffusion. It becomes translational in the 250K data set, while the interfibrilar water is still frozen contributing only an elastic scattering signal. Finally, at 265K the melting of the interfibrilar water dominates the 265K spectra. The diffusion coefficients of the water at 250K and 265K were 0.85±0.04 × 10-10 m2sec-1 and 1.77±0.09 × 10-10 m2sec-1, respectively. This indicates that the water associated with cellulose is somewhat restricted compared to bulk super cooled water at 268K which has a self-diffusion coefficient 9.41 × 10-10 m2sec-1.1 This study provides new insight into the dynamics of water surrounding crystalline cellulose and will aid in developing models to understand the mechanism of cellulose breakdown and its interaction with water, acids, and cellulases in an effort to optimize the cellulose digestion process.

Reference

1. W. S. Price, H. Ide, and Y. Arata, Y., "Self-Diffusion of Supercooled Water to 238 K Using PGSE NMR Diffusion Measurements," J. Phys. Chem. A 103, 448–450 (1999).

Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy (DOE) under contract no. DE-AC05-00OR22725. This program is supported by the Office of Biological and Environmental Research in the DOE Office of Science.

186. Morphological Changes in the Cellulose and Lignin Components of Biomass Occur at Different Stages During Steam Pretreatment

Sai Venkatesh Pingali1* (pingalis@ornl.gov), Hugh O'Neill,1 Yoshiharu Nishiyama,1 Lilin He,1 Yuri B. Melnichenko,1 Volker Urban,1 Loukas Petridis,1 Paul Langan,1 and Brian H. Davison1

1Oak Ridge National Laboratory, Oak Ridge, Tennessee

http://www.ornl.gov/science-discovery/clean-energy/research-areas/systems-biology/bioenergy/dynamic-visualization-of-lignocellulose

Project Goals: Lignocellulosic biomass comprises the vast majority of biomass on Earth and has the potential to play a major role in generation of renewable biofuels if cost-effective conversion can be achieved. Largely composed of plant cell walls, it is a complex biological composite material that is recalcitrant to the structural deconstruction and enzymatic hydrolysis into sugars that is necessary for fermentation to bioethanol. The Scientific Focus Area in Biofuels is developing "Dynamic Visualization of Lignocellulose Degradation by Integration of Neutron Scattering Imaging and Computer Simulation" for multiple-length scale, real-time imaging of biomass during pretreatment and enzymatic hydrolysis. This is providing fundamental information about the structure and deconstruction of plant cell walls that is needed to drive improvements in the conversion of renewable lignocellulosic biomass to biofuels.

Understanding the structural changes in biomass during pretreatment is critical for increasing biomass conversion efficiencies. Currently biomass analysis techniques are limited to comparing biomass structure before and after pretreatment. However, it is extremely desirable to understand how the structure of lignocellulose changes during a pretreatment regime. Our previous work showed the formation of lignin globules and the aggregation of cellulose fibers in pretreated biomass. However, this work did not provide any insights into the processes that drive these structural changes. In order to address this knowledge gap we developed an approach to monitor morphological changes as they occur during steam explosion pretreatment (SEP) by placing a pressure reaction cell in a neutron beam and collecting time-resolved small-angle neutron scattering data.1 Changes to cellulose morphology occurred mainly in the heating phase of pretreatment, whereas changes in lignin morphology occurred mainly in the holding and cooling phases. During the heating stage, water is irreversibly expelled from cellulose macrofibrils as the elemental fibers coalesce. On the other hand, there is little change to lignin during the heating phase, but aggregates begin to appear and increase in number during the holding phase. Lignin aggregates continue to appear and increase in size during the cooling phase. These results support previous MD simulations that suggest that lignin aggregation occurs during SEP as soon as lignin has phase separated from hemicellulose.2 The results presented here indicate that this phase separation starts at ~140oC during SEP. This experiment demonstrates the unique information that in situ SANS studies of pretreatment can provide. This approach could be useful in optimizing the heating, holding and cooling stages of pretreatments to allow the exact size and nature of lignin aggregates to be controlled in order to enhance enzyme accessibility to cellulose and therefore the efficiency of biomass conversion.

References

2. S. Pingali, H. M. O'Neill, Y. Nishiyama, L. He, Y. B. Melnichenko, V. Urban, L. Petridis, P. Langan, and B. H. Davison, "Morphological Changes in the Cellulose and Lignin Components of Biomass Occur at Different Stages During Steam Pretreatment, Cellulose 21, 873–878 (2014).
3. P. Langan, L. Petridis, H. M. O'Neill, S.V. Pingali, M. Foston, Y. Nishiyama, R. Schulz, B. Lindner, B. L. Hanson, S. Harton, W. T. Heller, V. Urban, B. R. Evans, S. Gnanakaran, A. J. Ragauskas, J. C. Smith, and B. H. Davison (2014) "Common processes drive the thermochemical pretreatment of lignocellulosic biomass," Green Chem. 16, 63 (2014).

This research is funded by the Genomic Science Program, Office of Biological and Environmental Research, U.S. Department of Energy, under FWP ERKP752. The Center for Structural Molecular Biology (CSMB) and the Bio---SANS beam line is supported by the Office of Biological and Environmental Research, using facilities supported by the U. S. Department of Energy, managed by UT-Battelle, LLC under contract No. DE-AC05-00OR22725. This Research on Bio-SANS and GP-SANS at Oak Ridge National Laboratory's High Flux Isotope Reactor was sponsored by the Scientific User Facilities Division, Office of Basic Energy Sciences, Department of Energy.

187. Abiotic Stress Networks Converging on FT2 to Control Growth in Populus

Xiaoyan Sheng1, Rita Tiexiera1, Mihir Mandal1, Jason A. Holliday1 and Amy M. Brunner1* (abrunner@vt.edu)

1Virginia Tech, Dept. of Forest Resources & Environmental Conservation, Blacksburg, VA

Project Goals:
Woody feedstock plants such as Populus show substantial natural variation in growth responses to various environmental conditions; however, the complex, multi-genic nature of these traits slows breeding for optimal biomass production in specific climates and on marginal lands. The identification of FT2 as an integrator of multiple abiotic signals in Populus (Hsu et al. 2011) provides a focal point for discovering and integrating transcriptional, protein-DNA interaction (PDI) and protein-protein interaction (PPI) networks controlling growth in response to environmental conditions. This project is focusing on two abiotic conditions that are key determinants of growth—daylength and nitrogen (N) availability. Our overarching hypothesis is that uncovering convergent and divergent regulatory networks that control responses to these two abiotic stresses will advance genetic improvement for adaptation to one or a combination of abiotic environmental conditions. To elucidate these networks, transcriptomic responses to daylength and N availability will be studied in multiple tissue/organs of Populus trichocarpa Nisqually-1. These responses will be studied in time series experiments that encompasses both cessation and resumption of shoot growth. The effects of the daylength and nutrient regimes on RNA abundance will be quantified by RNA- Seq. Three types of libraries will be prepared from each sample: total RNA, small RNA, and degradome. This data will serve as the basis for identification of differentially regulated genes and global gene regulatory network construction. A second aim is to identify targeted PPI and PDI networks using yeast- two-hybrid (Y2H) and Y1H to both validate and extend the regulatory networks reconstructed from the transcriptome experiments. This aim will initially focus on identifying networks both upstream and downstream of FT2. Understanding the mechanisms that determine periods of growth and quiescence will identify specific control points that can be manipulated to maximize growth in different environments. Based on the network development, a few genes will be selected for functional analysis in transgenic poplar to validate their utility for enhancing plant growth and yield under different environmental conditions.

References

1. Hsu CY , Adams J, Kim H, No K, Ma C, Strauss SH, Drnevich J, Vandervelde L, Ellis JD, Rice M, Wickett N, Gunter LE, Tuskan GA, Brunner AM, Page GP, Barakat A, Carlson JE, dePamphilis C, Luthe DS, Yuceer C (2011) FLOWERING LOCUS T duplication coordinates reproductive and vegetative growth in perennial poplar. Proc Natl Acad Sci USA 108:10756-10761.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0012574.

188. Exploiting natural diversity to Identify alleles and Mechanisms of Cold Adaption in Switchgrass

Michael D. Casler1, Gautam Sarath2, and C. Robin Buell3* (buell@msu.edu)

1USDA-ARS, U.S. Dairy Forage Research Center, 1925 Linden Dr., Madison, WI; 2USDA-ARS, Grain, Forage and Bioenergy Research Unit, 137 Keim Hall, East Campus, UNL, Lincoln, NE; 3Michigan State University, Dept. of Plant Biology, East Lansing MI

Project Goals

One of the proposed mechanisms to increase switchgrass biomass is to grow lowland cultivars in more northern latitudes where they have higher yield potential, due in part, to their significantly later flowering time at these latitudes. However, lowlands are not adapted to the colder winter conditions in northern climes and exhibit low survival rates when grown at latitudes outside their adapted range. Many unadapted populations possess a small frequency of individuals that can survive the winter, indicating that, within the collective genetic diversity of the population, alleles are present that confer cold tolerance. If these alleles could be catalogued and converted into molecular markers, they would facilitate accelerated breeding and provide a mechanism to improve the efficiency of breeding switchgrass cultivars with high biomass and cold hardiness. Our hypothesis is that alleles favorable to cold hardiness will be enriched in individuals that survive over---wintering conditions outside their native hardiness zone and that these alleles will be under---represented in individuals within the population that fail to survive over---wintering. Using allele frequencies determined through bulk segregant exome capture sequencing, we will identify genes, and specifically alleles, important to cold hardiness in 15 lowland switchgrass populations. Using bulk segregant transcriptome sequencing, we will identify transcripts and regulatory RNAs associated with successful senescence and cold hardiness that will permit prediction and validation of key metabolites involved in cold adaptation. Data from this project will be central to furthering our understanding of cold tolerance in switchgrass and in identify alleles that that can be used in breeding programs to increase switchgrass biomass.

This project was supported by Agriculture and Food Research Initiative Competitive Grant no. 2014-67009-22310 from the USDA National Institute of Food and Agriculture

189. A novel poplar biomass germplasm resource for functional genomics and breeding

Isabelle M. Henry1* (imhenry@ucdavis.edu), Matthew S. Zinkgraf2, Andrew T. Groover1,2 and Luca Comai1

1 University of California, Davis, 2 Pacific Southwest Research Station, U.S. Forest Service, Davis

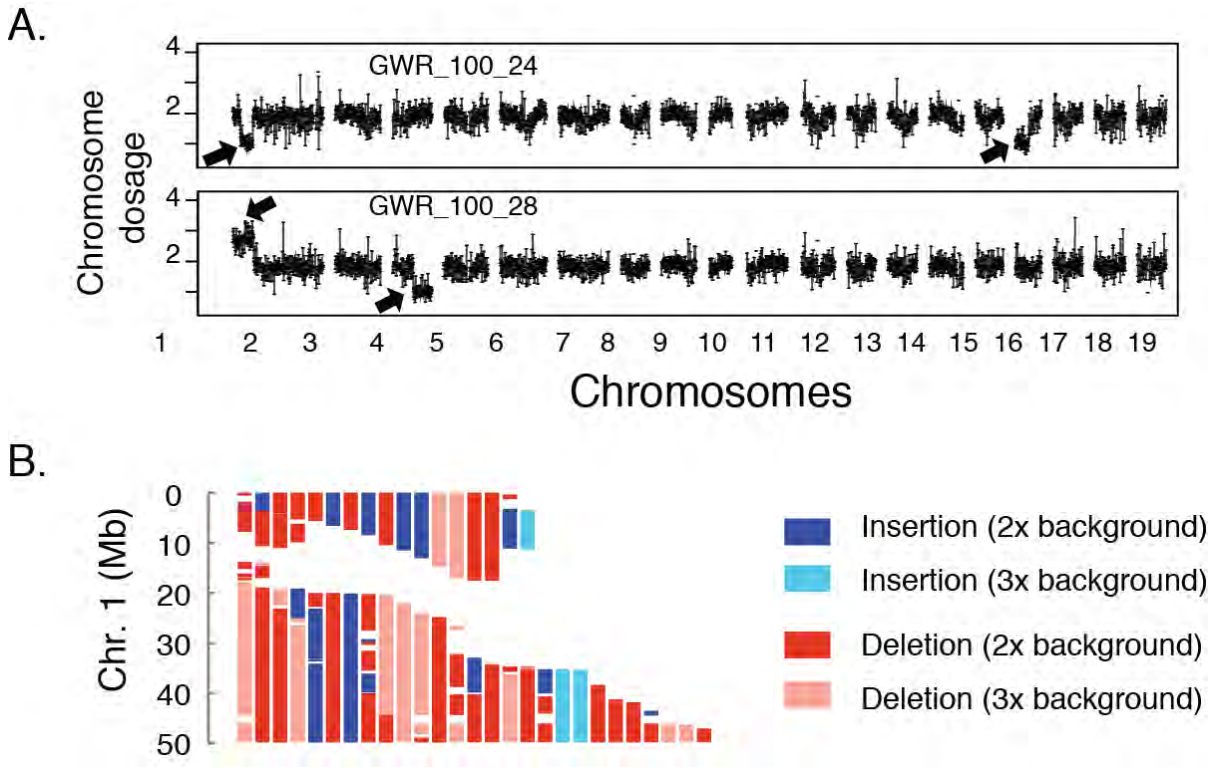
http://comailab.genomecenter.ucdavis.edu/index.php/Poplar
http://www.fs.fed.us/psw/programs/cb/staff/agroover/

Project Goals: We have produced a new population of interspecific poplar hybrids carrying insertions and deletions tiling the entire genome with dosage variation. This unique functional genomics resource can be used to investigate the role of gene dosage in poplar hybrid performance, and to identify and genes that contribute to poplar bioenergy traits. Our specific objectives are to i) finish characterize and maintain this resource ii) investigate the phenotypic effects of dosage change and iii) exploit the indel germplasm for functional genomics. Outputs will include a population of interspecific poplar hybrids carrying defined dosage variation and extensively characterized trait measurements. Possibly, our germplasm collection also includes cultivars directly usable for bioenergy applications. The approaches and tools developed can be generalized to other poplar hybrids and to germplasm development for other vegetatively-propagated species.

Interspecific hybridization is the most effective method for producing poplar clones for commercial applications, including biomass production. Previous research has shown that commercial F1 hybrids frequently exhibit dosage variation, i.e. variation in the number of pieces or entire chromosomes, and that these variants can have transgressive phenotypes desirable for biomass production. Poplar is an excellent system for studying and exploiting gene dosage because germplasm is propagated vegetatively for research and production, and advanced genome-enabled tools can be directly employed on the actual feedstock.

In the previous funding cycle, we have produced a population of interspecific hybrids carrying dosage lesions. Over 500 seedlings were subjected to high-precision dosage analysis using next-generation sequencing. Insertion and deletions were detected in approximately 55% of the progeny, and varied in length, position and number per individual, cumulatively covering >99% of the genome with an average of 10 dosage lesions (See Figure). Additionally, triploid individuals are also present within the population of dosage variant, which provide an additional layer of dosage variation. Preliminary phenotypic analysis confirmed that dosage variation is well tolerated and that samples with distinct morphological alterations are present within the population.

In this proposal, we will complete the development of this unique resource for the research community and start to use this germplasm as a functional genomics tool. Using precise phenotypic data, gene dosage information as well as associated gene expression information, we will investigate the role of gene dosage in poplar hybrid performance and identify chromosomal regions and genes that contribute to poplar bioenergy traits. Our efforts will add new, useful and unique tools that have both functional and practical applications for poplar breeding and is widely applicable to other biomass crops such as switchgrass, alfalfa, willow, and miscanthus.



Detection and characterization of dosage variants. A. For each sample, sequencing reads were aligned to the reference sequences and sorted into consecutive bins of 100 kbs, along the 19 chromosomes of the poplar genome (v3.0). Two different samples are represented here. Changes up or down in chromosome dosage correspond to the presence of an additional copy (insertion) or the deletion of a copy of a particular chromosome fragment, respectively (arrows). B. Lesions identified so far on chromosome 1. Lesions are tiled across the length of the chromosome. Dosage lesions vary in length (up to a whole chromosome) and a single individual carries up to 10 lesions. Deletions are most common (red and pink) but insertions are observed as well (blue). Background ploidy (2x or 3x) is indicated as well.

This research is supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grants no. DE-SC0007183.

190. Improving Cold Tolerance in Lowland Switchgrass

Katrien M. Devos1,2,* (kdevos@uga.edu), Srinivasa Chaluvadi3, Peng Qi1,2, Christian Schwoyer3, Joseph Young1, Ali Missaoui1, Orville M. Lindstrom4, Paul Schliekelman5, Jeffrey L. Bennetzen3

1 Institute of Plant Breeding, Genetics and Genomics, University of Georgia, Athens, GA 30602, USA 2 Department of Plant Biology, University of Georgia, Athens, GA 30602, USA
3 Department of Genetics, University of Georgia, Athens, GA 30602, USA 4 Department of Horticulture, University of Georgia, Griffin, GA 30223, USA 5 Department of Statistics, University of Georgia, Athens, GA 30622, USA

Project Goals:
The overall objective of our project is to investigate the effects of arbuscular mycorrhizal fungi (AMF) colonization on freezing tolerance and biomass production, to dissect the abilities of the host plant to interact with AMF and to tolerate sub-zero temperatures into their genetic components, and to identify expression quantitative trait loci (eQTL) for genes that are differentially regulated in upland and lowland ecotypes under cold-acclimatization and/or AMF symbiosis.

Switchgrass is a C4 grass that is native to the North-American prairies, and has potential as a cellulosic feedstock for bioenergy production. There are two ecotypes, upland and lowland, that vary in their predominant ploidy levels, overall morphology and regions of adaptation. The first aim of the project is to unravel the genetic basis of the cold-tolerance that is inherent to upland switchgrass but lacking in lowland ecotypes. Our second aim is to investigate whether colonization by arbuscular mycorrhizal fungi (AMF) provides a degree of cold-tolerance. The observed correlation between AMF levels in clonal genotypes grown in the field in Georgia and Oklahoma indicates that availability to AMF colonization is genetically determined. We have identified F1 sib plants from a cross between the upland genotype VS16 and the lowland genotype AP13 that vary greatly in their levels of colonization by AMF. Two crosses between different high and low AMF sibs have been generated and progeny have been planted in the field in Georgia (1 replicate). Since the F1 sibs carry 1 upland and 1 lowland chromosome, the generated populations should also be segregating for cold-tolerance. Efforts so far have been focused on (1) establishing an efficient protocol for producing large numbers of clonally identical plants which are needed for the freezing tests; (2) optimizing plant cultivation conditions to promote AMF colonization in the greenhouse; (3) initial testing of different freezing temperatures that best differentiate the cold tolerant VS16 and cold sensitive AP13 genotypes; and (4) developing a genotyping- by-sequencing methodology for use in switchgrass, an outbreeding tetraploid. We are now in a position to start the genetic mapping, and ramp up the freezing tests.

This research is supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0010743.

191. Genetic dissection of AM symbiosis to improve the sustainability of feedstock production

Maria J. Harrison* (mjh78@cornell.edu) and Zhangjun Fei

Boyce Thompson Institute for Plant Research, Tower Road, Ithaca, NY 14850

Project Goals:
Fertilizers comprise a significant proportion of biofeedstock production costs and large-scale fertilizer use has many negative environmental consequences. If biofeedstock production is to be economically viable and environmentally sustainable, increasing the efficiency of mineral nutrient acquisition is an important priority. One approach to this is to harness the soil microbiota, in particular, arbuscular mycorrhizal fungi, with which plants form stable mutualistic symbioses. Currently, the genetic basis of arbuscular mycorrhizal symbioses in feedstock species is largely unexplored. The overall goals of the proposed research are to determine which genes control development and functioning of AM symbiosis in feedstocks through studies of a feedstock model species, Brachypodium distachyon, as well as sorghum, a feedstock species.

Using B. distachyon, we aim to identify genes that are critical for development and functioning of AM symbiosis. This will be achieved through analysis of B. distachyon T-DNA insertion mutants and/or transgenic lines in which gene function has been down-regulated by RNAi or disrupted by genome editing using the CRISP/Cas9 system. B. distachyon T-DNA insertion mutants will be obtained from the DOE-funded insertion mutagenesis project, http://brachypodium.pw.usda.gov/TDNA/. Candidate genes have been selected based on previous transcriptional analyses of B. distachyon during AM symbiosis with three AM fungi, Glomus versiforme, Glomus intraradices (now Rhizophagus irregularis) and Gigaspora gigantea.

With support from the DOE Feedstock Genomics Program, Dr. Stephen Kresovich and colleagues are developing nested association mapping (NAM) populations for diverse sweet and cellulosic (bioenergy) sorghums. These NAM populations may also provide an opportunity to investigate variation in responsiveness to AM symbiosis in biofeedstock sorghum genotypes.
Relative to other crops, sorghum shows high nutrient use efficiency; however, in most production environments, fertilizers are still used and therefore improving nutrient capture would be economically and environmentally beneficial. As a first step, we will evaluate the parents of the NAM populations for responses to AM symbiosis including development of symbiosis, symbiotic mineral nutrient acquisition and increases in shoot biomass. For a selection of genotypes, root and shoot transcriptional responses to symbiosis will be analyzed by RNA-seq.
The profiles will provide insights into AM symbiosis and the effects of plant genotype on symbiotic gene expression.

References
1. Hong, J.J., Park, Y.S., Bravo, A., Bhattarai, K.K., Daniels, D.A., and Harrison, M.J. (2012). Diversity of morphology and function in arbuscular mycorrhizal symbioses in Brachypodium distachyon. Planta 236, 851-865.
2. Mainzer, L., Hong, J.J., Bhattarai, K.K. Battu, G., Daniels, D.A., Varala, K., Leveque- Tremblay, V., Park, Y.S., Hudson, M.E and Harrison, M.J. (2015), in preparation.

This research is supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0014260.

192. Abiotic stress induced Poplar transcriptome analyses.

Palitha Dharmawardhana1, Vindhya Amarasinghe1, Dylan Beorchia1, Teague Green1, Julian Preciado1, Sergei Filichkin1, Christopher M. Sullivan1,2, Matthew Geniza1, Stephen Rigoulot3 Amy Brunner3, Eric Beers4, Pankaj Jaiswal1* (jaiswalp@science.oregonstate.edu)

1Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR, 97331; 2Center for Genome Research and Biocomputing, Oregon State University, Corvallis, OR, 97331; 3Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA 24061; 4Department of Horticulture, Virginia Tech, Blacksburg, VA 24061

Project Goals:
1) To identify a whole genome-wide functional gene network (an interaction network of genes) for poplar using gene orthology based projections and identify subnetworks associated to abiotic stress tolerance and bioenergy related traits, 2) To identify a set of candidate genes which interact to produce abiotic stress resistant phenotypes, and to identify diagnostic genetic markers associated with the sub-networks, 3) To create an interactive online resource of regulatory and metabolic network of abiotic stress associated candidate genes and genetic markers for the poplar tree improvement community.

Abstract:
Improving abiotic stress of poplar clones used in plantation forestry is of significant economic importance. To better investigate the complex physiological, cellular, and molecular processes that underlie abiotic stress response in poplar we are investigating drought, heat, salt and cold stress induced transcriptomes from the reference Populus trichocarpa Nisqually-1 genotype. The mRNA was isolated from leaf, xylem and root tissue types at three time points including the control untreated and after early and prolonged stress time points. The RNA-Seq based transcriptome data is currently being analyzed for (1) identifying genes in the weighted co-expressed gene network, (2) compare the expressed gene network with the baseline poplar functional gene network constructed by orthology driven projection from the Arabidopsis probabilistic gene network, (3) reference genome guided and de-novo assembly of transcripts, (4) identifying gene loci and transcripts showing stress induced alternative splicing patterns with intron retention events (5) improving reference genome annotation by identifying novel gene loci and transcript isoforms. To validate key gene interactions and identify novel protein-protein interactions in an Y2H screening, cDNA libraries are being constructed from salt and drought stress treated poplar tissues. Following baseline experiments on reference Nisqually-1 genotype, we are in the process of conducting similar coexpression analyses studies of drought and salt resistant genotypes of poplar hybrids.

Publication:
1. Filichkin, Sergei A., Jason S. Cumbie, P. Dharmawardhana, P. Jaiswal, Jeff H. Chang, Saiprasad G. Palusa, A. S. N. Reddy, M. Megraw and Todd C. Mockler (2015). Environmental Stresses Modulate Abundance and Timing of Alternatively Spliced Circadian Transcripts in Arabidopsis. Molecular Plant 8(2): 207-227.

Funding:
This research project is supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0008570.

193. Genetic Control of Flowering in Switchgrass

Paul Grabowski1, Carl-Erik Tornqvist2, Megan Taylor3, Michael D. Casler1 and Yiwei Jiang3 * (yjiang@purdue.edu)

1USDA-ARS, U.S. Dairy Forage Research Center, Madison, WI; 2Great Lakes Bioenergy Research Center and Department of Agronomy, University of Wisconsin, Madison, WI; 3Department of Agronomy, Purdue University, West Lafayette, IN.

Project Goals: The goal of this project is to identify the genes involved in flowering time and to elucidate genetic mechanisms controlling the flowering responses of switchgrass. The specific objectives are to: 1) Conduct a genome-wide association study (GWAS) of flowering time in both lowland and upland panels; 2) Validate genes associated with flowering time in segregating populations; and 3) Examine transcriptome profiling with extreme genotypes using RNA sequencing.

The timing of phase change from juvenile (vegetative) to adult with reproductive competence is a key factor influencing biomass yield of switchgrass. A decline in biomass yield is typically observed in switchgrass immediately following completion of flowering. In temperate regions of the USA, if flowering time can be delayed about 4-5 weeks, biomass yield can potentially increase 30-50%. The use of late-flowering switchgrass genotypes has proven an effective mechanism to increase biomass production in the northern USA. Southern populations (lowland ecotypes) of swtichgrass can be 4–6 weeks later in flowering time than upland types, and selection of late flowering genotypes with southern origin, but adapted to northern conditions, can extend vegetative growth. To develop a rational strategy for creating improved switchgrass with late flowering, it would be beneficial to have a better understanding of the genes that control flowering time in switchgrass. Through GWAS of 1.3 million single nucleotide polymorphisms with 510 diverse individuals from both lowland and upland panels, we have identified 10 genes significantly associated with flowering time. One candidate gene is a homolog of Flowering Locus T (FT). We have developed 3 mapping populations derived from upland (early flowering) and lowland (late flowering) materials and established the populations in the field in DeKalb, IL and West Lafayette, IN. Flowering time will be evaluated in two locations over multiple years. We will genotype the mapping populations using exome capture and genotyping-by-sequencing, and further detect and verify genetic associations in the mapping populations of switchgrass. Knowledge generated from the project will aid breeding programs in developing late flowering varieties of switchgrass that fully utilize the growing season and achieve high biomass yield.

This project is supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0010631.

194. Accelerated Development of Optimal Feedstock for Bioenergy Using Genome-Wide Selection

Matricio Munoz,1 Gary F. Peter,2,3,4 and Matias Kirst2,3,4* (mkirst@ufl.edu)

1Department of Agronomy, University of Florida, Gainesville; 2Plant Molecular and Cellular Biology Program, University of Florida, Gainesville; 3School of Forest Resources and Conservation, University of Florida, Gainesville; 4Genetics Institute, University of Florida, Gainesville

Project Goals: Our overall goals are to hyper-accelerate pine breeding, to generate cultivars of loblolly and slash pine tailored to produce high energy yields, that are ready for deployment within the duration of this project. To achieve these goals we are using genome-wide selection prediction models to identify and generate crosses designed to support the short-term needs of the bioenergy industry. Genome-wide selection prediction models were previously developed for this population, which has been grown at multiple sites, and was phenotyped for growth, disease resistance and juvenile wood composition traits (lignin, cellulose, and terpene content). In parallel, we are also genotyping and phenotyping advanced breeding populations of loblolly and slash pine that encompass a broader genetic diversity of each species. For these populations, genome-wide selection prediction models will be generated for growth traits, wood chemistry and terpene flow. These genetically diverse populations will serve as the foundation for the next generation of advanced pine feedstocks for the bioenergy, biofuels and renewable chemicals industries.

Southern pines, such as Pinus taeda (loblolly pine) and Pinus elliottii (slash pine), are a proven sustainable source of renewable biomass for bioenergy and renewable chemicals. The S.E. forest industry produces ~18% of the global supply of industrial roundwood, 25% of the global pulp supply and 40% of the global pine chemical supply. This industry also generates 77% of all industrial biomass energy in the U.S. by burning wood waste and lignin at high thermal efficiencies with combined gas cycle turbines. The 93 million hectares of standing pines in the S.E. include 15.6 million hectares of genetically improved plantations; all grow with low water, fertilizer and herbicide inputs on land that is largely not suitable for cultivation of food crops. This large extant southern pine resource, with its well established supply chain and predictable year round supply is a promising source of bioenergy. To sustainably meet this potential it will be necessary to develop pine trees that produce higher energy and renewable chemical yields per acre, per year. Because of the wide natural genetic variation in southern pine growth rates and wood properties, traditional breeding can achieve this potential. However, breeding of pines is a long and complex process, due to the fact that each breeding cycle takes several decades and large field trails. Thus, current traditional breeding practices are too costly and slow to rapidly develop pines to reach the productivity required.

We recently demonstrated the potential that genome-wide selection, based on molecular markers, has to more than double genetic gains in tree breeding. These gains are achieved using prediction models for (i) early genotypic selection of elite individuals, leading to a dramatic reduction in the length of the breeding cycle; and for (ii) selecting crosses that generate optimal allelic combinations in the progeny, reducing the complexity and costs of breeding programs.

The first objective of this project is to use genome-wide selection to rapidly develop highly productive loblolly pine families and cultivars, to support the short term needs of the bioenergy industry, including disease-resistant, fast growing germplasm with high cellulose, high lignin or high terpene content. We previously phenotyped and genotyped the loblolly pine reference breeding population CCLONES, which has been grown on sites spanning the main breeding zones. This population has been phenotyped for biomass growth, wood composition and disease resistance traits. In this project, the phenotypic and

genotypic data was used to develop advanced models that include additive and dominance effects, which improves prediction accuracies significantly compared to traditional methods. These improved methods have now been used for mate-pair allocation, to identify the crosses predicted to generate the highest yielding families. In the remaining ~ 3 years of this project, the best crosses will be made and progeny screened to identify elite individuals for bioenergy and biofuel production.

The first aim we will apply advanced GWS prediction models for mate-pair allocation and early identification of elite individuals to an existing population, creating a resource for immediate development of highly productive loblolly pine families and cultivars for bioenergy. However, CCLONES has limited genetic diversity, includes only one of the two pine species (loblolly and slash pine) widely planted in the S.E. U.S., and has a relatively small size. To address these limitations and create a resource for long-term improvement of southern pines for bioenergy, in our second objective we are genotyping and phenotyping large, genetically diverse breeding populations of loblolly and slash pine. Those populations comprise the most advanced third generation genetic material from the Cooperative Forest Genetics Research Program (CFGRP) at UF. For slash pine, these large populations have already been established in field sites. For loblolly pine, field trials will be planted in the winter 2014-2015. DNA has already been extracted from these populations, and genotyping and phenotyping will be completed over the next 24 months. As part of the in-kind support provided by members of the CFGRP, measurements of growth for both loblolly and slash pine will proceed beyond the duration of this project, until the trials reach 12 years of age.

This project is funded by the USDA National Institute of Food Agriculture - Institute of Bioenergy, Climate and Environment, competitive grant no. 2013-67009-21200.

195. Demonstrating Agrobacterium-mediated transformation of Pennycress (Thlaspi arvense) by expressing EaDAcT to alter seed oil composition

Michaela McGinn1, Sunil Bansal2, Cynthia Cass1, Timothy Durrett2, and John Sedbrook1 Presented on behalf of John Sedbrook (CoLead PI) by Michael Marks*3 (PI)

1School of Biological Sciences, Illinois State University, Normal, IL 61790; 2Department of Biochemistry and Molecular Biophysics, Kansas State University, Manhattan, KS 66506; 3University of Minnesota

http://biology.illinoisstate.edu/jcsedbr/

Project Goals: Develop pennycress into a new oilseed crop that can be intercropped between corn and soybeans in the Upper Midwest.

Pennycress (Thlapsi arvense) is a Brassica species related to Arabidopsis and rapeseed that holds considerable agronomic and economic potential due to its unique combination of attributes including extreme cold tolerance, rapid growth, over-wintering growth habit, and natural ability to produce copious amounts of seeds high in oil and protein content.

Pennycress could be grown on millions of hectares of farmland throughout temperate regions of the world including the U.S. Midwest Corn Belt, e.g. being planted in the fall near the time of corn harvest and harvested the following spring in time for planting soybeans. While pennycress could potentially generate billions of liters of oil annually without displacing food crops or requiring land use changes, work must be done to develop pennycress varieties having improved seed germination and in-field stand establishment as well as improved seed oil and meal quality. To facilitate research on pennycress, we have developed a pennycress genetic transformation protocol using an Agrobacterium- mediated floral dip method. We found that, as with camelina, pennycress transformation was enhanced by applying a vacuum to floral spikelets while submerged in a sucrose- and Silwet L77-containing Agrobacterium solution. Hygromycin and DsRED worked well for selection of/screening for transformants, whereas pennycress seedlings exhibited naturally high resistance to kanamycin. We will discuss targets and approaches to genetically improving pennycress and will present data showing that we have successfully generated pennycress plants stably transformed with the Euonymus alatus diacylglycerol acetyltransferase (EaDAcT) gene; these plants produce seeds accumulating 3-acetyl-1,2-diacyl-sn-glycerols (acTAGs), a low---viscosity oil.

This project was supported by USDA National Institute of Food and Agriculture: USDA / 2014-67009-22305

196. Using mutation-based breeding approaches to domesticate pennycress into a new oilseed crop

Evan Johnson, Erin Daniels, Kevin Dorn, and M. David Marks*

University of Minnesota, St. Paul

http://pennycress.umn.edu

Project Goals: Develop pennycress into a new oilseed crop that can be intercropped between corn and soybeans in the Upper Midwest.

Thlaspi arvense (field pennycress; pennycress herein) is being targeted as a new winter-grown oilseed cover crop and biodiesel feedstock [1]. Pennycress is a member of the Brassicaceae family and is related to model species Arabidopsis thaliana and current oilseed crop species Brassicas napus and rapa [2]. Unlike these latter two species, pennycress can be grown in the interval between the corn/soybean rotation in the Midwestern United States. It can be seeded into standing corn, allowed to overwinter and then harvested in the spring, allowing for a summer crop of short-season soybeans [3]. Thus, it has potential to provide a new source of biofuel without requiring new land or greatly changing current farming practices. Additionally, it provides a winter cover on land that is traditionally left barren, which will reduce nutrient leaching and soil erosion.

Pennycress was chosen as a potential new crop species because it already harbors a number of desirable traits. These include a high oilseed yield (1100 to 2250 kg/ha), a high seed oil content (30-40% by seed weight), extreme cold tolerance, fast seed maturation in the spring, good plant stature (1>meter), and seed pods that exhibit less shattering compared to those of related species. However, many improvements need to be made to make pennycress a successful new crop species. In a recent review, Sedbrook et al. described strategies to rapidly improve and domesticate pennycress, particularly mutation breeding [1]. We will use both forward mutant screens, and reverse genetic approaches that will rely on our recently developed genomic resources for pennycress, which include both transcriptome and genome assemblies [4, 5]. We know from work with Arabidopsis that mutations can confer what can be considered agronomically desirable phenotypes. For example, various mutations in Arabidopsis have been shown to increase seed size, decrease seed dormancy, improve oil quality, speed up flowering and reduce seed shattering.

To begin our mutation-based breeding strategy, seeds were treated with three types of mutagens: EMS, fast neutrons, and gamma rays. These were planted in the fall of 2013 and M2 seeds were collected in the spring of 2014. That fall 1,000 rows were sowed, each containing seeds derived from pools of 10 M1 plants. To date, we are in the process of phenotypically screening M2 plants derived from 10,000 M1 plants. Already, over 100 mutants having phenotypes of interest have been identified. These include mutants with agronomically important traits such as early maturation, larger seeds, and light-colored seed coats, which we know from Arabidopsis are likely to produce seeds with higher oil content and reduced dormancy [6]. In the spring of 2015, we will screen the 1000 rows of plants in the field for additional traits of interest. We anticipate finding many more mutants with desirable traits. In addition, we are banking seeds and tissues that will used for TILLING, with the ultimate goal of directly identifying mutations in genes that are known to produce desirable traits when mutated in Arabidopsis [7].

Literature Cited

2. Sedbrook, J.C., W.B. Phippen, and M.D. Marks, New approaches to facilitate rapid domestication of a wild plant to an oilseed crop: example pennycress (Thlaspi arvense L.). Plant Sci, 2014. 227: p.

122-32.
3. Franzke, A., et al., Cabbage family affairs: the evolutionary history of Brassicaceae. Trends Plant Sci, 2011. 16(2): p. 108-16.
4. Phippen, W.B. and M.E. Phippen, Soybean Seed Yield and Quality as a Response to Field Pennycress Residue. Crop Science, 2012. 52(6): p. 2767-2773.
5. Dorn, K.M., et al., De novo assembly of the pennycress (Thlaspi arvense) transcriptome provides new tools for the development of a winter cover crop and biodiesel feedstock. The Plant Journal, 2013.
6. Dorn, K.M., et al., A draft genome of field pennycress (Thlaspi arvense) provides tools for the domestication of a new winter biofuel crop. DNA Research, 2015. in press.
7. Debeaujon, I., K.M. Leon-Kloosterziel, and M. Koornneef, Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. Plant Physiol, 2000. 122(2): p. 403-14.
8. Tsai, H., et al., Discovery of rare mutations in populations: TILLING by sequencing. Plant physiology, 2011. 156(3): p. 1257-68.

This project was supported by USDA National Institute of Food and Agriculture: USDA / 2014-67009-22305

197. Pyramiding genes and alleles for improving energy cane biomass yield

Ray Ming1, Chifumi Nagai2, Qingyi Yu3* (qyu@ag.tamu.edu)

1University of Illinois at Urbana-Champaign, Urbana, Illinois; 2Hawaii Agriculture Research Center, Kunia, Hawaii; and 3Texas A&M AgriLife Research, Dallas, Texas

http://www.life.illinois.edu/ming/LabWebPage/Home.html

Project Goals: Our long term goal is to establish a new paradigm to accelerate energy cane breeding programs and maximize the biomass yield for biofuel production. Our specific objectives are: (1) Phenotyping extreme segregants of the F2 population for exploring the molecular basis of high biomass yield from transgressive segregation; (2) Mapping genes affecting biomass yield by transcriptome sequencing of the extreme segregants; (3) Identifying differentially expressed genes and alleles through analyses of transcriptomes of extreme segregants from the F2 population; (4) Developing gene- and allele-specific markers for implementation of marker-assisted selection in energy cane breeding programs. As a C4 plant, sugarcane/energy cane has been recognized as one of the world's most efficient crops in converting solar energy into chemical energy. Traditional energy cane and sugarcane breeding via interspecific hybridization and backcrossing to S. officinarum improved stress tolerance and recovered high sugar and biomass yield. However, this approach reduced the genetic diversity of sugarcane and energy cane breeding materials and limited the potential maximizing biomass yield. We are developing a new paradigm to accelerate energy cane breeding programs and maximize the biomass yield for biofuel production.

An F2 population with 2616 individuals was created from an interspecific cross between S. officinarum LA Purple (2n = 80, x = 10) x S. robustum MOL5829 (2n = 80, x = 10). This population showed transgressive segregation with high yielding clones substantially exceeding the biomass yield of both parents. Field trial of the extreme segregants was carried out for three years in Hawaii. Biomass yield of the top 10 F2 clones ranged from 71.5 to 122.1 MT/ha in 12 months, and the best performing clone showed 338% yield increase compared to its high yield parent LA Purple. The bottom 10 F2 clones had estimated biomass yield ranging from 3.4 to 8.2 MT/ha in 12 months, and the worst performing clone showed yield decrease to only about 10% of LA Purple yield. It should be noted that sugarcane biomass yield in Hawaii is 30 MT/ha per year, and the potential biomass yield of sugarcane/energy cane is 140 MT/ha per year.

Transcriptome sequencing of the 3rd, 9th, and 15th internodes and the 1st dew lap leaf from the 14 top and 8 low biomass segregants of the F2 population was carried out in order to map genes or alleles affecting biomass yield in energy cane. Differential gene expression analysis was conducted between the high and low biomass groups. A total of 2,475 genes were up-regulated and 799 genes were down-regulated in high biomass group. GO terms analysis indicated that the differentially expressed genes between the high and low biomass groups were enriched in the cell wall modification, catabolic process, and carbohydrate catabolic process. The genes that encode cell wall and pectin modifying enzymes were up-regulated in the high biomass group.

Twenty percent of the genes in the breakdown pathway of homogalacturonan, one of the major pectin polysaccharides, were up-regulated in the high biomass group. In contrast, no significant difference of the genes involved in pectin polysaccharides biosynthesis was detected between the high and low biomass groups. For the genes in the pathway producing the lignin precursors, genes involved in caffeoyl-CoA synthesis were highly up-regulated, and genes in converting caffeoyl-CoA into downstream products were down-regulated in the high biomass group. Among the 13 genes involved in UDP-D-glucose

synthesis pathway, 4 of them were up-regulated in the high biomass group.

In summary, genes regulating the biosynthesis and breakdown of the cell wall components play essential roles in achieving high biomass yield in energy cane. Increased cell expansion and cell division in high biomass yield energy cane was caused by fast turnover of cell wall components. High biomass yield in energy cane was achieved through promoting primary cell wall production without over-lignification of secondary cell wall.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0010686.

198. Genomics of Energy Sorghum's Water Use Efficiency/Drought Resilience

John Mullet1*(jmullet@tamu.edu), William Rooney1, and Marc Libault2 1Texas A&M University, 2University of Oklahoma

Project Goals/Description: The overall goal of the proposed research, funded through Plant Feedstock Genomics for Bioenergy, is to increase the water use efficiency, drought resilience, and yield of high biomass energy sorghum and other C4 bioenergy grasses. Energy sorghum hybrids have been developed that have high biomass yield, excellent greenhouse gas displacement values, and good nitrogen use efficiency (1, 2, 3). Further increases in biomass yield will require the development of energy sorghum with improved water use efficiency, root systems with excellent water extraction properties and drought resilience. This goal will be accomplished by carrying out the following research objectives: (i) identify traits and molecular responses that improve the water use efficiency and drought resilience of energy sorghum using root-trait lysimeters and field analysis, (ii) characterize the genetic basis of variation in water use efficiency and drought resilience using the energy sorghum association panel and RIL populations, and (iii) test the utility of traits that modulate water use efficiency and drought resilience in energy sorghum hybrids through marker-assisted breeding.

References

1. Mullet, J.E., Morishige, D.T., McCormick, R., Truong, S., Hilley, J, McKinley, B., Anderson, R., Olson, S., and Rooney, W. (2014) Energy Sorghum – A Genetic Model for the Design of C4 Grass Bioenergy Crops. Journal of Experimental Botany 65: 3479-3489 (doi:10.1093/jsb/eru229)
2. Olson, Sara, Kimberley Ritter, Jim Medley, Ted Wilson, William Rooney, and John E. Mullet (2013) Energy Sorghum Hybrids: Functional Dynamics of High Nitrogen Use Efficiency. Biomass and Bioenergy 56: 307-316.
3. Olson, S.N., Ritter, K., Rooney, W., Kemanian, A., McCarl, B.A., Zhang, Y., Hall, S., Packer, D., and Mullet, J. (2012) High Biomass Yield Energy Sorghum: Developing a Genetic Model for C4 Grass Bioenergy Plants. Biofuels, Bioproducts and Biorefining 6: 640-655 (DOI: 10:1002/bbb)

Funding for this research project is provided through the Plant Feedstock Genomics to Bioenergy Program.

199. Genomics of Bioenergy Grass Architecture

Andrew H. Paterson1*, Wenqian Kong1, Dong Zhang1, Changsoo Kim1, Hui Guo1, Valorie Goff1

1Plant Genome Mapping Laboratory, University of Georgia, Athens GA

Project Goals: Understanding the genetics of plant architecture has taken on new importance with invigorated efforts to develop plant genotypes optimized for production of biomass for use in fuels or chemical feedstocks. We will increase knowledge of genetic determinants of components of plant architecture that are important to the 'design' of crop genotypes optimized for production of biomass from a range of environments, also characterizing allelic and haplotype variation in salient sorghum genes toward their deterministic utilization, and enabling new integrative queries of these and other results to accelerate discovery and utilization of genetic determinants in biomass crop improvement. This project fits closely with DOE missions of advancing energy security, scientific discovery, and environmental responsibility, and leverages and enhances the value of many existing resources.

Abstract Text: Variations in plant architecture are central to the yield potential of biomass crops under high-input conditions and are a key element of crop adaptation to marginal environments under which most cellulosic biomass production may occur. Optimal biomass productivity in temperate latitudes and/or under perennial production systems may require substantial changes to architecture of plants of tropical origin that have previously been adapted to annual cultivation. Sorghum is of intrinsic importance as one of the few crops suited to all proposed approaches for renewable fuel production, i.e. from starch, sugar, and/or cellulose; and as a model for functional genomics of bioenergy plants, in view of its C4 metabolism (unlike rice), its small genome (unlike many), and its low level of gene duplication resulting from ~70 million years of abstinence from genome-wide duplication (unlike switchgrass, Miscanthus, maize, sugarcane, and other grasses). Its remarkable diversity of morphologies make sorghum a particularly attractive model for dissecting the genetic basis of grass morphological diversity. Utilizing these advantages, we are employing forward genetics in three populations that broadly sample the eusorghum clade to obtain baseline QTL data for components of plant architecture and related traits, also assessing inter-relationships with one another and with perenniality; association genetics of positional and/or functional candidate genes to narrow the locations of QTLs, reveal haplotype diversity in trait-controlling regions, investigate roles in quantitative variation of major genes identified in other taxa, provide insight into the evolution and distribution of phenotypic variation in components of plant architecture and related traits useful toward numerous applications in sorghum genetics and breeding, and perhaps even identify functional variants in some instances; and facilitate integrative use of positional, diversity, and mutant information in discovery and utilization of genetic variation by using Gramene trait ontologies to 'interleave' (in silico) three well-characterized but to date isolated genetic resources. Collectively, these expected results will provide a foundation for deterministic improvement of sorghum itself, accelerate progress in dissecting the genetic control of plant architecture in the complex polyploid genomes of closely-related Saccharinae biofuel crops, and nurture new dimensions in plant functional genomics with novel integrative queries that may accelerate progress toward long-held goals and provide a powerful attraction to new investigators.

Publications:

1. Wenqian Kong, Huizhe Jin, Changsoo Kim, Valorie H. Goff, Tae-Ho Lee and Andrew H. Paterson. 2014. Genetic analysis of vegetative branching in sorghum. Theoretical and Applied Genetics 127:2387-2403. DOI 10.1007/s00122-014-2384-x

2. Dong Zhang, Wenqian Kong, Jon Robertson, Valorie H. Goff, Ethan Epps, Alexandra Kerr, Gabriel Mills, Jay Cromwell, Yelena Lugin, Christine Phillips, Andrew H. Paterson 201#. Comparative

genetics of inflorescence and plant height components in divergent cereal lineages represented by sorghum (Panicoidae) and rice (Oryzoidae).

This research was supported by the USDA National Institute of Food and Agriculture (NIFA), AFRI Competitive Grant no. 2012-67009-19718.

200. Global Analysis of Alternative Splicing and Epigenetic Regulation of Gene Expression in Response to Drought Stress in Sorghum

Salah E. Abdel-Ghany1, Michael Hamilton2, Fahad Ullah2, Asa Ben-Hur2, and A.S.N. Reddy1* (reddy@clostate.edu)

1Dept. of Biology, Colorado State University, Fort Collins, CO, 80523; 2Department of Computer Sciences, CSU, Fort Collins, CO, 805423.

Abiotic stresses including drought are major limiting factors of crop yields and cause significant crop losses. Acquisition of stress tolerance to abiotic stresses requires coordinated regulation of a multitude of biochemical and physiological changes, and most of these changes depend on alterations in gene expression. The major goal of this proposal is to investigate drought-induced global changes in alternative splicing events and epigenetic modifications that are important for drought tolerance in Sorghum. The specific objectives of this project are: 1) Perform global analysis of differential regulation of expression and alternative splicing of primary transcripts from both protein coding and non-coding genes in drought sensitive and tolerant cultivars. 2) Perform global analysis of accessible chromatin landscape changes in drought sensitive and tolerant cultivars in the presence and absence of drought stress. 3) Analyze genome-wide changes in histone modifications in response to drought stress in drought sensitive and tolerant cultivars in the presence and absence of drought stress. The focus of this objective will be on global analysis of stress-induced changes in histone modifications. Our ultimate goal is to correlate chromatin modifications to changes in gene expression and alternative splicing, and to identify genes and regulated splicing events that play a role in drought tolerance.

To identify two drought tolerant and two susceptible cultivars of sorghum for our proposed experiments we screened seven sorghum varieties, which include the sequenced BTx623 cultivar. Based on the growth phenotype we selected two drought tolerant and two drought susceptible cultivars. To investigate differential regulation of gene expression and alternative splicing in response to drought we performed RNA-Seq for all four lines in control and drought-treated seedlings. Two biological replicates were used for each sample. Prior to RNA-Seq we confirmed the efficacy of drought treatment by testing the expression of known drought-responsive marker genes using either semi- quantitative PCR or real-time PCR (qPCR). Analysis of RNA-Seq data has revealed that between 1300 and 2000 genes are differentially expressed in response to treatment and a large majority were up-regulated. A large fraction of the up-regulated genes were common among all four cultivars. However, each cultivar had a unique set of up- regulated genes in response to drought treatment. Selected candidate genes that are differentially expressed or alternatively spliced were confirmed using semi-quantitative and quantitative RT-PCR. To study differential alternative splicing we developed a tool called iDiffIR that detects statistically significant differences in intron retention across conditions. Preliminary analysis using this tool has detected hundreds of differential intron retention events in response to drought treatment, indicating that splicing is involved in the response to drought.

As the Sorghum transcriptome is not very well annotated, with very few known splice variants per gene, we sequenced the transcriptome of one cultivar (BTx623) using the Pacific Biosciences Iso-Seq platform that provides long reads and permits accurate prediction of splice variants. This is a new technology with very limited tools for the analysis of such data, especially due to the high error rate of the resulting reads. We developed a new pipeline called TAPIS (Transcriptome Analysis Pipeline for Isoform Sequencing) that predicts both isoforms and alternative polyadenylation sites and uses either the reference genome or Illumina reads for error correction. Out of 33,032 genes in version 2 of the Sorghum annotations, we found 7,108 genes with one or more reads covering both first and last exons. Analysis of this data revealed >2000 novel alternative splicing events, more than doubling the number of known intron

retention events. We also uncovered over 600 loci that were misannotated and evidence for around 500 novel genes, some of which are similar to known protein-coding genes in other plants. As the 3' ends of transcripts are well-covered by Iso-Seq reads, we were also able to uncover the landscape of polyadenylation sites. Transcripts from over 2000 genes were found to have two or more polyadenylation sites. Of these, around 4% were in the coding region or the 5' UTR.

Intron retention is the predominant form of alternative splicing in plants. However, very little is known about the regulatory sequences in pre-mRNAs that regulate this event. To identify putative elements that contribute to intron excision or retention, we have developed a computational pipeline that uses all known IR and intron excision events in Sorghum and other land plants to detect over-represented hexamers that are conserved across species. We have found over-represented hexamers that fall either in the flanking exons, or in the intron itself. A large fraction of these hexamers are also known to regulate splicing in animals. In view of their conservation and similarity to known splicing regulatory elements in animals, the elements we detected are likely to be functional.

In our proposal, we suggested to use DNAaseI-Seq to analyze open chromatin. However, a new method called DNase-FLASH (DNase I–released fragment-length analysis of hypersensitivity) that is superior to DANseI-Seq was reported recently (Vierstra et al., Naure Methods 11, 66–72, 2014). We have been standardizing this method for Sorghum. We anticipate sending DNA samples form DNase-FLASH and chromatin immunoprecipitation in the near future.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0010733

201. Deciphering Natural Allelic Variation in Switchgrass for Biomass Yield and Quality Using a Nested Association Mapping Population

Malay C. Saha1,2* (mcsaha@noble.org), Desalegn D. Serba1,2, Hem S. Bhandari3, Shawn Kaeppler4, E. Charles Brummer5

1Forage Improvement Division, The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401; 2BioEnergy Science Center (BESC), Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA; 3Department of Plant Sciences, University of Tennessee, 2431 Joe Johnson Drive, Knoxville, TN 37996; 4Department of Agronomy, University of Wisconsin, 1575 Linden Drive, Madison, WI 53706; 5Plant Breeding Center, Department of Plant Sciences, University of California, Davis, One Shields Ave., Davis, CA 95616

Project Goals: Long-term goal of this project is to understand the genetic basis of the key biofeedstock traits of biomass yield and composition in order to accelerate development of superior cultivars of switchgrass. The specific objectives are to: i) develop a nested association mapping (NAM) population of 2,000 plants and construct a genetic map for this population; ii) evaluate the population at multiple environments to identify QTLs and molecular markers associated with biomass yield, feedstock quality and other agronomically important traits; iii) validate marker-QTL associations in breeding populations as part of a marker-assisted breeding program. Validated markers co- segregated with key traits for biomass yield and composition will be used to initiate a marker-assisted and/or genomic selection program to develop improved switchgrass cultivar(s).

Switchgrass (Panicum virgatum L.) biomass yield and feedstock quality improvement are priority research areas for bioenergy feedstock development. Understanding the genetic basis of quantitative traits is essential to predictive crop improvement. Nested association mapping (NAM) is a technique especially designed for identifying and dissecting the genetic architecture of complex traits. NAM enables high power and high resolution in QTL detection through joint linkage-association analysis by capturing the best features of bi-parental and association mapping approaches. A switchgrass NAM population was developed by crossing 15 diverse genotypes with specific characteristics, selected from natural variants, to a recurrent parent, AP13, a genotype used for the genome sequencing. Confirmed F1s (n=10) from each cross were then chain-crossed and a final population of 2,000 progenies were developed.

A total of 2,350 genotypes (2000 progeny genotypes from 15 families, 30 ramets of AP13, three copies of grandparents, two copies of F1 parents and Alamo checks) were planted in the field at Knoxville, TN and Ardmore, OK following the Alpha Lattice design. Phenotypic data were collected on biomass yield and related traits. In 2014, Average plant height of different families varied from 64-300 cm. At Knoxville, average yield of grandparents was 0.75 kg plant-1, F1 parents was 0.83 kg plant-1 and AP13 was 0.44 kg plant-1. Average yield of progeny families varied from 0.17 to 0.82 kg plant-1. In 2013, the family average yield ranged from 0.06-0.12 kg plant-1. Wide variability was observed within each family. Several families produced more biomass than their parental genotypes. We just completed harvesting of the population at Ardmore, OK.

Grandparents of the NAM population have been sequenced at the JGI following the NGS technology. After quality filter, on average we received 28-66 Gb sequence data from each parent. Preliminary analysis of the sequence data revealed that 78-98 % of the sequence reads can be mapped to the reference switchgrass genome sequence. A total of 249 million reads from the parent CDV09_05 were mapped to the switchgrass genome. Number of reads assembled to each switchgrass chromosome varied from 10.3 to 20.6 million. Genomic DNA was extracted from the whole population and genotyping with exome capture is in progress at the University of Wisconsin.

Acknowledgement: This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-S0018448.

202. Transcriptome Analysis of Diploid and Triploid Species Hybrids of Shrub Willow

Agnes Chan1, Yongwook Choi1, Craig H. Carlson2, Michelle J. Serapiglia2, Christopher D. Town1, and Lawrence B. Smart2,* (lbs33@cornell.edu)

1J. Craig Venter Institute, Rockville, MD; 2Horticulture Section, School of Integrative Plant Science, Cornell University, NYSAES, Geneva, NY

http://willow.cals.cornell.edu

Project Goals: Our goal is to understand how the gene expression patterns in willow hybrids are related to their yield potential and other traits important for biofuels production. Specifically, we will test if there is a bias in the expression of key genes from one parent versus the other in species hybrids, and whether there is a gene dosage effect skewing gene expression patterns in triploid progeny compared with their diploid and tetraploid parents.

Interspecific hybridization has been a primary approach in breeding shrub willow (Salix spp.) as an energy crop, as hybrids often display heterosis for yield. Expanding available genomic resources for shrub willow will not only facilitate breeding efforts but also contribute to our understanding of hybrid vigor in outcrossing woody perennials. Utilizing RNA-Seq, transcriptome variation in hybrids of intra- and interspecific crosses of shrub willow is being investigated to determine genome-wide dosage-dependent allelic bias and differential patterns of cis- and trans-regulation within and between hybrid categories. Specifically, shoot-tip and stem internode libraries of 10 random progeny from intra- and interspecific families and their parents have been sequenced to identify alleles that correlate with heterosis for biomass yield and variation in lignocellulosic composition. Allele-specific expression will be confirmed using independent assays and will subsequently be used as a resource for developing predictive models for heterosis in hybrid shrub willow.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0008375.

203. Structural Polymorphisms as Causes of Heterosis in Populus

Upendra Kumar Devisetty1, Kori Ault1, Alejandro Riveros Walker2, Sandra Simon2, Stephen Difazio2, Austin Himes3, Brian Stanton3, Todd Rosenstiel4 and Steven H. Strauss1*
(Steve.Strauss@oregonstate.edu)

1 Oregon State University, Corvallis OR; 2 West Virginia University, Morgantown WV; 3Greenwood Resources, Portland, OR and 4Portland State University, Portland, OR

http://people.forestry.oregonstate.edu/steve-strauss/

Project Goals: The main objectives of the project are to characterize the extent of structural polymorphisms (SP) that exist between and within species in cottonwood (poplar) trees that are used to produce wood and bioenergy, and examine their relationships to phenotypes such as growth, stress and breeding efficiency. The ultimate goal is to identify the extent to which SPs-primarily insertions and deletions of genes-could improve hybrid breeding in Populus compared to alternative approaches such as SNPs (single nucleotide polymorphisms) using genomic selection methods. We are using phenotypic and genomic information from existing research trials as well as generating substantial new phenotypic and genomic data.

Project background: Heterosis refers to the superior performance and stress tolerance of highly heterozygous groups or individuals. It is widely known for its applicability in maize. Many perennial biomass programs also employ heterosis either in the form of polyploids (e.g., Miscanthus) or in the form of interspecific hybrids (e.g., Populus, Salix, Acacia and Eucalyptus). The proposed research will test the predictions of a fundamental theory of heterosis (dominance) in relation to environmental stress, and its application to marker- aided selection to improve the efficiency of Populus breeding for woody biomass production. We will first analyze SPs in three distinct populations that serve specific roles in the project: a) in a wild P. trichocarpa collection of 1,100 wild genotypes where SP effects can be finely dissected due to relatively low LD; b) in an outbred hybrid backcross collection of 337 genotypes of (P. trichocarpa x P. deltoides) × P. deltoides, to jointly examine the effects of between- as well as within-species variation in SPs; and c) in F1 hybrid families of P. deltoides × P. nigra (DN), P. deltoides × P. maximowiczii (DM) and P. deltoides × P. koreana (DK) comprised of 371 progeny in total, to evaluate the efficiency of detection and prediction impact of SPs within diverse hybrids. We will generate SP genotype data using bait-capture/resequencing, and test the degree to which SPs are associated with variation in productivity traits (stem height, diameter, wood density, chlorophyll content, chlorophyll fluorescence; and carbon assimilation and respiration under imposed heat stress). Finally we will test whether the selected SPs-putatively focusing on selection against homozygosity for gene-containing deletions can improve selection efficiency and economic gain within families over phenotypic and genomic selection approaches.

Progress to date: Currently, a high quality reference genome is available for P. trichocarpa, but not for other poplar species. We will therefore carry out an assembly to produce reference genomes for the other species in our project. The genome of P. deltoides (WY94) was first assembled using Illumina paired end reads of 100bp using ABYSS denovo assembler. The generated assembly has an N50 of 10,490 using k-mer of size 63, below our requirements for this project. Further work on genome assembly for this clone will await its genome assembly by JGI-as part of a separately funded Community Sequencing Project. The other three reference genomes needed for the project are being analyzed through access to data from collaborators (P. nigra), and de novo sequencing (P. maximowiczii and P. koreana). We have generated Illumina MiSeq (2x300bp) coverage totaling 12X for each of these species and are currently generating an additional 40X of Illumina HiSeq sequences (2x100bp). In addition, we will generate mate pair

libraries and/or long read sequencing (PACBio) to span large indels and create more contiguous scaffolds using hybrid assembly approaches.

For genotyping of 1,100 trees of P. trichocarpa, we have used the 15X resequencing data available from DOE to perform SP discovery. For genotyping of the hybrid and backcross populations, we will first use a sequence-capture assay devised to efficiently target the genome containing the most polymorphic gene containing SPs within and among species. We will sequence a discovery population of 24 individuals per species to a depth of 15X for discovery of segregating SPs, and will then devise a sequence capture assay for these loci. Thus far we have made progress in optimizing high-throughput library preparation protocol for this purpose.

We made substantial progress in phenotypic assessments in summer 2014. Working with GreenWood Resources, we established a new population of F1 hybrid families in Jefferson, Oregon in spring 2014 that had 98% survival. We made phenotypic measurements here and in two other populations. Preliminary phenotypic analyses indicated that average photosynthesis measured in both light-adapted and dark-adapted leaves was significantly different for genotypes at increasing temperatures for both association and F1 populations. The dark-adapted chlorophyll fluorescence varied significantly among F1 families but not among genotypes within families. The light-adapted chlorophyll fluorescence was significant for F1 families and is also significant for genotypes in two out of 6 families. In the P. trichocarpa association population we also assayed light- and dark-adapted fluorescence and identified loci underlying these traits using genome- wide association analyses with our existing SNP dataset.

In the coming year we plan to do more phenotyping, including chlorophyll fluorescence, gas exchange as a function of temperature, leaf area, specific leaf mass, stomatal density, and height/DBH. Currently we are working on indel discovery in the hybrid and backcross populations; optimizing the bait-capture probe design for their determination; and then combining high-throughput DNA library production with bait-capture to generate SP genotype data. Finally, we will be developing a genomic selection model to assess if the produced genomic data and selection of SPs could be used to incrementally improve molecular accelerate breeding.

This project is supported by the USDA National Institute of Food Agriculture - Institute of Bioenergy, Climate and Environment, competitive grant no. 2013-67009-21008

204. Dual Effect of Tubulin Manipulation on Populus Wood Formation and Stomatal Behavior

Hao Hu1,2, Yingying Zhu1,2, Prashant Swamy1, Sivakumar Pattathil3, Hui Xiao4, Liang-Jiao Xue1,2, Jeng-Der Chung5, Gary F. Peter6, Michael G. Hahn3, Shawn D. Mansfield7, Scott A. Harding1,2, Chung-Jui Tsai1,2,*

1Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA; 2Department of Genetics, University of Georgia, Athens, GA; 3Complex Carbohydrate Research Center, University of Georgia, Athens, GA; 4Laboratory for Macromolecular Analysis and Proteomics, Albert Einstein College of Medicine, Bronx, NY; 5Division of Silviculture, Taiwan Forestry Research Institute, Taipei, Taiwan; 6School of Forest Resources and Conservation, University of Florida, Gainesville, FL; 7Department of Wood Science, University of British Columbia, Vancouver, BC.

Project Goals

Cortical microtubules play important roles in the regulation of plant morphogenesis, cell wall biogenesis and stomatal behavior, presumably by governing cellulose microfibril orientation. However, genetic manipulation of tubulins often leads to abnormal plant development, making it difficult to investigate the underlying structural bases. This project developed multiple suites of transgenic Populus with altered alpha- (TUA) and beta-tubulin (TUB) expression and/or post- translational modifications (PTMs), either singly or in combination, in a plant-wide, guard cell- specific or xylem-specific manner. We showed that it is feasible to obtain morphologically normal transgenic Populus with altered tubulin expression and/or PTMs. However, biased expression between TUA and TUB transgenes, as well as between leaves and xylem was evident, regardless of the promoter used. While transgene over-expression was consistently achieved in leaves, transgene expression was low in xylem, suggesting that high levels of tubulin transgene expression were not tolerated in vascular tissues during regeneration of transformants.

Counterintuitive to the proposed role of microtubules in directing microfibril deposition, cellulose content and microfibril angle were not changed in transgenic wood, nor were hemicelluloses and lignin. However, glycome profiling revealed increased cell wall glycan extractability, especially for lignin-bound pectin and xylan polysaccharides. This was corroborated by increased expression of cell wall-modifying enzymes in xylem based on RNA- Seq analysis, and by altered syringyl-to-guaiacyl monolignol ratio in the transgenics. A subset of transgenic plants exhibited prolonged epidermal cell expansion and reduced structural pectin and xyloglucan extractability in mature leaves. The results suggest that growth-compatible, mild perturbations of tubulins affected primarily noncellulosic polysaccharides, especially pectins, in transgenic Populus. Guard cell responses to drought and light were slower in the transgenic than wild-type plants. The findings were less consistent with an effect of tubulin perturbations on microtubule stability, but more consistent with an effect of altered pectin networks on cell wall flexibility critical for reversible stomatal movement. Our work so far suggests that pectin networks are more sensitive than cellulose to subtle tubulin perturbation, and that the tubulin effects on pectin incorporation may be conditioned by the degree of lignification in xylem and leaves during cell wall biogenesis.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0008470.

205. Plant Microbe Interfaces: Proteome Characterization of Plant-Bacterial and Fungal- Bacterial Interactions

Gregory B. Hurst1* (hurstgb@ornl.gov), Karuna Chourey1, David J. Weston2, Dale A. Pelletier2, Timothy J. Tschaplinski2, Sara Jawdy2, Tse-Yuan Lu2, Cora Guennoc3, Aurélie Deveau3, Jessy Labbé2, Chongle Pan4, Zhou Li4, Jessie Uehling5, Rytas Vilgalys5, Gerald A. Tuskan2, Mitchel J. Doktycz2

1Chemical Sciences Division, Oak Ridge National Laboratory, Oak Ridge TN; 2Biosciences Division, Oak Ridge National Laboratory, Oak Ridge TN; 3INRA, Tree-Microbe Interactions, UMR1136, Champenoux, France; 4Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge TN; 5Biology Department, Duke University, Durham, NC.

http://PMI.ornl.gov

Project Goals: The goal of the PMI SFA is to understand the genome-dependent molecular and cellular events involved in establishing and maintaining beneficial interactions between plants and microbes. Populus and its associated microbial community serves as the experimental system for understanding how these molecular events manifest themselves within the spatially, structurally, and temporally complex scales of natural systems. To achieve this goal, we focus on 1) characterizing host and environmental drivers fro diversity and function in the Populus microbiome, 2) utilizing microbial model system studies to elucidate Populus-microbial interactions at the molecular level and dissecting the signals and pathways responsible for initiating and maintaining microbial relationships and 3) develop metabolic and genomic modeling of these interactions to aid in interpreting the molecular mechanisms shaping the Populus-microbial interface.

We are studying protein expression patterns, and changes in those patterns, in experiments involving bacteria and fungi isolated from the root microbiomes of Populus trees.

Populus proteome response to a constructed soil microbial community. We have acquired leaf proteome data from Populus specimens inoculated with a 3-member bacterial community (see poster by Weston et al.), as well as from axenic control Populus specimens. The goal of this study is to determine whether observed phenotype differences are accompanied by changes in the Populus proteome. Proteome data were obtained for leaves 1, 2, and 3 weeks post- inoculation. Approximately 2000 protein groups were identified for each individual measurement, summing to >5600 protein groups across the entire experiment, of which nearly 1700 exceed a minimum abundance criterion for quantitative comparison. The 20 most abundant proteins observed for each treatment account for 16 to 18% of the total protein abundance, and are dominated by ribosomal proteins, photosystem I and II proteins, peptidyl-prolyl cis-trans isomerases, and RuBisCo. Not surprisingly, the leaf proteomes showed only background-level signals from proteins of the inoculated bacterial strains. Despite clear growth phenotype differences between inoculated and uninoculated plants at the 3 week time point, as well as marked differences in metabolite profiles from a similar inoculation experiment, we observed only a few proteins that differed in abundance between the inoculated and control plants. However, the numbers of differentially abundant proteins increased with time. The fraction of observed proteins in several KOG functional categories (RNA processing and modification; energy production and conversion; translation, ribosomal structure and biogenesis; amino acid transport and metabolism; carbohydrate transport and metabolism; posttranslational modification, protein turnover, chaperone; and cytoskeleton) is higher than the corresponding fraction in the full predicted Populus proteome, while the observed fraction in the signal transduction mechanisms KOG is less than the full predicted proteome. The results indicate that observed growth phenotype and metabolic profiles do not strongly correlate with changes in protein abundance, suggesting future studies such as the roles of post-translational modification in regulation of enzyme

activity and deeper proteome measurements to interrogate low- abundance proteins such as transcription factors.

Effect of helper bacteria on the proteome of Laccaria bicolor. A Pseudomonas strain (GM41) isolated within the PMI project from the Populus microbiome has previously been shown to enhance the growth and root colonization by the mycorrhizal fungus L. bicolor. In initial proteome measurements of co-cultures of the fungus with this Mycorrhizal Helper Bacterial isolate (MHB; Labbé et al., 2014), we have identified both bacterial and fungal proteins that are present only in the co-culture, but not in cultures of either organism grown alone. We are examining these proteins for evidence that will identify mechanisms underlying the MHB phenomenon.

Proteome of Mortierella elongata and a bacterial endosymbiont. Mortierella elongata is a commonly isolated soil fungus that plays an important role in forest soil ecosystem processes by affecting plant growth, carbon sequestration, and nitrogen cycling. This fast-growing fungus hosts a bacterial endosymbiont, Candidatus Glomeribacter sp. We have applied mass spectrometry-based quantitative proteomics to investigate changes in protein abundance, as well as post-translational modifications. This analysis reveals that the abundances of over 400 fungal proteins significantly changed in response to the reduced abundance of endosymbiont in normal growth medium, providing functional insights into fungal host-endosymbiont interaction.

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. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02- 05CH11231.

206. Plant-Microbe Interfaces: Correlotypes: Discovering complex, heterogenous genotypes in the Populus pan genome responsible for phenotypes and microbiomic associations.

Debra A. Weighill1,2, Gerald A. Tuskan1, David W. Ussery1 and Dan A. Jacobson1* (jacobsonda@ornl.gov)

1Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; 2Bredesen Center for Interdisciplinary Research and Graduate Education, University of Tennessee, Knoxville, TN.

http://PMI.ornl.gov

Project Goals: The goal of the PMI SFA is to understand the genome-dependent molecular and cellular events involved in establishing and maintaining beneficial interactions between plants and microbes. Populus and its associated microbial community serves as the experimental system for understanding how these molecular events manifest themselves within the spatially, structurally, and temporally complex scales of natural systems. To achieve this goal, we focus on 1) characterizing host and environmental drivers for diversity and function in the Populus microbiome, 2) utilizing microbial model system studies to elucidate Populus-microbial interactions at the molecular level and dissecting the signals and pathways responsible for initiating and maintaining microbial relationships and 3) develop metabolic and genomic modeling of these interactions to aid in interpreting the molecular mechanisms shaping the Populus-microbial interface.

Genome-wide association studies (GWAS) have focused on the analysis of individual single nucleotide polymorphisms (SNPs) in an attempt to find single alleles responsible for a phenotype. Although this has proven useful, it does not account for the fact that many phenotypes are the result of a combination of a broad range of genomic variants and cannot be simply described or controlled by a single gene or variant. Here we propose to develop a method with which to find the complex, heterogeneous collections of genomic variants responsible for many phenotypes. A number of different types of genome variants can all affect the phenotype of an organism including SNPs, small insertions/deletions (INDELs) and larger INDELs that equate to gene knockouts or replications.

Biological organisms are complex systems that are composed of pleiotropic functional networks of interacting molecules and macro-molecules. By the very nature of these pleiotropic networks many phenotypes are multigenic and, as such, will not follow classical Mendelian laws of inheritance and are thus less amenable to discovery by traditional linkage disequilibrium or genome-wide association methods that effectively only consider one SNP at a time or very localized genomic regions such as haplotypes. Genome sequences have previously been generated for over 1000 genotypes of Populus. We are mapping the resulting 100 billion reads to the reference genome in order to confirm bi-allelic and multi-allelic SNPs as well as discover small and large INDELs. It appears that roughly 10 billion reads will not map to the Populus reference genome. We are treating this as a pooled metagenome and assembling it as such. Preliminary evidence suggests that the majority of these reads actually form the Populus pan genome. We can also detect thousands of different microbial and viral species in these reads, thus constituting the endophytic microbiome. We intend to use this meta-assembly to define a pan genome for Populus. We will then map the reads from individual samples to this overall pan genome in order to determine the full compliment of SNPs, large and small INDELs present across these 1000 genomes.

Complex phenotypes are the result of somewhat heterogeneous collections of genome variants. However, the effects of these variants are collectively subject to selective pressure and, as such, their co-occurrence

can be seen as genome-wide correlations. We will calculate the correlations between all pairs of genome variants. The correlations above 0.68 will be used to create a correlation network. Breadth first searches will be done on this correlation network in order to determine an exhaustive collection of sets of variants that we will refer to as correlotypes. We will use correlotypes in combination with a new set-based agglomerative statistical method in order to associate collections of heterogeneous genomic variants with complex phenotypes. We will be testing these correlotype profiles against a range of phenotypic variables, including morphological, microbiomic and molecular profiles, resulting in thousands of phenotypes to test for complex genotypic associations.

To our knowledge, genome variant correlotypes have never been used at this scale before. One set of correlations from 4 million SNPs requires 32 trillion correlation coefficients to be calculated which would require a Petabyte of disk space should we chose to store them all. We will be generating many different sets of correlotypes based on global and local (phenotypic partitioning) correlations across the entire Populus pan genome. Thus, this will present the need for significant high performance computing and storage resources. The networks that result from these comparisons will contain millions of nodes and probably billions of edges.

The methods being developed here are designed to accelerate advances in plant-based bio-energy feed stocks, crop improvement and to further elucidate plant-microbial interactions. Although this approach will initially be used on the Populus 1000 genomes dataset the method itself is species agnostic and can be used in any project wishing to tie complex phenotypes to profiles of heterogeneous genomic variants. Not only have genomic variant correlotypes never been attempted at this scale but, they have also never been tested for significance with the sophisticated combinatorial GSA-based approach that we are developing.

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. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02-05CH11231.

207. Plant-Microbe Interfaces: Probing the Molecular Mechanisms of Plant-Bacterial Interactions

Amber Bible1, Yasu Oda2, Jennifer Morrell-Falvey1* (morrelljl1@ornl.gov), Amy Schaefer2, Dale Pelletier1, Carrie Harwood2, and Mitchel Doktycz1

1Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; 2Department of Microbiology, University of Washington School of Medicine, Seattle, WA.

http://pmi.ornl.gov

Project Goals: The goal of the PMI SFA is to understand the genome-dependent molecular and cellular events involved in establishing and maintaining beneficial interactions between plants and microbes. Populus and its associated microbial community serves as the experimental system for understanding how these molecular events manifest themselves within the spatially, structurally, and temporally complex scales of natural systems. To achieve this goal, we focus on 1) characterizing host and environmental drivers for diversity and function in the Populus microbiome, 2) utilizing microbial model system studies to elucidate Populus-microbial interactions at the molecular level and dissecting the signals and pathways responsible for initiating and maintaining microbial relationships and 3) develop metabolic and genomic modeling of these interactions to aid in interpreting the molecular mechanisms shaping the Populus-microbial interface.

In this project, we are focused on understanding Populus–microbe interactions at the molecular level to dissect the signals and pathways important for initiating and maintaining symbiotic relationships with Populus. Our goal is to elucidate molecular, spatial, and temporal dynamics involved in Populus-microbe interactions using systems biology approaches and directed analytical methodologies. We are interested in how bacteria selectively respond to and become associated with Populus; and how microbially induced molecular and cellular events impact plant growth, health, and fitness. Ultimately, these data will be used to construct model plant- microbial communities to better understand the underlying rules to community assembly and the functional contributions that result from arrangements of multiple organisms.

Our current research is focused on dissecting the signaling pathways involved in plant-bacterial interactions using select Populus-derived bacterial isolates that were chosen based on phenotypic screens and genome inventory data. One area of focus in these isolates is cyclic-di-GMP signaling, which often controls exopolysaccharide (EPS) production, biofilm formation, motility, and other colonization factors. In the robust root colonizer Pantoea sp. YR343, we have employed transposon mutagenesis using a strain that overexpresses the diguanylate cyclase orf2884 in order to identify gene products that function in response to high levels of cyclic-di- GMP production. We have identified at least 60 different gene products, several of which are predicted to have a role in exopolysaccharide (EPS) production and transport, as well as several transcription factors. RNAseq analyses of the c-di-GMP overexpressing strain identified many more candidate genes whose expression is controlled by c-di-GMP levels. By combining these datasets, we have identified a subset of genes on which to focus our initial colonization studies. We are currently working to characterize the functions of these gene products in biofilm formation, EPS production, motility and plant colonization. These mutants provide a set of tools for developing a broader understanding of the molecular mechanisms involved in root colonization by bacteria.

A genomic analysis of Populus isolates belonging to the Proteobacteria, revealed that many possess a newly discovered, but not very well explored plant–bacterial signaling system called the OryR system. Originally discovered in the rice pathogen Xanthomonas oryzae, OryR, which is a transcription regulator, responds to an unknown signal in rice plant extracts to control virulence. We found that an OryR

homologue, now called PipR, in the Populus endophyte, Pseudomonas sp. GM79 responds to Populus leaf macerates to activate gene expression. The genomic region surrounding the GM79 plant-responsive pipR gene includes two genes annotated as peptidases and also genes coding for a putative ABC-type peptide transporter. Pseudomonas sp. GM79 strains that we constructed with mutations in the putative peptidases showed increased responses to Populus macerates. A strain with a mutation in a gene coding for the putative ABC- type peptides transporter did not respond to Populus leaf macerates. We hypothesize that the plant signal(s) enters the bacterial cells by active transport and that the peptidases affect the activity of the signal. We have partially purified the signal and the purified material can be partially inactivated by one of the peptidases. We believe that a better understanding of these OryR-type plant signal receptors and their plant signals is of general importance as they occur in dozens of bacterial species that are associated with economically important plants. Our analysis of the PipR system in a Populus-associated strain opens up the door to studies of a specific Populus-bacterial interaction that is previously unexplored.

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. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02- 05CH11231.

208. Plant-Microbe Interfaces: Carbon Utilization Reflects Nutritional Environment of Pseudomonas Isolates

Collin M Timm1, Se Ran Jun1, Tse-Yuan Lu1, Priya Ranjan2, Tim Tschaplinski1, David J Weston1, and Dale A Pelletier1* (pelletierda@ornl.gov)

1Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; 2University of Tennessee, Knoxville, TN.

http://pmi.ornl.gov

Project Goals: The goal of the PMI SFA is to understand the genome-dependent molecular and cellular events involved in establishing and maintaining beneficial interactions between plants and microbes. Populus and its associated microbial community serves as the experimental system for understanding how these molecular events manifest themselves within the spatially, structurally, and temporally complex scales of natural systems. To achieve this goal, we focus on 1) characterizing host and environmental drivers for diversity and function in the Populus microbiome, 2) utilizing microbial model system studies to elucidate Populus-microbial interactions at the molecular level and dissecting the signals and pathways responsible for initiating and maintaining microbial relationships and 3) develop metabolic and genomic modeling of these interactions to aid in interpreting the molecular mechanisms shaping the Populus-microbial interface.

Plant roots are a rich source of carbon in a carbon-poor soil environment and are host to a diverse microbiome. Root zones can be segregated into two distinct compartments: the endosphere, defined as the internal environment of the root, and the rhizosphere, the volume of soil directly influenced by the presence of the root. The rhizosphere is high in organic acids that are generated by active exudation from the root. Bacterial strains from the genus Pseudomonas were isolated from both the rhizosphere and endosphere compartments of Populus deltoides root samples. Analysis of shotgun whole genome sequencing results revealed that rhizosphere isolates have significantly smaller genomes ($p<0.05$) and have functions biased toward metabolic processes and chemotaxis. Metabolic reconstructions were built, utilizing KBase tools, for each genome and were used to predict substrate utilization bias that may reflect nutritional environment of isolate. Genome informed predictions were tested using substrate utilization data for growth on a panel of 190 carbon substrates (Biolog PM1 and PM2 plates). Among the carbon substrates tested, 43 where utilized by all strains, 78 showed differential utilization and 69 were not utilized by any strain. While single molecules were not predictive of isolation compartment, grouping the molecules into larger classes based on functional groups showed that endosphere isolates are biased in utilization of modified sugars and nucleosides, and rhizosphere isolates are biased towards utilization of carboxylic acids and amino acids. These compound groups reflect the environment of isolation for the strains. This work provides insight into the chemical makeup of the endosphere and rhizosphere compartments, and highlights the importance of metabolism in studies of microbiome structure and function.

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.

209. Plant-Microbe Interfaces: Understanding the factors shaping microbiome structure and function within Populus species.

Michael S. Robeson II1,3, Melissa A. Cregger1, Migun Shakya1,4, Nathan Cude1,5, Gregory M. Bonito2,6, Zamin K. Yang1, Nancy Engle1, Mitchel J. Doktycz1, Intawat Nookeaw1, Mircea Podar1, Susannah G. Tringe7, Dale A. Pelletier1, Wellington Muchero1, Timothy J. Tshaplinski1, Gerald A. Tuskan1, David W. Ussery1, Rytas Vilgalys2, and Christopher W. Schadt1* (schadtcw@ornl.gov)

1Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; 2Department of Microbiology, Duke University, Durham, NC; 3Colorado State University, Fort Collins, CO; 4Dartmouth College, Hanover, NH; 5Novozymes Corporation, Raleigh, NC; 6Michigan State University, East Lansing, MI; 7DOE Joint Genome Institute, Walnut Creek, CA.

http://PMI.ornl.gov

Project Goals: The goal of the PMI SFA is to understand the genome-dependent molecular and cellular events involved in establishing and maintaining beneficial interactions between plants and microbes. Populus and its associated microbial community serves as the experimental system for understanding how these molecular events manifest themselves within the spatially, structurally, and temporally complex scales of natural systems. To achieve this goal, we focus on 1) characterizing host and environmental drivers for diversity and function in the Populus microbiome, 2) utilizing microbial model system studies to elucidate Populus-microbial interactions at the molecular level and dissecting the signals and pathways responsible for initiating and maintaining microbial relationships and 3) develop metabolic and genomic modeling of these interactions to aid in interpreting the molecular mechanisms shaping the Populus-microbial interface.

Populus spp. (Poplar and Cottonwood) are broadly distributed in the temperate environments of North America. They are typical of important riparian habitats, making them ideal ecological model species. Additionally, their rapid growth rates, ability to form both ecto- and endomycorrhizae, clonal propagation, and growth on land otherwise not suitable for food production, make them good candidates for bioenergy production scenarios. Because of these reasons, Populus are also now emerging as a model system for understanding the role of the plant microbiome. We have been investigating microbial communities of Populus deltoides, P. trichocarpa and their TxD hybrids in natural riparian habitats as well as clonal plantation populations. Community level microbiome structure is examined with Roche-454 and Illumina- MiSeq analyses of rRNA amplicons, and we are currently developing techniques for metagenomic analysis. Microbiome data for both bacteria and fungi are analyzed against the corresponding bulk soil properties and chemistry, tree phenotypic measures, tree metabolomics data, and tree genotype data in order to understand how such properties influence microbiome structure. Most work to date has focused on root endosphere and rhizosphere communities, but we have also recently examined a variety of different plant tissue/sample type across the ecosystem that is a Populus tree.

In our studies of P. deltoides we have shown that the environments of rhizo- and endosphere compartments feature highly developed, diverse and to a large degree often exclusive communities of bacteria and fungi. Endophytic bacterial diversity is found to be highly variable, but typically contains tenfold lower diversity than the rhizosphere, suggesting root tissues provide a distinct environment supporting relatively few species more heavily dominated by Actinobacteria and γ-Proteobacteria when compared with the rhizosphere. Populus spp. appear to be highly enriched for Pseudomonas fluorescens-like species/OTUs in the endosphere when compared to rhizosphere habitats, and as compared to the endophytic habitats of surrounding (non-Populus) tree species. Fungal endophytic communities are more diverse than corresponding bacterial diversity, but less diverse than corresponding fungal rhizosphere

communities. Both fungal and bacterial rhizosphere samples showed distinct phylogenetic composition patterns compared to the more variable endophyte samples. Contrary to initial expectations, Populus deltoides has low natural levels of colonization by ectomycorrhizal (ECM) and arbuscular mycorrhizal fungi, but high levels of presumed fungal endophytic taxa such as, Mortierrella, Ilyonectria and members of the Atractiellales. However at the overall community level endosphere, rhizosphere, and soil communities more closely resemble others of the same type, regardless of the plant species from under which they were collected.

Overall P. trichocarpa communities studied in the Western US in two separate Oregon common gardens have shown that rhizosphere and soil microbiomes significantly group by the common garden site at which they were sampled: Clatskanie and Corvallis. There is also a significant separation of endophyte microbiome communities by common garden although not as strong as soil and rhizosphere. While at the overall community level there appeared to be no pattern of Populus genotype specific selection among clones, at the OTU level certain bacteria appear to be enriched or depleted in the endosphere of several clones across all replicates. Pseudomonad OTUs are highly enriched in both the rhizosphere and endophyte samples while being nearly undetectable within bulk soil. Additionally, in contrast to our prior work in P. deltoides, Acinetobacter were also prominent in many P. trichocarpa endosphere samples. Finally, the effects of P. trichocarpa genotype on the composition of its root microbiome appear to be limited compared to the effects of local soil environment at the community level, but OTUs/species may respond to genotypic/phenotypic specific cues. We are currently reanalyzing these results in the context of newly available metabolomic profiles to clarify mechanisms for such responses.

We have also been developing methods for enriching microbiomes from plant tissues for metagenomic sequencing that avoid host DNA background contamination. Ongoing research will take advantage of these metagenomic techniques as well as 1) move beyond the rooting zone to total microbiome studies of Populus inclusive of stem and leaf tissue, and 2) functional studies of defined phenotypic and genotypic variants of Populus to clarify the mechanisms and effects of microbiome interactions in natural and agroforestry environments.

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. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02- 05CH11231.

210. Plant-Microbe Interfaces: Constructed plant and bacterial communities for understanding and predicting community function.

Collin M. Timm1, Jeremiah Henning1,2, Sarah Jawdy1, Lee Gunter1, Nancy Engle1, Tim Tschaplinski1, Gerald A. Tuskan1, Mitchel J. Dokytcz1, Dale A. Pelletier1, and David J. Weston1* (westondj@ornl.gov)

1Oak Ridge National Laboratory, Oak Ridge, TN; 2Graduate School of Genome Science and Technology, University of Tennessee-Knoxville, Knoxville, TN.

http://pmi.ornl.gov

Project Goals: The goal of the PMI SFA is to understand the genome-dependent molecular and cellular events involved in establishing and maintaining beneficial interactions between plants and microbes. Populus and its associated microbial community serves as the experimental system for understanding how these molecular events manifest themselves within the spatially, structurally, and temporally complex scales of natural systems. To achieve this goal, we focus on 1) characterizing host and environmental drivers for diversity and function in the Populus microbiome, 2) utilizing microbial model system studies to elucidate Populus-microbial interactions at the molecular level and dissecting the signals and pathways responsible for initiating and maintaining microbial relationships and 3) develop metabolic and genomic modeling of these interactions to aid in interpreting the molecular mechanisms shaping the Populus-microbial interface.

Our goal is to study and understand the community function of the Populus host plant and its associated microbiome by leveraging our 1,039 genome sequenced Populus clones, collection of >2700 bacterial isolates, functional genomic and physiological assays from plants with individual microbes, and community data from wild Populus trees. Our results indicate that strains alone and in combinations of 10 or less can have a drastic influence on plant morphology and biochemistry. For example, the colonization of Pseudomonas fluorescens strain GM41 on plant roots has a strong effect on leaf metabolome profiles. This effect is also observed in combination with Burkholderia sp. BT03, which increases root biomass in individual and four member communities. The abundance of GM41 and BT03 are equal in individual or mixed experiments, suggesting that these strains occupy different root regions or niche space. As an initial attempt to model plant–microbe interactions, we recently coupled an N-fixing metabolic model of a diazotrophic endophyte with a plant host physiological model to inform nitrogen status and constrain the uncertainty surrounding net photosynthesis predictions and community biomass estimates. To extend this model and further evaluate how host carbon status influences the microbiome, we recently imposed a light stress limitation that altered carbon allocation patterns. We are currently investigating how this treatment will influence root exudate composition and concentration that are growth substrates for the microbiome. These data and modeling endeavors will be used to generate hypotheses for testing community biodesign strategies for sustainable bioenergy feedstock production.

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. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02- 05CH11231.

211. Plant-Microbe Interfaces: Fungal-bacterial interactions within Populus rhizosphere

Cora Guennoc1,2, Jessie Uehling3, Gregory Bonito3, Rytas Vilgalys3, Francis Martin2, Mitchel J. Doktycz1, Gerald A. Tuskan1, Aurelie Deveau2 and Jessy Labbé1* (labbejj@ornl.gov)

1Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; 2INRA, Tree-Microbe Interactions, UMR1136, Champenoux, France; 3Biology Department, Duke University, Durham, NC.

http://PMI.ornl.gov

Project Goals: The goal of the PMI SFA is to understand the genome-dependent molecular and cellular events involved in establishing and maintaining beneficial interactions between plants and microbes. Populus and its associated microbial community serves as the experimental system for understanding how these molecular events manifest themselves within the spatially, structurally, and temporally complex scales of natural systems. To achieve this goal, we focus on 1) characterizing host and environmental drivers fro diversity and function in the Populus microbiome, 2) utilizing microbial model system studies to elucidate Populus-microbial interactions at the molecular level and dissecting the signals and pathways responsible for initiating and maintaining microbial relationships and 3) develop metabolic and genomic modeling of these interactions to aid in interpreting the molecular mechanisms shaping the Populus-microbial interface.

The molecular events leading to recognition and colonization of a host plant by beneficial microorganisms are poorly understood. Our ongoing research is aimed at identifying and isolating microbes associated with natural Populus ecosystems in order to determine molecular, genetic and cellular events involved in recognition and establishment of beneficial microbial interactions with Populus. Several research investigations to enhance our understanding of plant- microbe interactions focus on elucidating the genetic and molecular mechanisms of the interactions of the host plant and the bacterial network associated with fungal partners within the natural Populus ecosystems.

The mycorrhizal symbiosis is the most widespread plant-microbe association that supports forest growth and sustainability. In soil, roots and fungi are surrounded by diverse microbial communities, which modulate the mycorrhizal symbiosis. This community includes the so-called mycorrhiza helper bacteria (MHB), which are thought to assist mycorrhizal formation and symbiosis. Other bacteria are confined to living inside of fungal hyphal cells and cannot sustain life independently outside of their hosts. Because very little is known about the role of these helper and endosymbiotic bacteria in Populus-fungi interactions, this project is aimed at dissecting the signaling mechanisms underlying Populus–fungal–bacterial interactions. To this end, our experimental approach integrates transcriptomes, metabolomes, proteomes, and genomes. We have sequenced genomes of several endosymbionts and free-living helper bacteria, constructed mutant libraries and have begun mutant phenotype screening. We demonstrate that some helper bacterial strains influence Populus–L. bicolor colonization and some mutants are affected in their beneficial effect. In parallel we have characterized diverse endosymbiotic bacterial communities in several fungal Mortierella elongata strains, a fast-growing coenocytic fungus found in association with Populus. We then cleared bacterial endosymbionts from the strains and observed various effects on fungal host fitness in specific conditions. Hereby we provide new insights into the mechanism of multi-partite interaction between Populus and its complex microbial communities with clear synergistic effects on plant growth, stress tolerance, and fitness.

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. The work conducted

by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02- 05CH11231.

212. Plant-Microbe Interfaces: Genome re-sequencing reveals a species-specific whole-gene deletion associated with Populus-Laccaria mycorrhizal symbiosis

Wellington Muchero1* (mucherow@ornl.gov), Juan Wang1, Jessy Labbé1, Olaf Czarnecki1, Jin- Gui Chen1, Priya Ranjan2, Xiaohan Yang1, Luke Evans3, Stephen DiFazio3, Anthony Bryan1, Sara Jawdy1, Lee Gunter1, Wendy Schackwitz4, Joel Martin4, Christopher W. Schadt1, Gerald A. Tuskan1

1Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; 2University of Tennessee; Knoxville, TN; 3West Virginia University; Morgantown, WV; 4DOE Joint Genome Institute; Walnut Creek, CA.

http://PMI.ornl.gov

Project Goals: The goal of the PMI SFA is to understand the genome-dependent molecular and cellular events involved in establishing and maintaining beneficial interactions between plants and microbes. Populus and its associated microbial community serves as the experimental system for understanding how these molecular events manifest themselves within the spatially, structurally, and temporally complex scales of natural systems. To achieve this goal, we focus on 1) characterizing host and environmental drivers for diversity and function in the Populus microbiome, 2) utilizing microbial model system studies to elucidate Populus-microbial interactions at the molecular level and dissecting the signals and pathways responsible for initiating and maintaining microbial relationships and 3) develop metabolic and genomic modeling of these interactions to aid in interpreting the molecular mechanisms shaping the Populus-microbial interface.

Mycorrhizal symbiosis between perennial plants and fungal associates has critical implications for diverse phenomena including global carbon, water and nutrient cycling, as well as agricultural and forestry productivity with limited inputs on marginal croplands. As such, characterizing the molecular genetics underlying such interactions holds tremendous potential in engineering biological systems for enhanced carbon sequestration and sustainable biomass production.

In the perennial bioenergy feedstock Populus, numerous studies have demonstrated the species-dependent colonization efficiency by the fungal symbiont Laccaria bicolor, with P. trichocarpa exhibiting high levels of mycorrhization compared to P. deltoides. This highly species-specific attribute of the interaction presents an opportune platform for the discovery of host genetic factors governing mycorrhizal interactions using inter-specific hybrids. To this end, we identified a major quantitative trait locus (QTL) contributing up to 60% of the phenotypic variance explained (PVE) in colonization of P. trichocarpa x P. deltoides F1 hybrids by Laccaria. Genome anchoring of this QTL using single nucleotide polymorphism (SNP) markers with known physical positions revealed its co-location with a region harboring tandemly repeated lectin-type receptor kinases. Alignment of the P. trichocarpa and P. deltoides re-sequenced parental genomes suggested major structural differences in this region including a whole-gene deletion event in P. deltoides involving a D-mannose lectin receptor kinase. Analysis of allelic effects of the indel revealed that individuals carrying a full copy of the gene exhibited 2X more colonization by the fungal symbiont compared with individuals missing segments of the same gene. Further, we screened pure P. trichocarpa and P. deltoides natural variants to assess penetrance of the indel in the species' natural habitats. We could not detect a full copy of the gene in any of the 60 P. deltoides genotypes collected from diverse geographical origins in eastern United States whereas the gene was highly conserved in 673 re-sequenced P. trichocarpa genomes evaluated. Since D-mannose receptor kinases have been implicated in innate immunity and self-incompatibility responses, which require highly specific recognition of cells and microorganisms, we hypothesize that this indel polymorphism contributes substantially to the

species-specificity observed in Populus interaction with Laccaria. Transgenic validation of putative effects of the receptor kinase on mycorrhization is currently underway and results of these analyses will be presented.

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.

213. Plant-Microbe Interfaces: Comparative Genomics of Populus endosphere microbiomes.

Se-Ran Jun1, Sanjeev Dahal2, Qian Zhang2, Visanu Wanchai1, Tse-Yuan Lu1, Miriam Land1, Intawat Nookaew1, Chris W. Schadt1, Collin M. Timm1, Dale A. Pelletier1, and David W. Ussery1* (usserydw@ornl.gov),

1Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; 2Graduate School of Genome Sciences and Technology, University of Tennessee, Knoxville, Tennessee

http://PMI.ornl.gov http://dtree.ornl.gov

Project Goals: The goal of the PMI SFA is to understand the genome-dependent molecular and cellular events involved in establishing and maintaining beneficial interactions between plants and microbes. Populus and its associated microbial community serves as the experimental system for understanding how these molecular events manifest themselves within the spatially, structurally, and temporally complex scales of natural systems. To achieve this goal, we focus on 1) characterizing host and environmental drivers for diversity and function in the Populus microbiome, 2) utilizing microbial model system studies to elucidate Populus-microbial interactions at the molecular level and dissecting the signals and pathways responsible for initiating and maintaining microbial relationships and 3) develop metabolic and genomic modeling of these interactions to aid in interpreting the molecular mechanisms shaping the Populus-microbial interface.

Populus is a widely studied model woody plant species and a potential cellulosic feedstock for biofuels. These trees are also host to a wide variety of microbial associations within their roots and rhizosphere and thus serve as a model to study interactions between plants and microorganisms. The Populus microbiome contains perhaps a thousand different bacterial species, although most are in low abundance. Over the past several years, the PMI SFA has isolated >3000 bacterial strains from Populus roots representing >425 distinct OTUs. The genomes for >100 of these strains were sequenced in collaboration with the DOE Joint Genome Institute. Here we compare three different genera for which we have genomes for >15 strains each. These include Rhizobium strains, Pseudomonas strains, and Chryseobacterium strains. Members of these genera are commonly found in the rhizosphere and endosphere compartments of Populus roots. Comparative genomics of these endophytes are done, within the broader context of all known sequenced genomes form the same genera. We describe a set of 31 Rhizobium strains that we isolated from the endosphere and rhizosphere compartments of Populus roots. We compare 162 genomes, from 18 different names species within the Rhizobium genus, as well as a set of 45 genomes without a species designation (e.g., Rhizobium sp. CF142). For the Pseudomonas genus, we compared all publicly available Pseudomonas genomes (65 complete and 1008 draft genomes) including 21 strains sequenced as part of the PMI SFA. For the Chryseobacterium genus, we compare 18 PMI strains with 36 publically available strains. Additionally, phenotypic differences in plant root architecture during co-cultivation, phosphate solubilizing activity and carbon utilization indicate functional diversity among the isolates. For comparative genomics-based systematics, we clustered genomes based on genomic relatedness based on average amino acid identity (AAI). The local relationship of pairwise genomic relatedness combined with global relationship of genomic clusters on a whole proteome tree resolves species (e.g., Pseudomonas fuscovaginae, Pseudomonas nitroreducens) not settled by 16S rRNA gene sequence analysis. Here we will present results from “core” and “pan” genome analyses. The core genomes of genomic clusters showed very similar functional distributions, which suggests that specific genes to genomic clusters may explain the influence of microorganisms’ ecology on their functional change. Furthermore, we examined the differences of strains by pathway profiles, carbohydrate-active enzyme profiles, and other specific genes. In addition, metabolic models were generated for specific strains and experiments performed, and

their results were compared with predictions based on the genome sequences. This analysis provides insight into the genotype/phenotype relationship and identification of species-specific gene families, and in some cases, it is possible to predict unique functions/ecological niches for a given species.

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. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02- 05CH11231.

214. Assessing In Situ Metabolic Carbon Dynamics in a Phototrophic Microbial Mat

James Moran1* (james.moran@pnnl.gov), Mary Lipton1, Alexandra Cory1, Eric Huang1, Krystin Riha1, Jennifer Mobberly1, William Nelson1, Marina Gristsenko1, Young-Mo Kim1, Thomas Metz1, and James Fredrickson1

1Pacific Northwest National Laboratory, Richland, Washington

http://www.pnnl.gov/biology/programs/fsfa/

Project Goals: The PNNL FSFA is focused on discovering fundamental principles that determine microbial community structure, function, and dynamics, and applying these principles to rational design of microbial communities for desirable outcomes. Our strategy integrates field studies of microbial mat communities and laboratory studies of field-derived consortia and isolates to gain a mechanistic understanding of the environmental and biotic drivers of community higher-order properties such as resistance and resilience. Genome sequence-enabled approaches, including chemical probes directed against specific functions, are being used to elucidate underlying interaction mechanisms. These data are in turn being used to inform community metabolic, spatial and regulatory network models. Our research plan supports DOE goals to achieve a predictive understanding of microbes and microbial communities and to provide foundational knowledge enabling rational design of microbial systems.

We tracked C from stable isotope (13C) labeled substrates into a phototrophic microbial mat system to assess metabolic interactions responsible for directing C flow over a complete diel cycle. We used substrates directed towards photoautotrophic (13C bicarbonate addition) and heterotrophic (13C-glucose and 13C-acetate additions) members of the community to specifically target C exchange between these functional groups. The different fates of the 13C-substrates were assessed in terms of chemical transformation (catabolic remineralization versus anabolic uptake to different biomass constituents), spatial localization (along a depth profile), functional utilization (protein and storage polymer synthesis), and phylogenetic activity (specific protein synthesis).

Hot Lake is a hypersaline, meromictic lake located in an endorheic basin in north-central Washington which, despite extreme salinity and seasonal water temperatures (> 55 ˚C), hosts dense, phototrophic benthic microbial mats. We extracted mature mat from this system (figure 1) and performed lake-side, batch incubations using ambient light, native lake water, and while maintaining natural temperature. A single pulse of labeled substrate (13C bicarbonate, glucose, or acetate) was added to different incubations and we performed timeseries subsampling around a complete diel cycle (noon to noon).

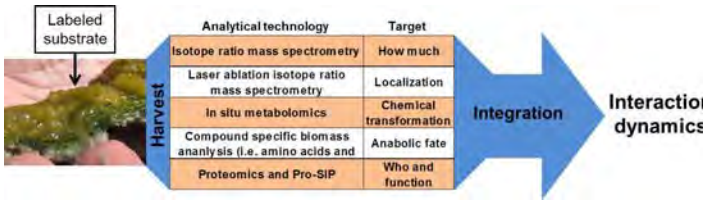


Figure 1: Experimental approach. Mat is harvested then incubated in the presence of ^{13}C-labeled substrates and subsampled over a diel timeseries. A suite of analytical techniques are applied to track overall carbon uptake into the mat and its spatial, chemical, and functional fate.

Bulk stable isotope mass spectrometry revealed distinct patterns between the different substrates in terms of uptake into mat biomass. Bicarbonate and acetate uptake tracked the diel cycle (presumably governed by light and oxygen availability respectively) while glucose showed a nearly uniform uptake rate despite large-scale diel shifts in redox conditions within the mat. Sample washing with dilute HCl removed a significant fraction of 13C associated with bicarbonate and glucose uptake, suggesting conversion of 13C

from these substrates into to acid-labile or water soluble components. In comparison, little acetate-derived 13C was removed by the HCl washing. We speculate that a portion of the removed C can be attributed to acid labile and water soluble EPS components (i.e. polysaccharides), suggesting preferential conversion of bicarbonate and glucose-derived C to EPS versus acetate C. 13C-metabololite analysis demonstrated anabolic uptake of applied substrates into various monomers (i.e., glucose conversion in a distinct C6 unit directly into trehalose) that serve in osmotic regulation or as metabolic storage compounds. Similar processes (using minimal chemical bond transformations for synthesis) may enable high rates of conversion of extracellular glucose C to EPS components. Interestingly, however, is the lack of water-soluble, acid labile C derived from acetate 13C. We are exploring whether acetate may be preferentially shunted towards other compound classes (e.g. lipids).

In incubations with either of the organic substrates (glucose and acetate) we observed a large 13C flow towards remineralization to dissolved inorganic carbon (DIC) at multiple points in the diel cycle. We additionally observed conversion of C from all substrates to solid-phase, inorganic carbon (e.g., carbonate) although this represented a relatively small C flux and the mechanism for this conversion is unclear, but likely includes intermediate remineralization to DIC.

Proteomic analysis was enabled by metagenome sequences of unicyanobacterial consortia (UCC) developed from this system (See Nelson poster for details). For all three substrates, there was a higher C flux into phototroph associated proteins versus those of heterotrophic organisms. Both bicarbonate and glucose derived 13C showed a larger flux into heterotrophic proteins than acetate 13C. The relative high abundance of bicarbonate derived 13C in the heterotroph proteins highlights the close C coupling between autotrophs and heterotrophs; autotroph fixed C being converted to heterotroph proteins as quickly as or more quickly than exogenously applied glucose or acetate respectively. We also observed diel dependence in conversion of 13C substrates to protein with the bulk of conversion not occurring until after the diel dark period. We hypothesize the mat undergoes day-time nitrogen limitation and that freshly fixed N2 during the night, when the mat goes anoxic, may be needed for protein synthesis.

Taken together, our data highlights the different C conversions and interactions revealed in this mat by the combined stable isotope, metabolomics, and proteomic approach we applied. Specifically, the three applied substrates showed distinct patterns of uptake into microbial biomass; bicarbonate 13C preferentially converted to protein, glucose 13C showing the highest conversion to water/HCl soluble compounds, and acetate 13C to water/HCl insoluble compounds. Further, the higher conversion of bicarbonate 13C to heterotroph associated proteins highlights the tight metabolic interactions between these groups as autotrophically fixed C was preferred for protein synthesis versus exogenous C.

This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Genome Sciences Program (GSP). This contribution originates from the GSP Foundational Scientific Focus Area (FSFA) at the Pacific Northwest National Laboratory (PNNL).

215. Metagenomic Sequencing of Parallel Consortia Enables Genome Reconstruction of Rare and Closely Related Organisms

William C. Nelson1* (william.nelson@pnnl.gov), Margie F. Romine1, Eric Huang1, Mary S. Lipton1, Stephen R. Lindemann1, and James K. Fredrickson1

1Pacific Northwest National Laboratory, Richland, Washington

http://www.pnnl.gov/biology/programs/fsfa/

Project Goals: The PNNL FSFA is focused on discovering fundamental principles that determine microbial community structure, function, and dynamics, and applying these principles to rational design of microbial communities for desirable outcomes. Our strategy integrates field studies of microbial mat communities and laboratory studies of field-derived consortia and isolates to gain a mechanistic understanding of the environmental and biotic drivers of community higher-order properties such as resistance and resilience. Genome sequence-enabled approaches, including chemical probes directed against specific functions, are being used to elucidate underlying interaction mechanisms. These data are in turn being used to inform community metabolic, spatial and regulatory network models. Our research plan supports DOE goals to achieve a predictive understanding of microbes and microbial communities and to provide foundational knowledge enabling rational design of microbial systems.

Metagenomic sequencing has enabled the de novo reconstruction of microbial genomes from environmental communities. Biological complexity has limited this approach to low-diversity communities and abundant members of complex communities. The genomic complement of each member population must be determined, however, to gain a predictive understanding of the dynamics of interactions between populations and the contributions of these interactions to overall community function.

We performed short-read metagenomic sequencing of two unicyanobacterial consortia cultivated from saline (MgSO4) phototrophic microbial mats (previously described in [1]) and reconstructed member species' genomes. These are the most complex communities for which near-complete, species-resolved (and in some cases strain-resolved) genomic information has been generated. Near-complete genome sequences were obtained for 17 of the 20 member species found in the two cultures. Our use of parallel consortia (consortia cultivated under the same conditions from the same natural community and having overlapping species composition) allowed acquisition of near-complete genome sequence for organisms that are rare in the natural community, and detection and deconvolution of microdiversity. Differential membership and abundances between the two consortia revealed inter- and intraspecies microdiversity, resolving two Halomonas sp. and two Rhodobacteraceae sp. ecotypes previously clustered by amplicon analysis. Differences in functional potential (~850 genes in the Halomonas spp.) between the resolved organisms suggest distinct roles in the community. This work demonstrates how significant functional diversity can be hidden within an OTU, and how subtle variations between genomes might result in distinct niches.

In addition, we have been able to assess the accuracy of common practices in genome reconstruction, such as scaffold segregation by differential read coverage and conserved single-copy gene analysis, through comparison of our reconstructed genomes against available complete genome sequences from isolates cultured in parallel to the metagenomic study. Comparison of the reconstructed genomes to the complete genomes of isolated strains demonstrated the sensitivity and specificity of the reconstruction process across a wide range of sequence coverage (9x – 2700x).

The utility of such datasets in interpreting environmental data was demonstrated by screening an environmental proteomic data set against the reconstructed genomes and identifying peptides from 19 of the organisms in the consortia.

Our work represents a critical step toward understanding the mechanisms that drive microbial community formation, function, and response to environmental stimuli. The genomic information generated allows prediction of the functions of member taxa (see Posters by M. Romine and S. Lindemann) and provides a foundation against which experimental transcriptomic and proteomic data can be evaluated. By enabling in toto community genomic reconstruction, the parallel consortia approach serves as a useful intermediate between environmental metagenomics and single-cell amplified genomics.

References:

1. Cole JK, Hutchison JR, Renslow RS, Kim YM, Chrisler WB, Engelmann HE, Dohnalkova AC, Hu D, Metz TO, Fredrickson JK, Lindemann SR: Phototrophic biofilm assembly in microbial-mat-derived unicyanobacterial consortia: model systems for the study of autotroph-heterotroph interactions. Front Microbiol 2014, 5:109.

This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Genome Sciences Program (GSP). This contribution originates from the GSP Foundational Scientific Focus Area (FSFA) at the Pacific Northwest National Laboratory (PNNL). Sequencing was done at the DOE Joint Genome Institute under Contract No. DE-AC02-05CH11231 and Community Science Project 701.

216. Omics and Chemical Probe Strategies for Assessing Interactions Among Microbial Community Members

Margaret Romine1* (Margie.romine@pnnl.gov), Lindsey Anderson1, Eric Huang1, Steve Lindemann1, Mary Lipton1, Jennifer Mobberly1, Premchendar Nandhikonda1, Bill Nelson1, Dmitry Rodionov2, Aaron Wright1, and Jim Fredrickson1

1Pacific Northwest National Laboratory, Richland, Washington 2Sanford-Burnham Medical Research Institute, La Jolla, California

http://www.pnnl.gov/biology/programs/fsfa/

Project Goals: The PNNL FSFA is focused on discovering fundamental principles that determine microbial community structure, function, and dynamics, and applying these principles to rational design of microbial communities for desirable outcomes. Our strategy integrates field studies of microbial mat communities and laboratory studies of field-derived consortia and isolates to gain a mechanistic understanding of the environmental and biotic drivers of community higher-order properties such as resistance and resilience. Genome sequence-enabled approaches, including chemical probes directed against specific functions, are being used to elucidate underlying interaction mechanisms. These data are in turn being used to inform community metabolic, spatial and regulatory network models. Our research plan supports DOE goals to achieve a predictive understanding of microbes and microbial communities and to provide foundational knowledge enabling rational design of microbial systems.

We have used metagenome sequence to assemble complete or near-complete genome sequences from nearly all the members (~20) of two unicyanobacterial consortia (UCC) developed from benthic microbial mats associated with a saline (MgSO4) lake in northern Washington State (see Nelson and Moran posters for details). These compositionally-stable consortia have been maintained under constant light in defined media containing salts, trace minerals, and CO2 as the primary carbon source. This defined system has allowed us to use single organism metabolic reconstruction strategies to infer member function and metabolic requirements and to use these genomic analyses collectively to predict interactions among members and the identity of genes and processes involved. We describe two examples of how the assembled genomes provide detailed insights into microbial community interactions than would not be possible with typical incomplete metagenome sequence data. We also describe development of chemical probes intended to provide visualization of individuals contributing to carbon and nutrient interactions.

Maintaining the consortia under constant light promotes continuous O2 production and highly oxidizing conditions. Under these conditions Fe is present as Fe3+ and that a strong complexing ligand (citrate) be provided to maintain its solubility. We conducted regulon and comparative genomic analysis to assess how different members acquire Fe3+ and maintain internal iron homeostasis. Analyses revealed that while most of the members sense iron levels directly using Fur, the seven Rhodobacteraceae and one Rhizobiales member sense it indirectly via the heme-binding Irr regulator. A review of these predicted regulons suggest that the gamma proteobacterial members have an expanded repetoir of iron uptake systems and that the Halomonas sp. are the only members with detectable siderophore biosynthetic capability. All of the members possess a ferric uptake system (FbpABC) and utilize the SUF rather than ISC iron-sulfur assembly system, likely due to its insensitivity to aerobic stress. Proteomic analysis of the UCC OSCr consortia detected ferric uptake system components from 11 members but provided no evidence for siderophore utilization. On average the number of proteins involved in Fe homeostasis detected accounted for two percent of the total for each organism which is high considering how few proteins are involved in iron metabolism and how few are detected for each organism.

It is well recognized that vitamin auxotrophy and prototrophy provide the basis for metabolic interactions between phototrophs and heterotrophs. By analyzing our UCC we are able to provide the first insights into how auxotrophy might be maintained in a self-sustaining multi-member community. As no external source of vitamin is provided to our cultures, auxotrophs must acquire this resource from other members or employ a metabolic strategy that obviates the need for vitamins that they cannot synthesize. Genomic analysis revealed that only the cyanobacteria and two Halomonas sp. are capable of synthesizing all required vitamins (B1, B2, B3, B5, B6, B7, B9, and B12) or coenzymes derived from them. While all members could produce FAD/FMN (B2) and pyridoxine (B6), auxotrophy for the remaining vitamins are widespread among consortia members. We also identified the vitamin-dependent metabolic processes to assess the burden that sharing these resources might have on the community. This analysis revealed an unexpected relationship between the ability to produce B12 and need for it to mediate essential processes. In general, the auxotrophs encode more enzymes that require B12 than the five producers do. Both Halomonas sp. are B12 opportunists, yet they utilize enzymes or pathways that obviate its requirement under all conditions except growth on ethanolamine suggesting a potential role for these organisms as producers of this commodity. We also discovered that B12 likely acts as a regulatory sensor in the phototrophic Rhodobactericeae and Algoriphagus marincola, controlling photosynthesis and carotenoid biosynthesis, respectively. Its dual role as a sensor and coenzyme suggests that it may be a key determinant controlling the community dynamics.

In order to identify proteins involved in these processes and to monitor the exchange of metabolites under different conditions we have developed activity-based protein probes including mimics of vitamins, amino acids, and sugars. These probes are designed so that they can be taken up by living cells, fixed to targets that they bind, and assayed by image analysis, flow cytometry, or mass spectrometric proteomic characterization. We have validated the specificity of the vitamin probes and successfully used them to identify transporters and enzymes in several different UCC taxa. We have initiated investigations that involve use of the probes to image the distribution of cells that assimilate them under different conditions so that the impact of changing culture and/or environmental conditions on resource allocation can be determined.

This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Genome Sciences Program (GSP). This contribution originates from the GSP Foundational Scientific Focus Area (FSFA) at the Pacific Northwest National Laboratory (PNNL). The work conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of DOE under Contract No. DE-AC02-05CH11231 and Community Sequencing Project 701. A portion of this work (proteomics) was performed in the William R. Wiley Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by OBER and located at PNNL.

217. Using Genome-Resolved Metagenomics to Uncover Trophic Structure in Microbial Communities: Its Influence on Community Dynamics

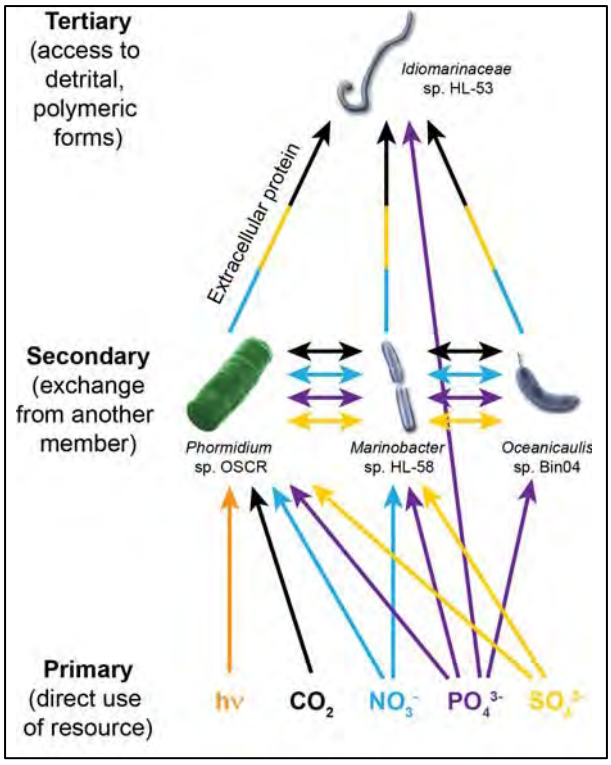
Stephen R. Lindemann1* (Stephen.lindemann@pnnl.gov), Jessica K. Cole1, Jennifer M. Mobberley1, Christopher R. Anderton1, James J. Moran1, William C. Nelson1, Margaret F. Romine1, and James K. Fredrickson1

1Pacific Northwest National Laboratory, Richland, Washington

http://www.pnnl.gov/biology/programs/fsfa/

Project Goals: The PNNL FSFA is focused on discovering fundamental principles that determine microbial community structure, function, and dynamics, and applying these principles to rational design of microbial communities for desirable outcomes. Our strategy integrates field studies of microbial mat communities and laboratory studies of field-derived consortia and isolates to gain a mechanistic understanding of the environmental and biotic drivers of community higher-order properties such as resistance and resilience. Genome sequence-enabled approaches, including chemical probes directed against specific functions, are being used to elucidate underlying interaction mechanisms. These data are in turn being used to inform community metabolic, spatial and regulatory network models. Our research plan supports DOE goals to achieve a predictive understanding of microbes and microbial communities and to provide foundational knowledge enabling rational design of microbial systems.

In contrast to macroorganisms, which tend to occupy clear positions in food webs based upon producer-consumer relationships, the majority of known microbes encode the capacity to assimilate at least some of their nutrients directly in their mineral forms as well as through exchange of organic metabolites. However, the ability to directly assimilate mineral resources (e.g., CO2, NO3-, PO43-, SO42-) is not uniformly distributed among members of microbial communities, constraining some organisms to acquire lacking nutrients through interactions with other members. The standard categorizations of microbial metabolisms (photoautotroph, chemoautotroph, photoheterotroph, and chemoheterotroph) have focused solely upon carbon and energy acquisition. We suggest that trophic levels for microbes should be examined for each element required for life, as such trophic interactions govern the flow of these elements through communities. These flows determine which community members are able to grow, shaping community dynamics. We hypothesize that, under conditions where an organism has primary access to a nutrient, its relative abundance will increase over conditions where its access to the same nutrient is constrained by interspecies interactions. Figure 1. Conceptual diagram of trophic levels potentially occupied by four of the species in consortium UCC-O under standard cultivation conditions. Idiomarinaceae sp. HL-53 can also occupy the secondary trophic levels for all nutrients; these interactions are not pictured for the sake of clarity. Using species-resolved



metagenome bins (see Nelson et al. poster) generated from unicyanobacterial consortia1 derived from microbial mats populating epsomitic Hot Lake, WA2, we evaluated each member species for the genome-encoded mechanisms by which it may acquire energy, carbon, nitrogen, phosphorus, and sulfur. These genome-based predictions were used to assign these organisms to one or more trophic levels they could potentially occupy under given cultivation conditions. Organisms with a predicted capability to assimilate a nutrient directly from an inorganic source were assigned to the primary trophic level, those that can acquire an element via a small molecule produced within the consortium to the secondary trophic level, and those that can depolymerize extracellular macromolecules and salvage monomers to the tertiary trophic level. These assignments permitted a reconstruction of the possible trophic networks within the community. We further investigated the impact of changes in trophic levels with respect to nitrogen upon the community's composition, spatial structure, and patterns of element flow by amending growth medium with ammonium. Because the UCCs are routinely cultivated with excess NO3- as the nitrogen source, the primary organisms by which nitrogen enters the consortia will be those capable of reducing NO3- to NH4+ under standard conditions1. Both cyanobacteria possess cytoplasmic nitrate and nitrite reductases, as do all of the consortial gammaproteobacteria except for Idiomarinaceae sp. HL-53, allowing them to serve as primary producers with respect to nitrogen acquisition. All members of the consortia contain Amt-like NH4+ transporters and can assimilate NH4+ via glutamine synthetase, meaning all of our members can occupy the first trophic level of nitrogen when NH4+ is present. As expected, supplementation with ammonium essentially eliminated nitrate uptake. Ammonium amendment significantly altered the community composition and the biofilm structure of UCC-O in succession, unexpectedly suppressing the growth of species predicted to be unable to directly incorporate NO3- (e.g., Idiomarinaceae sp. HL-53). Furthermore, ammonium amendment increased the total biovolume and density of organisms in UCC-O, and eventually terminated in a crash in autotrophic biomass late in succession. In contrast, ammonium amendment of UCC-A resulted in only minor changes to successional dynamics. Because the consortia share the same suite of heterotrophic members1, these results suggest that these members occupy significantly different niches in each consortium with respect to nitrogen flow that are difficult to predict from the presence or absence of nitrate assimilation pathways. These data suggested large-scale alteration in carbon and nitrogen flow between members, a conclusion further supported by changes in the spatial patterns of H13CO3- and 15NH4+ or 15NO3- as visualized using nanoSIMS. In sum, these results suggest that lack of primary access to a nutrient may not significantly impede an organism's growth within a consortium, but that community compositional and spatial structure are likely determined by large-scale patterns of resource flow. This informs the design of consortia by suggesting that supplementation with resources predicted to favor expansion of a desired species are likely to have unpredictable and distributed impacts upon community structure and function.

References:

1. Cole, J. K. et al. Phototrophic biofilm assembly in microbial-mat-derived unicyanobacterial consortia: model systems for the study of autotroph-heterotroph interactions. Syst. Microbiol. 5 109 (2014).
2. Lindemann, S. R. et al. The epsomitic phototrophic microbial mat of Hot Lake, Washington: community structural responses to seasonal cycling. Front. Microb. Physiol. Metab. 4, 323 (2013).

This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Genome Sciences Program (GSP). This contribution originates from the GSP Foundational Scientific Focus Area (FSFA) at the Pacific Northwest National Laboratory (PNNL). The work conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of DOE under Contract No. DE-AC02-05CH11231 and Community Sequencing Project 701. A portion of this work (nanoSIMS) was performed in the William R. Wiley Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by OBER and located at PNNL.

218. Pan-omics Biological Applications in the Genome Sciences Program

Mary S. Lipton*(mary.lipton@pnnl.gov), Joshua N. Adkins, Thomas O. Metz, Samuel H. Payne, Stephen J. Callister, Kristin E. Burnum, Eric Huang, Aaron T. Wright, Amy A. Boaro, Richard White, Carrie D. Nicora, Meagan C. Burnet, Richard D. Smith

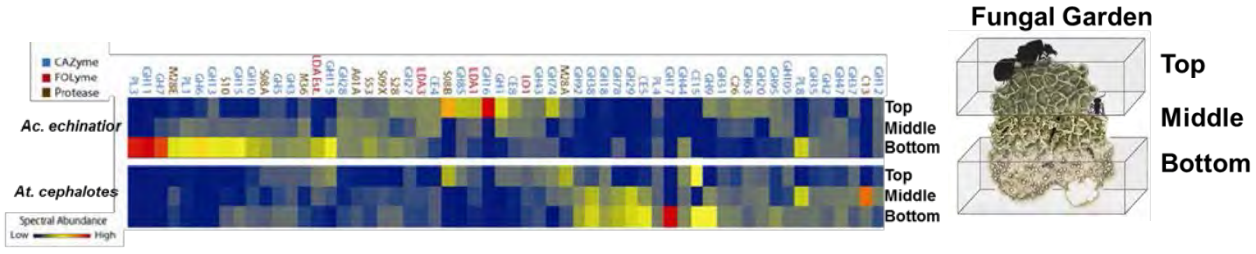
Biological Sciences Division, Pacific Northwest National Laboratory, Richland WA 99352

http://www.omics.pnnl.gov

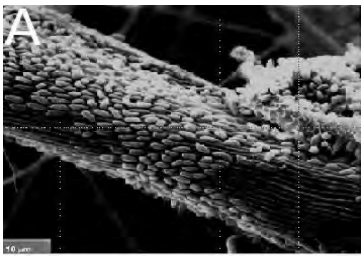
Project Goal: The Pan-omics Program includes efforts that involve a broad spectrum of analytical measurements acquired using advanced mass spectrometry-based measurement pipelines. The Biological Applications portion of the Pan-omics Program works in collaboration with GSP scientists, and utilizes these measurements to advance the fundamental understanding of biological systems important to DOE missions. These pan- omics measurements, integrated with genomic and transcriptomics data, will increasingly enable predictive models, as well as the monitoring, manipulation or control of complex biological systems. Currently, collaborative projects range from studies involving novel microbial species to highly diverse ecosystems including interactions between bacteria, fungi, plants, and insects. The understanding of the complex regulatory and functional interactions necessary to provide predictions for optimal production for reengineered organisms or the elucidation of emergent properties in climate and carbon-cycling process will benefit from a range of pan-omics measurement capabilities.

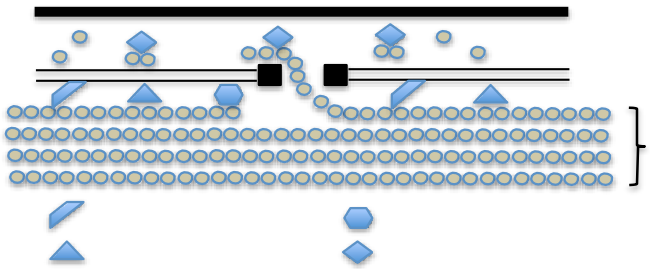
We present below several examples from present collaborations efforts conducted under the Biological Applications portion of the Pan-omics Program. These examples illustrate misson specific areas of interest to the DOE, including terrestrial carbon cycling, ligno-cellulose deconstruction, and the impacts of climate change on ecosystems.

Leaf-cutter ants are dominant herbivores in ecosystems throughout the Neotropics that feed on fungal gardens cultivated on fresh foliar biomass. While the cycling of nutrients that takes place in these specialized microbial ecosystems is still not well understood, pan-omic investigations in collaboration with Dr. Cameron Currie (University of Wisconsin; Madison) have revealed that: the fungus, Leucoagaricus gongylophorus, plays a dominant role in breaking down cellulose, lignin, and pectin, while the bacterial community and the fungus break down simple oligosaccharides and hemicelluloses



Relative abundances of L. gongylophorus Carbohydrate-Active Enzymes (CAZyme) (blue), Lignin-degrading FOLyme (magenta), and Proteases (brown) protein families in the top (T), middle (M), and bottom (B) strata of Ac. echinatior and At. cephalotes gardens. regrading organisms ever cultured from the cow rumen, but the mechanism of cellulose degradation involving the direct attachment of the organism to the plant material has never been elucidated. Working with Dr. Garrett Suen (University of Wisconsin; Madison), pan-omics measurements have identified a novel mechanism for cellulose deconstruction that falls between cellulosome associated and fully secreted mechanisms.





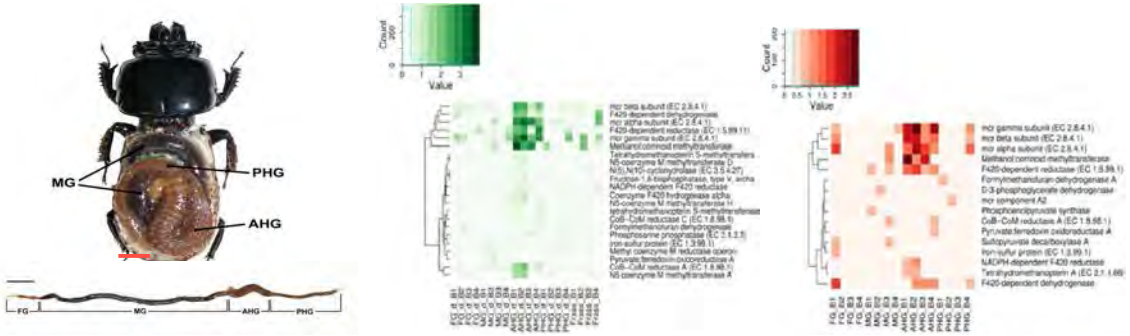
Cytoplasmic membrane Periplasmic space Outer membrane

Cellulose Fibers

Fibroslime protein
= CAZy protein
= Pilin structure= Cellulase

A proposed mechanism for cellulose deconstruction is being evaluated using pan-omics measurements, where plant fiber is mediated by fibroslime proteins, pilin like structures, and CAZy by way of ABC transporters into periplasmic space where cellulose is deconstructed through interaction with cellulases.

In phototrophic dominated microbial communities, carbon excretion by microbial mat primary producers, Cyanobacteria, is key to our understanding of carbon cycling but carbon excretion for extracellular matrix production is not well understood. In collaboration with Drs. Jennifer Pett- Ridge and Rhona Stuart (LLNL), pan-omic investigation of microbial mat extracellular matrix basic composition revealed a reservoir of starch-like (α-linked) oligosaccharides, cyanobacterial exoproteins and extracellular enzyme activity that are involved in breakdown of carbohydrates, Changes in extracellular protein abundance and enzyme activity of the Cyanobacteria in culture are diel and light dependent, and seem to be most active during the day.



Proteomics measurments determine that methanogenesis activity is localized in the anterior hind gut of the Passalid beetles.

Passalid beetles colonize decaying hardwoods, posses a high potential for lignocellulose degradation, and are a potential source of efficient xylose fermenting yeasts used for cellulosic ethanol production. Recent meta proteomic investigations into the microbial communities resident in the gut of the beetle, in collaboration with Dr. Eoin Brodie (LBNL) and Dr. Jennifer Pett-Ridge (LLNL) reveal nitrogenase and CAZY activities distributed through out the 4 regions of the gut, but methanogenesis activities localized to the anterior hind gut.

219. Pan-Omics: Activity-Based Proteomics

Aaron T. Wright1* (email: aaron.wright@pnnl.gov), Natalie C. Sadler1, Paul D. Piehowski1, Stephen J. Callister1, Richard D. Smith1

1Pacific Northwest National Laboratory, Biological Sciences Division, Richland, Washington, 99352

Project Goals: As part of the Pan-Omics Program we seek to enable a more complete assessment of the functional proteome of microbes and microbial communities by developing and deploying novel strategies in activity-based proteomics (ABP), which uses synthetic chemistry to create chemical probes that can measure enzyme functions, protein-substrate interactions, and protein redox regulation directly in living systems. Our research program has centered on increasing the technical capabilities of this technology by creating a multimodal probe suite for in situ studies, with accessibility to several analysis methods, and to correlate post-translational modifications to protein function, in conjunction with the use of advanced mass spectrometry-based platforms. Additionally, we are employing multiplexed probe studies to concomitantly address multiple functional proteins, or to analyze an active metabolic pathway more fully. Together, these ABP studies help link protein function, regulation, and interactions to the dynamic cellular physiology of microbial systems.

Activity-based proteomics addresses the fact that enzymatic function does not necessarily correlate with protein abundance, as cellular localization, allosteric and cofactor interactions, regulatory dynamics, and other details that mitigate function are generally not known. Activity-based probes can be used to monitor functional enzyme activity, redox regulation, and protein-substrate interactions on a proteome-wide basis. ABP not only provides a method to directly evaluate enzyme and substrate binding, and links between the proteome and metabolome, but when coupled to mass spectrometry- based top-down characterization methods can also provide quantitative information on specific proteoforms and PTMs that instigate activities (Figure 1). ABP can also provide information on the expression and functional status of proteins both in vitro and in living systems using chemical probes.

ABP involves the labeling of specific protein targets through the use of probes containing a click chemistry moiety (e.g., azide or alkyne), and effective enrichment of labeled proteins through click chemistry and affinity methods. ABP probes consist of three moieties: (i) a reactive group that forms an irreversible covalent bond with a target protein, (ii) a binding group that biases the probes toward a protein or protein family and may also impart cell permeability via active transport or diffusion, and (iii) a "clickable" reporter tag for rapid and sensitive measurement of labeled enzymes. Each of these three groups can be varied to create suites of probes.

When synthesizing new probes we attempt to keep the size small, thereby minimizing undesirable impacts on reactivity with the target proteins and maximizing cell permeability. Live cell or community labeling of active proteins enables the activity of a target to be monitored while maintaining the native cellular physiology, and prevents complications from the release of endogenous enzyme inhibitors following cell disruption. In addition to live studies we are increasingly multiplexing chemical probes to observe the complexity of myriad simultaneously functioning enzyme activities (Figure 2). Finally, since ABP captured proteins can be eluted at both the intact protein level or at the peptide level (i.e. following enzymatic digestion), both top-down and bottom-up MS-based approaches can be applied for a more comprehensive characterization of proteoforms (often revealing the specific sites of modifications that drive activity).

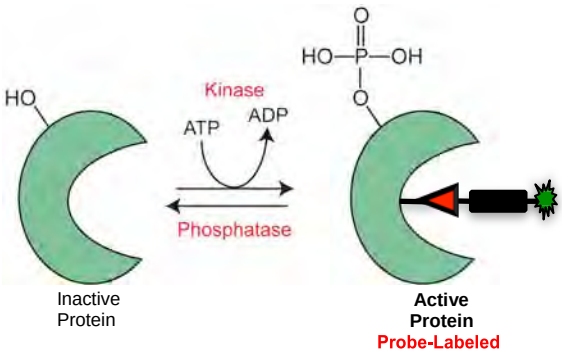


Figure 1. Intact ABP enables more effective characterization of active proteoforms.

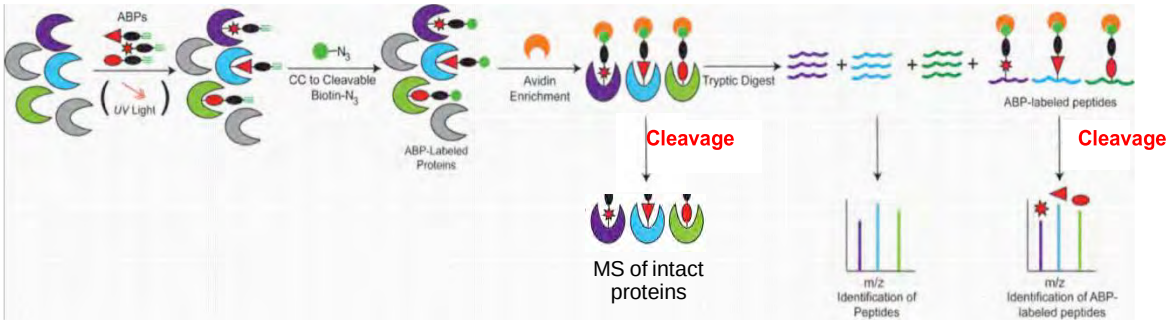


Figure 2. Multiplexed ABP approach for mass spectrometry-based intact and bottom-up proteoform characterization. A probe is added to a living cell or lysate, and only reacts with the active proteoform of the target enzyme. Click chemistry is used to attach biotin or other cleavable (e.g., by chemical reduction, light, or enzymatically) enrichment moiety, followed by enzyme capture and subsequent processing for MS analyses.

This research under the Pan-omics Program at PNNL was supported by the Genomic Science Program (GSP), Office of Biological and Environmental Research (OBER), U.S. Department of Energy (DOE), and is a contribution of the Pacific Northwest National Laboratory (PNNL) Foundational Scientific Focus Area. The work conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of DOE under Contract No. DE-AC02-05CH11231. A portion of this work was performed in the William R. Wiley Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by OBER and located at PNNL.

220. Pan-omics: Development of a Universal Next Generation Pan-omics Measurement Pipeline

Richard D. Smith*, Thomas O. Metz, Gordon A. Anderson, Yehia M. Ibrahim, Erin S. Baker, Ian K. Webb, Joshua N. Adkins, Carrie D. Nicora, Jennifer E. Kyle, Paul D. Piehowski, and Ronald J. Moore

Biological Sciences Division, Pacific Northwest National Laboratory, Richland WA 99352

http://www.omics.pnnl.gov

Project Goals: An aim of the Pan-omics Program at PNNL is to develop and apply advanced mass spectrometry (MS)-based platforms and bioinformatics approaches for 'pan-omics' measurements that will greatly accelerate the understanding of biological systems important to the U.S. Department of Energy and its Genomic Science Program. In this context, we are developing a new ion mobility separations (IMS) technology that can provide dramatically increased throughput and resolution to facilitate faster and better measurements. The approach uses novel Structures for Lossless Ion Manipulations (SLIM) to provide fast separations that exceed the resolving power of liquid chromatography and provide the basis for either large increases in throughput (e.g. 100-fold) or depth of 'ome coverage (i.e. completeness). The new SLIM IMS-MS platform will be applicable to all MS- based pan-omics measurements, can be interfaced to essentially any MS platform, and will provide a basis for future massively parallel MS arrays.

Our general approach for next generation pan-omics measurements includes a coherent universal pipeline involving the fractionation of biological samples prior to a new SLIM IMS-MS analysis platform (Figure 1). This pipeline uses a "one pot" approach to isolate proteomes, metabolomes, and lipidomes from the same parent sample, followed by fractionation of each sub-sample, e.g. using chromatographic separations. The degree of fractionation will depend on the desired depth of the measurements, as well as the capabilities of the new IMS-MS platform. Initially, we plan to prepare ~50 to 100 fractions for each sample, but larger numbers will be practical if deeper coverage is necessary. The analysis of each fraction will require only seconds, as compared to approximately one hour today. Because of the robust nature of these steps and their ready automation, sample processing needs for our initial applications will be readily addressable based on the use of conventional LC approaches and the developing SLIM IMS technology in conjunction with existing MS platforms. However, future advanced multiplexed and highly parallel MS platforms will also be enabled by these approaches and will provide the potential for measurement throughput that is even greater, and indeed comparable to that achieved with the fastest genomic technologies.

The new IMS-MS platform will employ a radically new approach to achieve ultra-fast mobility separations based upon SLIM. SLIM technology (Figure 2) has conceptual similarities with integrated electronic circuits, but instead of moving electrons, we use electric fields to create pathways, switches, etc for manipulating ions in the gas phase. SLIM enables sequences of ion separations, transfers, and trapping in the space between two surfaces positioned ~4 mm apart. We have developed key components that provide for lossless linear ion transport, ion transport around a corner (i.e. a 90-degree bend), ion switches (to direct ions to one of two paths), ion elevators (for transporting ions between different levels of multilevel SLIM devices), and ion traps (for trapping, accumulation, and reaction of ions). The capabilities are being initially used to enable long path length, high resolution IMS, as well as necessary ancillary capabilities such as extended ion trapping and accumulation necessary to achieve practical and high dynamic range MS-based measurements. The planned developments will facilitate understanding of biological systems and simultaneously enable more effective extraction of new knowledge and insights.
Figure 1. Overall Pan-omics pipeline being developed.

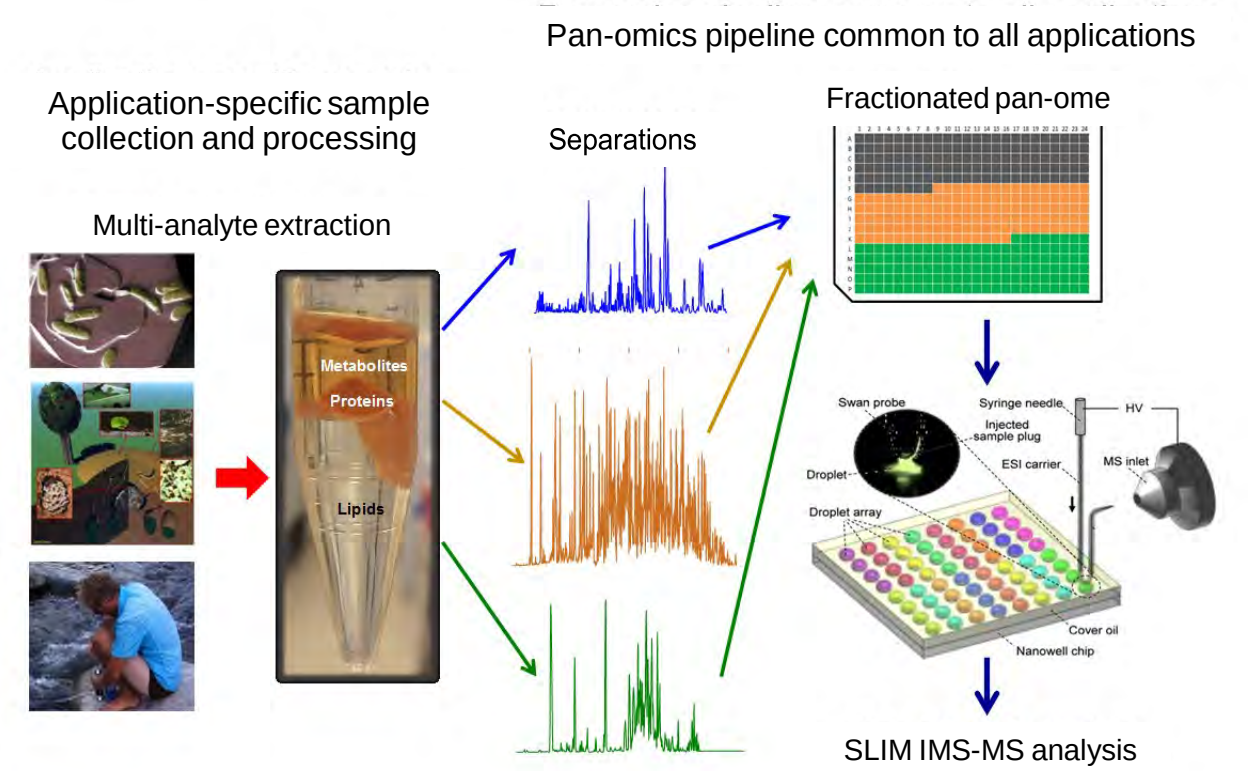
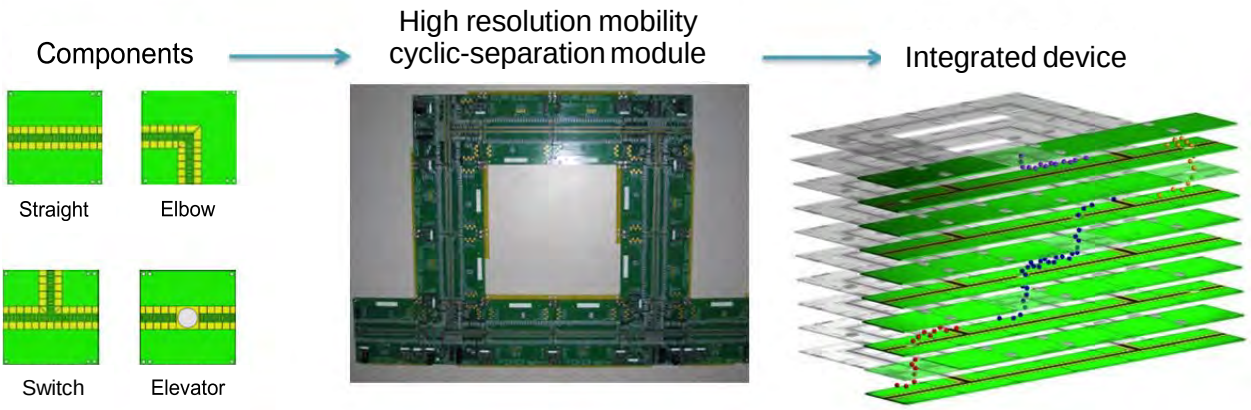


Figure 2. The new SLIM technology separations platform being developed is constructed using printed circuit board (PCB) technologies and involves the use of new approaches for fast gas phase ion mobility separations that replace on-line LC separations.



221. Pan-omics: Bioinformatics Developments

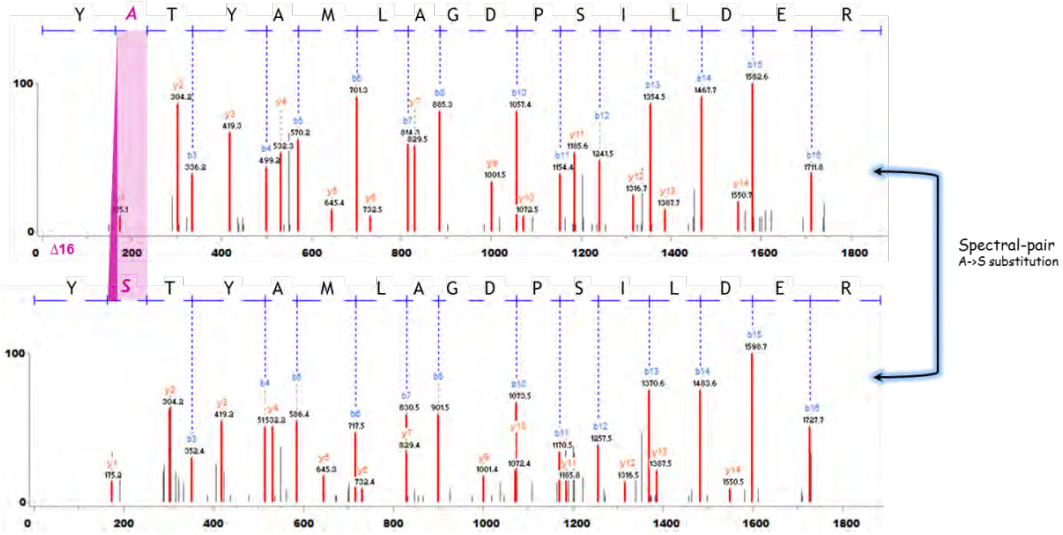
Samuel H Payne, Jason E McDermott, Nick Cramer, Joshua N Adkins* (joshua.adkins@pnnl.gov), Mary S Lipton, Richard D Smith

Biological Sciences Division, Pacific Northwest National Laboratory, Richland WA

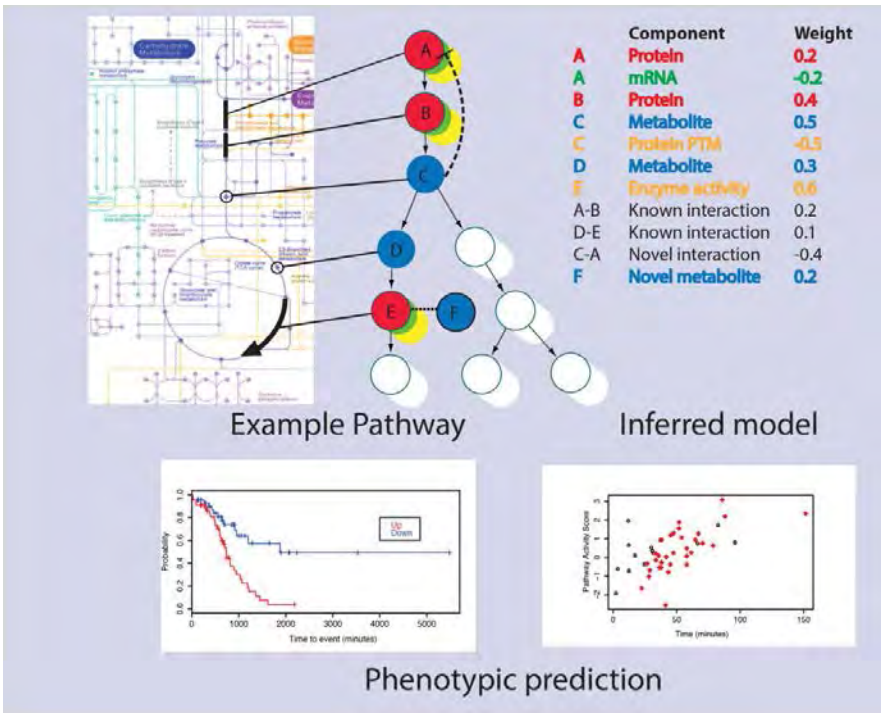
http://omics.pnl.gov

Project Goals: The Pan-omics Program scope includes creating rich multi-omic datasets for GSP relevant biological systems such as microbial communities and plants. The goal of the bioinformatics component within the Pan-omics Program is to build tools that can deal with increasing data throughput, and both speed and assist in data analyses and interpretation to help advance the knowledge about these environments/systems. We have recently focused on three critical areas: improved methods for proteomic characterization of microbial communities and other natural systems, integrating diverse pan-omic measurements, and advanced methods for collaborative and visual data analysis. These three efforts follow the natural flow of data as it is processed, analyzed and interpreted for biological conclusions.

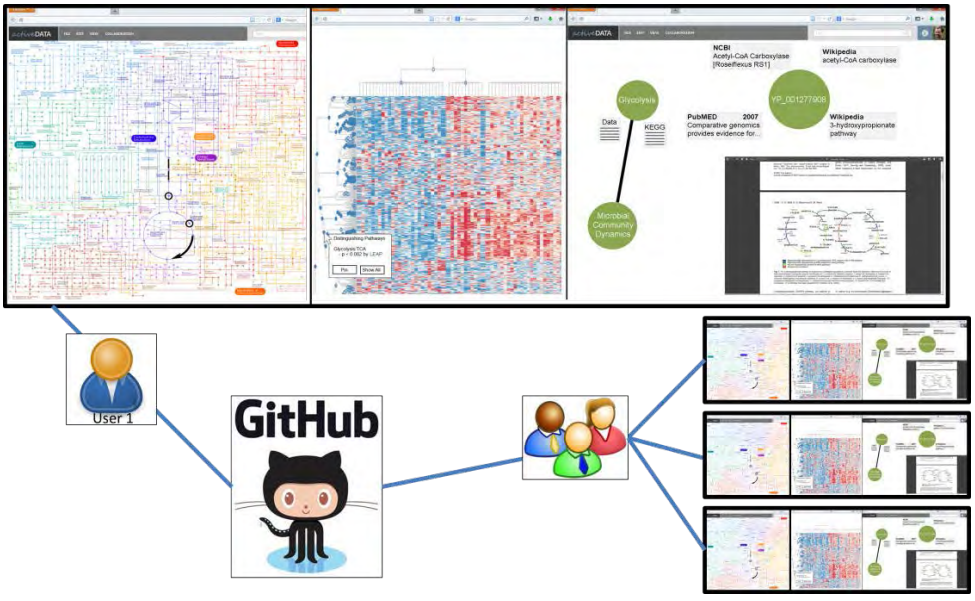
For mass spectrometry-based metaproteomic measurements where the corresponding metagenome information is unavailable, or where strain heterogeneity prevents complete resolution of the metagenome, traditional proteomics data analysis algorithms are insufficient. We are developing a new method, called Library Network Alignment that utilizes previously annotated spectra in a library to help identify the spectra from the metaproteomics sample. The core of this algorithm is the ability to recognize similar spectra from similar peptides and translate the sequence identification. The figure below is an example of spectral alignment between peptides with an S-> A mutation.



After quantifying proteins, metabolites, mRNA, etc. researchers need a way to coherently identify which biomolecules are correlated to their experimental inquiry. LEAP is a lasso-based statistical regression tool that identifies components of networks and pathways that are predictive of phenotype and experimental variables. LEAP integrates all pan-omic measurements as well as other high-throughput technologies into a single model, thus leveraging the unique information of each data-type.



Analysis of large and heterogeneous data requires multiple domain experts, and is a common pan-omics bottleneck. Moreover, the biological conclusions which are derived from data are most often discovered by non-computational scientists. Therefore, data analysis need to be both collaborative and visually accessible. We are developing the Active Data Canvas as a visual analytics tool that greatly facilitates collaboration and data browsing by providing on-the-fly statistical analyses, integrated and interactive data visualization, and a sharable thought-space for annotating significant data or conclusions.



222. Semantic Index of Phenotypic and Genotypic Data

Charles T. Parker1* (chuck.t.parker@namesforlife.com), Nenad Krdzavac,2 Kevin Petersen,2 Grace Rodriguez,1 and George M. Garrity1,2

1NamesforLife, LLC; East Lansing, Michigan and 2Michigan State University; East Lansing, Michigan

http://www.ontology.namesforlife.com

Project Goals: The goal of this project is to develop a semantic data resource that can serve as a basis for predictive modeling of microbial phenotype. The core technical objectives are twofold: (1) to build a database of normalized phenotypic descriptions (observational data) using the primary taxonomic literature of bacterial and archaeal type strains, and (2) to construct an ontology with reasoning capabilities to make accurate phenotypic and environmental inferences based on that data. This project is tightly coupled with ongoing DOE projects (the Genomic Encyclopedia of Bacteria and Archaea, the Microbial Earth Project, the Community Science Project) and with two key publications, Standards in Genomic Sciences (SIGS) and the International Journal of Systematic and Evolutionary Microbiology (IJSEM).

The DOE Systems Biology Knowledgebase (KBase) was envisioned to provide a framework for modeling dynamic cellular processes of microorganisms, plants and metacommunities. The KBase will enable rapid iteration of experiments that draw on a wide variety of data and allow researchers to infer how cells and communities respond to natural or induced perturbations, and ultimately to predict outcomes.

Predictive models rely on high quality input data, but not all data are of similar quality nor are they amenable to computational analysis without extensive cleaning, interpretation and normalization. Key among those needed to make the KBase fully operational are phenotypic data, which are more complex than sequence data, occur in a wide variety of forms, often use complex and non-uniform descriptors and are scattered about specialized databases and scientific and technical literature. Incorporating phenotypic information into the KBase requires expertise in harvesting, modeling and interpreting these data. The Semantic Index of Phenotypic and Genotypic Data will address this problem by providing a resource of reference phenotypic data for all validly published type strains of Bacteria and Archaea, based on concepts and observational data drawn from the primary taxonomic literature. In the Phase I project we developed software to construct and analyze a corpus of this literature and to extract putative feature domain vocabularies comprising approximately 40,000 candidate phenotypic terms used in new and emended descriptions of the 12,937 distinct type strains of Bacteria and Archaea. In Phase II, these vocabularies have served as the basis for developing a phenotypic ontology, a repository of phenotypic data that is undergoing normalization of phenotypic descriptions for each species. We have found that many of the phenotypes applied to microbes describe a combination of quantitative environmental conditions and qualitative growth and metabolic capabilities. Such terms have proven difficult to implement in query systems because of their context-based interpretations and conceptual overlap across multiple feature domains. We have furthered our work on novel design patterns for ontology development [1] that address these problems and remove barriers to machine reasoning over these complex terms, while preserving the bi-directional mapping back to human interpretation at multiple levels of abstraction. A PCT patent application on this method was filed in Q3 2014 and the preliminary examination has found all claims to be novel, non-obvious and industrially applicable. Critical to implementing this system has been the adoption of a SPARQL Inferencing Notation reasoner coupled with a triplestore rule-based reasoner. This approach resolves ambiguity attributed to the semantic equivalence and imprecision of

phenotypic terms arising in literature and in databases.

In order to better facilitate access to knowledge extracted from the literature and encoded in the ontology, we are implementing a special-purpose web portal to accommodate query and retrieval of biological resources by term or concept, with a multi-tier query platform conforming to current search standards and backed by Semantic Web and Linked Data query standards. In addition to linking to the primary literature, related ontologies and source data, we are also incorporating public data from NCBI, USDA and the Joint Genome Institute in order to provide researchers with immediate access to the appropriate resources for a set of strains, along with consistent, accurate interpretations of available knowledge about those strains that are usable for predictive modeling and in other research and commercial applications. As part of our commercialization activities, we continue to develop several software components that resulted from this project into a commercial semantic search and document analysis platform with end products being used in Standards in Genomic Sciences and the International Journal of Systematic and Evolutionary Microbiology.

Publications

Parker, CT, Garrity, GM and Krdzavac, NB. Systems and Methods for Establishing Semantic Equivalence Between Concepts. U.S. Patent Application No. PCT/US2014/056808. Filed September 22, 2014. Washington, DC: U.S. Patent and Trademark Office.

Funding for this project was provided through the DOE SBIR/STTR program (DE-SC0006191). Public funding for development of the NamesforLife infrastructure was received from the DOE SBIR/STTR program (DE-FG02-07ER86321), the Michigan Small Business Technology Development Corporation, the Michigan Strategic Fund, and the Michigan Universities Commercialization Initiative.

223. Exploiting Plant Growth-Promoting Root Endophytes to Improve Feedstock Nitrogen Acquisition

Meghan E. Feltcher1, Natalie Breakfield2, Hunter Cameron1, Sur Herrera Paredes1, Surge Biswas1, Amanda Panousis1, and Jeffery L. Dangl1

1University of North Carolina, Chapel Hill

http://labs.bio.unc.edu/dangl/projects/Functional_endophyte_manip.htm

Project Goals: It is increasingly clear that sustainability of plant feedstocks will require novel approaches to maintaining plant fitness in diverse soil environments with fluctuating nutrient and moisture levels. Plant health is co-dependent on the microbial communities intimately associated with roots (the root microbiome), which can positively modulate plant physiology by increasing bioavailable nutrients, stimulating root growth through phytohormone signaling, and priming the plant innate immune system to enhance pathogen resistance. A major goal of the Dangl lab is to define members of the root microbiome that are functionally relevant to plant health and productivity. Furthermore, we want to determine how these beneficial bacteria compete amongst a complex soil microbial community to occupy the plant root niche and confer plant growth-promoting activity. Because nitrogen is the limiting element of net primary production in most terrestrial ecosystems, we are interested in determining how root-associated bacteria can stimulate plant growth under limited nitrogen conditions. More specifically we want to 1) identify endophytic bacteria that stimulate plant growth under nitrogen stress and define the mechanisms by which plant biomass is increased, 2) determine the bacterial traits required for root endophytic compartment colonization and whether occupying this niche increases the success of plant-growth promoting bacteria in community settings, and finally, 3) define synthetic bacterial communities that confer robust plant resilience to nitrogen stress.

Plants employ several strategies to maximize nitrogen uptake amidst fluctuating soil nitrogen levels, such as root architecture modulations, regulation of nitrogen assimilation pathways, and maintenance of a root microbiota that improves root nitrogen acquisition. Exploiting this root microbiome in order to improve plant fitness is an attractive strategy for generating fertile crops while reducing dependence on chemical fertilizers. Specifically, there has been interest in isolating plant growth-promoting bacteria (PGPB) that improve plant nitrogen acquisition. Unfortunately, efforts to inoculate field crops with PGPB often fail, likely due to competition with native soil microbial communities and limited root endophytic compartment (EC) colonization efficiency. Thus, there is a demand to study EC colonization by PGPB both in the context of the root microbiome community and changing nitrogen conditions. Our group, in collaboration with the Department of Energy Joint Genome Institute (DOE-JGI), completed a bacterial census of the EC community of A. thaliana grown in two geographically distinct soils using 16S ribotyping. An important finding from this study was that the microbial composition of the root EC was both taxonomically distinct from and less diverse than that of bulk soil, suggesting there are plant- and/or bacterial- derived factors governing root EC community assemblage. To begin to understand what members of the root microbiome (specifically, endophytes) contribute to plant fitness under nutrient stresses and pathogen challenges, we isolated a diverse collection of putative endophytes from A. thaliana roots grown in the same two natural soils. This collection of nearly 600 isolates provides the basis for studying the root microbiome in a more reductionist laboratory setting, allowing us to develop screens for various plant growth-promoting activities and to construct synthetic soil and root communities that can be manipulated and studied under various nutrient stresses.

We hypothesize that some of these putative EC colonizers stimulate plant growth under nitrogen stress;

therefore, we are screening this diverse collection of putative root endophytes in A. thaliana mono-association experiments for both EC colonization and plant growth-promotion activity. Using a gnotobiotic vertical growth method on agar plates with varying nitrogen concentrations, we are able to monitor both bacterial-induced changes to rosette size as well as root architecture. Screening of over 60 isolates revealed that bacteria from diverse families stimulate rosette growth under sufficient nitrogen conditions such as Microbacteriaceae, Nocardiaceae, Burkholderiaceae, and Bacillaceae. However, many of these growth-promoting strains resulted in reduced A. thaliana shoot biomass under nitrogen stress. Those that increased rosette size under nitrogen deprivation include isolates from the Pseudomonadaceae and Comamonadaceae families. These mono-association studies have also revealed that many isolates robustly colonize the A. thaliana EC. To understand the bacterial genes required for EC colonization of diverse PGPB identified in these screens, we are using transposon mutagenesis coupled with high-throughput sequencing (TnSeq). Additionally, we are comparing PGPB to other taxonomically-related isolates in our collection (defined by 16S) in order to identify phenotypic diversity that could be exploited in whole genome sequence comparisons to look for the genetic determinants of growth promotion.

We have also constructed diverse synthetic communities of up to 53 isolates to determine how individual plant- growth promoting bacteria perform amongst a community, and to begin to identify root community structures that maximize plant growth. We demonstrated that re-colonization of A. thaliana roots is robust in our model community system and some synthetic communities can promote A. thaliana growth. Another important finding is that some isolates that colonize the EC robustly in mono-association experiments do not associate with the root in synthetic community experiments. This suggests they are outcompeted by other members of the community. Interestingly, several isolates that stunted A. thaliana growth in mono-association were unable to do so in the context of a community. This result suggests that other isolates either directly dampen bacterial activities detrimental to plant growth or stimulate the plant in a protective way. Current work is aimed at using mono- association data to construct synthetic communities in order to determine if individual PGPB activities are additive and transferable to more complex model systems (such as plants grown in pots with synthetic or natural soils). Also, to elucidate the mechanisms by which PGPB are stimulating plant biomass under nitrogen deprivation we are measuring root architectural changes associated with increased rosette size, and will use RNAseq analysis to determine how the plant nitrogen stress response is changed upon addition of PGPB. Results from these studies will further our understanding of the dynamics of natural root microbiome structures that directly impact the health of field crops, therefore aiding in the development of novel PGPB-based strategies that can compete in natural microbial soil communities to increase crop performance in an environmental friendly matter.

This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0010423 to JLD. JLD is an Investigator of the Howard Hughes Medical Institute, supported by the HHMI and the Gordon and Betty Moore Foundation (GBMF3030). MEF is a Ruth L. Kirschstein fellow, grant no. 1F32GM112345-01.

224. Linking genotype and phenotype in a model grass: Building tools for genome-wide- association studies in Brachypodium

Ludmila Tyler1, Scott J. Lee2,3, Nelson D. Young2, Gregory A. DeIulio1,3, Elena Benavente4, Michael Reagon5, Jessica Sysopha2, Riccardo M. Baldini6, Angelo Troìa7, Samuel P. Hazen2,* (hazen@bio.umass.edu), and Ana L. Caicedo2

1 Biochemistry and Molecular Biology Department, University of Massachusetts, Amherst, Massachusetts, USA. 2 Biology Department, University of Massachusetts, Amherst, Massachusetts, United States of America. 3 Plant Biology Graduate Program, University of Massachusetts, Amherst, Massachusetts, United States of America. 4 Department of Biotechnology, ETSIA, Technical University of Madrid, Madrid, Spain. 5 Department of Biology, The Ohio State University Lima, Lima, Ohio, United States of America. 6 Department of Biology, University of Florence, Florence, Italy. 7 Dipartimento STEBICEF, Sezione di Botanica ed Ecologia vegetale, Università di Palermo, Palermo, Italy

Project Goals: (1) Phenotype a germplasm collection of geographically diverse B. distachyon individuals to determine overall existing diversity for biofuel conversion quality. (2) Assess population structure and genetic diversity in our population using genotype-by-sequencing. (3) Determine the feasibility of using association mapping techniques to find genes of importance in B. distachyon, by assessing genome-wide linkage disequilibrium.

The small, annual grass Brachypodium distachyon, a close relative of wheat and barley, is a powerful model system for cereals and bioenergy grasses. Genome-wide association studies (GWAS) of natural variation can elucidate the genetic basis of complex traits, but have been largely prevented in B. distachyon by the lack of sufficient numbers of well- characterized accessions. Here, we report on genotyping-by-sequencing of 84 B. distachyon, eight B. hybridum, and three B. stacei accessions with diverse geographic origins, including Albania, Armenia, Georgia, Italy, Spain, and Turkey. Over 90,000 high- quality single-nucleotide polymorphisms (SNPs) distributed across the Bd21 reference genome were identified. Our results confirm the hybrid nature of the B. hybridum genome, which appears as a mosaic of B. distachyon-like and B. stacei-like sequences. Analysis of more than 50,000 SNPs for the B. distachyon accessions has revealed three distinct, genetically defined populations. Surprisingly, these genomic profiles are more readily explained by differences in flowering time than by geographic origin. High levels of differentiation in loci associated with floral development support the differences in flowering time between B. distachyon populations. GWAS investigations combining genotypic and phenotypic data also suggest the presence of one or more photoperiodism, circadian clock, and vernalization genes in loci associated with flowering time variation within B. distachyon populations. Our characterization expands the germplasm resources available for Brachypodium and demonstrates the feasibility of GWAS in this model grass.

This work was supported by the Office of Science, Office of Biological and Environmental Research of the U.S. Department of Energy grant DE-SC0006621.

225. Engineering Anaerobic Gut Fungi for Lignocellulose Breakdown

Kevin V. Solomon (ksolomon@engineering.ucsb.edu)*1, John K. Henske1, Charles H. Haitjema1, Diego Borges-Rivera2, Dawn A. Thompson2, Michelle A. O'Malley1

1 Department of Chemical Engineering, University of California-Santa Barbara, Santa Barbara, CA 93106-5080; 2 Broad Institute of MIT and Harvard, Cambridge, MA 02124

http://omalleylab.weebly.com

Project Goals: The goal of this project is to engineer anaerobic gut fungi as novel platform organisms for biofuel production from plant material. To accomplish this goal, a panel of anaerobic fungi will be isolated from different herbivores and screened for their ability to degrade lignocellulose. The basic metabolic networks that govern lignocellulose hydrolysis within anaerobic fungi will also be determined, and models will be generated to describe how important enzyme groups are coordinated during breakdown. Using this information, genetic transformation strategies to manipulate gut fungi will be developed, which would endow them with enhanced functionality against a range of industrially relevant substrates. Collectively, this information will establish the molecular framework for anaerobic fungal hydrolysis, and will guide in the development of lignocellulosic biofuels.

Anaerobic fungi are the primary colonizers of biomass within the digestive tract of large herbivores, where they have evolved unique abilities to break down lignin-rich cellulosic biomass through invasive, filamentous growth and the secretion of powerful lignocellulolytic enzymes and enzyme complexes (fungal cellulosomes). Despite these attractive abilities, considerably less genomic and metabolic data exists for gut fungi compared to well-studied anaerobic bacteria and aerobic fungi that hydrolyze cellulose. This presents a significant knowledge gap in understanding gut fungal function, substrate utilization, and metabolic flux, which has prohibited the genetic and functional modification of gut fungi. Recently, however, advances in sequencing technologies have made it possible to explore the dynamic metabolic networks within gut fungi for the first time. Our approach combines next-generation sequencing with physiological characterization to establish the critical knowledge base to understand lignocellulose breakdown by gut fungi. This project will (1) enable exploration of novel isolates for desirable enzymatic properties, (2) construct metabolic models to describe biomass degradation, and (3) develop new methods to metabolically engineer gut fungi for bioprocessing.

We isolated a panel of novel gut fungi from sheep, goat, giraffe, and elephants at the Santa Barbara Zoo. To date, four unique strains from the Piromyces, Neocallimastix, Anaeromyces, and Caecomyces genera have been obtained through roll tube isolation. Proliferation of the fungal isolates was monitored via fermentation gas production, and cellulosomes from each species were isolated through cellulose-precipitation. All of these isolates exhibited high enzymatic reactivity against a range of cellulosic and unpretreated substrates (reed canary grass, switch grass, beechwood xylan), which was repressed in the presence of simple sugars. Within isolated cellulosomes across multiple genera, striking similarities are observed for certain dockerin-fused glycosyl hydrolases, and these proteins are not secreted from fungi when simple sugars are present, supporting the hypothesis that these enzymes are catabolically regulated. We have collaborated with researchers at the DOE-JGI and PNNL EMSL through a 2014 Community Science Program to sequence the genomes/transcriptomes for these isolates, which reveals conserved xylan-degrading machinery and putative cellulosome structure.

Though fungal hydrolytic activity has been shown to be substrate dependent, the underlying regulation mechanisms that coordinate the action of cellulases and cellulosomes from gut fungi remain unknown. We hypothesize that cellulose-degrading machinery is catabolite-repressed to conserve cellular energy,

and our objective is to exploit this regulation mechanism to discover novel enzymes. To address this hypothesis, we have combined next-generation sequencing and proteomic approaches to examine cellulose-degrading enzyme production in a panel of gut fungi isolated from natural ecosystems. Through strand-specific RNAseq and use of the TRINITY assembly platform, we were able to assemble thousands of novel genes de novo from >27,000 transcripts without the need for genomic sequence information. RNA-Seq also elucidated global regulatory patterns in response to catabolite repression of biomass degradation, and in response to growth on cellulosic substrates of increasing complexity. Through these efforts, we have identified hundreds of transcripts encoding novel enzymes for biomass degradation, and the fungal transcriptome is particularly rich in GH6, GH48, and GH43-containing enzyme domains. As hypothesized, most of these transcripts are strongly repressed by the addition of simple sugars and are clustered within distinct 'regulons' of coordinated gene expression. Gene set enrichment analysis confirms the upregulation of these regulons across cellulosic substrates, and complementary proteomic analyses reveals that the composition the fungal cellulosomes is tuned by the presence of different substrates. More importantly, the functional enrichment of these regulons suggests a critical role for the divergent and unannotated transcripts that they contain. A dozen co-regulated transcripts from our screen bear no homology to known enzymes, and likely harbor previously undiscovered glycosyl hydrolase domains from nature, which we are currently investigating through protein crystallography in collaboration with Argonne National Laboratory, as well as high-throughput activity screening at NREL. Collectively, this information will establish the molecular framework for anaerobic fungal hydrolysis, ultimately allowing us to refactor this system in anaerobic fungi and beyond.

This Project is supported by the Office of Biological and Environmental Research through the DOE Office of Science

226. Engineering Synthetic Anaerobic Consortia for Sustainable Chemical Production from Biomass

John K. Henske (john.henske@gmail.com)*1, Kevin V. Solomon1, Charles H. Haitjema1, Sean P. Gilmore1, Michelle A. O'Malley1

1 Department of Chemical Engineering, University of California-Santa Barbara, Santa Barbara, CA 93106-5080

http://omalleylab.weebly.com

Project Goals: The goal of this project is to engineer anaerobic gut fungi as novel platform organisms for biofuel production from plant material. To accomplish this goal, a panel of anaerobic fungi will be isolated from different herbivores and screened for their ability to degrade lignocellulose. The basic metabolic networks that govern lignocellulose hydrolysis within anaerobic fungi will also be determined, and models will be generated to describe how important enzyme groups are coordinated during breakdown. Using this information, genetic transformation strategies to manipulate gut fungi will be developed, which would endow them with enhanced functionality against a range of industrially relevant substrates. Collectively, this information will establish the molecular framework for anaerobic fungal hydrolysis, and will guide in the development of lignocellulosic biofuels.

Anaerobic fungi, native to the digestive tracts of large herbivores, are capable of breaking down lignin-rich biomass. This is accomplished through invasive growth of mycelial root structures and the secretion of lignocellulolytic enzymes and complexes. Thus far, very little is known about these fungi on a genomic and metabolic level compared to cellulolytic bacteria and aerobic fungi. We have used next generation sequencing to identify biomass degrading enzymes through sequence homology and orthology analyses as well as through regulation patterns during growth on substrates of varying complexity. The enzymes contained within anaerobic fungi make them a valuable resource to improve bioprocessing methods. Based upon the ability of these fungi to degrade a variety of materials ranging in complexity, we have examined several co-culture systems to determine how these abilities may be enhanced and how they may be used for bio- based production.

Natively, gut fungi maintain a syntrophic relationship with archaeal methanogens. Fungi degrade the plant material and produce both CO2 and hydrogen that is then converted to methane gas by the methanogens. We have isolated cultures of anaerobic fungi with their dependent methanogens from herbivore fecal materials. These co-cultures establish a simplified system to model cooperative actions of the anaerobes and are stable in culture for > 60 weeks. (e.g. Methanocorpusculum) during biomass breakdown shows that fungal-methanogen syntrophy drastically accelerates cellulose and lignocellulose breakdown by the fungi. Furthermore, an excess of sugar hydrolysates and metabolites produced by anaerobic fungi enable metabolic linkage of non-native, facultative anaerobes to the consortium. Strains of S. cerevisiae and E. coli have been tested for their use in synthetic consortia with gut fungi, by which they consume the excess sugars (~5 g/L) left over from fungal cellulose breakdown. As anaerobic fungi are not yet genetically tractable, our strategy offers a path forward to make value-added products directly from crude lignocellulose.

227. Studying lipid accumulation mechanism in oleaginous yeast using hyperspectral SRS microscopy

Dan Fu1*(danfu@fas.harvard.edu), Shiyou Ding,2 and X. Sunney Xie1

1Harvard University, Cambridge, MA; and 2Michigan State University, East Lansing

Project Goals: Our research will combine advanced imaging, single-cell systems biology, and metabolic engineering approaches to understand the lipid accumulation mechanism of oleaginous yeast, and to decouple nitrogen regulation and sugar utilization on lipid production for lignocellulosic advanced biofuel production. The proposed research will focus on the analysis of transcriptomic data of single cells that will be correlated with the quantitative measurement of lipid production in vivo. Through genetic association study we will identify genes and transcriptional factors that correlate with lipid accumulation, and further verify these potential candidate genes or genome regions that contribute to microbial lipid accumulation by metabolic engineering.

The Xie group at Harvard University develops advanced imaging and sequencing techniques to quantitatively characterize the genome, transcriptome, and metabolome of single cells. Using these novel tools, we hope to understand lipid accumulation mechanism of oleaginous yeasts and identify yeast mutants that are capable of high lipid production using biomass hydrolysates. Although the mechanism of lipid production has been explored with genome and multi-omic analysis and documented recently in certain yeast species such as R. toruloides, the exact regulatory network governing lipid accumulation to maximize the microbial lipid production has not yet been completely unraveled. In addition, it has been reported that the content, morphology (i.e., number and size), and chemical composition of lipid droplets produced by oleaginous yeast vary in different strains and under different growth conditions. To understand the lipid production mechanism and identify genetic features responsive to lipid accumulation in the presence of pentose and nitrogen, we propose to develop an automated chemical imaging and single cell transcriptomics method to correlate the lipid accumulation with the transcriptional profiles at the single-cell level.

Our first goal is to develop a high throughput single cell imaging technique that can visualize and quantify lipid production of yeasts under a wide variety of culture conditions and with different sugar sources. Our group has pioneered stimulated Raman scattering (SRS) microscopy, a technique that visualizes molecules based on their intrinsic vibrational contrasts without any need for labels. It offers exciting new opportunities in studying lipid production in vivo with high spatial and temporal resolution based on C-H vibrational signature from lipids. The number and size of lipid droplets within yeast cells can be easily quantified. However, it has limited chemical resolvability due to the strong spectral overlap among many biomolecular species in the C-H region. We tackled this challenge with two technical innovations: 1) hyperspectral SRS imaging uses spectral information to distinguish closely related lipid species; 2) deuterium tracing SRS exploits deuterium labeled metabolic substrates to quantify their conversion to cellular components. We have demonstrated the capability of these novel imaging approaches in both single cell organisms such as yeasts as well as multicellular organisms such as C. elegans. Triglycerides and cholesteryl ester, two major neutral lipids, can be distinguished based on their spectral difference in the C-H region. Uptake of different fatty acids into individual lipid droplets in living cells can be quantified with deuterium tracing. We further demonstrated that lipogenesis can be monitored at subcellular resolution by using deuterated glucose as the metabolic substrate. These unprecedented capabilities open up new ways to understand lipid accumulation in oleaginous yeast. We will build a flow cytometry platform that uses SRS signal from lipids to screen for single cells with high lipid production. In combination with single cell sequencing, we expect to unravel the genetic elements that regulate yeast

lipid accumulation and engineer mutants with robust and high yield lipid production.

XSX acknowledges the U.S. Department of Energy's BER Genomic Science Program (DE-SC0012411) for supporting this research.

228. Isotopic and click chemistry-based approaches for tracking activity in single cells and newly synthesized proteins in uncultured microorganisms catalyzing anaerobic methane oxidation

Victoria J. Orphan1*, Roland Hatzenpichler1, Jeffrey Marlow1 Grayson Chadwick1, Shawn E McGlynn1, Chris Kempes2, Connor Skennerton1, Roland Hatzenpichler1, Hang Yu1, Karuna Chourey3, Chongle Pan3, Robert L Hettich3, Christof Miele4, Yimeng Shi4 (vorphan@gps.caltech.edu)

1 California Institute of Technology, Pasadena, CA 2 NASA Ames Research Center, Mountain View, CA 3 Oakridge National Research Lab, Oakridge, TN 4University of Georgia, Athens, GA

www.gps.caltech.edu/people/vorphan/profile

Project Goals: Quantifying the activity of uncultured microbial cells within the environment has been a long standing challenge, but represents a key element in understandinging the function, dynamics and resiliency of microbial communities and consortia. Here we describe two complementary methodologies for measuring cellular activity (protein synthesis) by environmental microorganisms using 15N stable isotope probing (SIP) and bioorthogonal noncanonical amino acid tagging, or BONCAT, a click-chemistry method which labels newly synthesized proteins after assimilation of a methionine analog (Hatzenpichler et al., 2014). These methods were coupled with fluorescence and ion imaging (nanoSIMS) techniques for determination of the activity of individual cells within structured microbial consortia mediating the anaerobic oxidation of methane (AOM). At the level of the proteome, 15N-SIP and BONCAT are being coupled with environmental proteomics, allowing for the identification of newly synthesized microbial proteins in environmental samples incubated under different geochemical conditions.

Combinations of BONCAT and 15N-SIP being used to develop a more comprehensive understanding of the potential mechanisms underpinning the anaerobic oxidation of methane (AOM)- a globally important sink for methane. In methane-saturated anoxic sediments, uncultured methane-oxidizing 'ANME' archaea are frequently observed in structured consortia with sulfate-reducing deltaproteobacteria. This physical association has been historically interpreted as evidence for syntrophically-mediated AOM coupled with sulfate-reduction, but the specific mechanism(s) has remained elusive. Through activity-based tracking of methane-oxidizing archaea and associated bacteria, we are gaining new insights into the interspecies interactions occurring within these consortia and the specific pathways used for electron/ metabolite exchange during AOM. Protein- SIP analyses revealed information on the ecophysiology of slow growing methane- oxidizing consortia and specific proteins that are actively synthesized during AOM. Several ortholog proteins believed to be involved in the "reverse methanogenesis" and sulfate reduction pathways were detected, with roughly 1/3 of the total detected protein pool (704 proteins) showing evidence of 15N enrichment. A direct comparison of patterns in 15N labeling of proteins with the newly synthesized protein fraction detected by BONCAT is underway. FISH-nanoSIMS quantification of 15N assimilation by individual archaea and sulfate-reducing bacteria in AOM consortia revealed spatial patterns in activity that are inconsistent with conventional syntrophy based on molecular diffusion and instead appear to support an alternative interspecies electron transfer mechanism, likely facilitated by extracellular multiheme cytochromes. BONCAT experiments were used to track de novo protein synthesis in AOM consortia from sulfidic sediments and CH4 bioreactors over time, revealing enhanced protein synthesis by ANME-SRB consortia in the presence of CH4. The application of fluorescence-based cell sorting of individual, translationally active and FISH-stained consortia from sediments for the generation of targeted ANME consortia metagenomes is in progress, and data from this genome sequencing effort will be valuable for further refinement of our proteomics reference database and will facilitate comparative

genomics studies of the distinct methane-oxidizing archaeal lineages.

Supported by the Office of Biological and Environmental Research

229. The Genetics of Divergence Between Upland and Lowland Ecotypes of Panicum virgatum

Elizabeth Milano1, David Lowry2, and Thomas Juenger1*(tjuenger@austin.utexas.edu)

1University of Texas at Austin, Department of Integrative Biology, Austin, Texas; 2Michigan State University, Department of Plant Biology, East Lansing, Michigan.

https://w3.biosci.utexas.edu/juenger_lab/?q=switchgrass-overview

Project Goals: The overall goal of the proposed research is to utilize genetic and genomic analyses to better understand the growth and development of Panicum grasses. This project will characterize the phenotypic diversity in Panicum virgatum and Panicum hallii accessions to better understand growth architecture and developmental characteristics using QTL mapping, gene expression studies, and experimental manipulations.

Abstract: Panicum virgatum (switchgrass) is an outcrossing C4 perennial grass native to most of North America. Wild populations are predominantly found in tall grass prairie and agronomic cultivars are being developed as a feedstock for cellulosic bioethanol. Because of its economic importance, there is interest in developing genomic tools for marker-assisted breeding programs and crop development. However there is abundant natural variation in Panicum virgatum that lends itself well to the study local adaptation. Two predominant ecotypes are found in overlapping ranges across the Midwestern United States. The northern upland ecotype is adapted to colder climates and shorter growing seasons, whereas the lowland ecotype is found in riparian areas and has a longer growing season. We have developed a 4-way outbred mapping population of 400 individuals that captures the variation observed between upland and lowland ecotypes to study the underlying genetic architecture of locally adaptive traits. Genetic markers were developed using a double-digest RAD scheme and assembled using both a reference-based and de novo approach. The population is currently growing in Austin, Texas but will be clonally propagated and planted at several sites spanning a latitudinal gradient across the US. Thus far we have created a linkage map and found many QTL associated with traits measured in the initial field season. In future years we will compare QTL across environments to discover loci underlying trait divergence between locally adapted ecotypes.

Funding Statement: This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0008451.

230. Biofuels in the (arid) West: germplasm development for sustainable production of Camelina oilseed

John K. McKay, Colorado State University (PI), Edgar Cahoon, University of Nebraska – Lincoln (Co-PI) Luca Comai, University of California, Davis (Co-PI) Timothy Durrett, Kansas State University (Co-PI), Jack L. Mullen, Colorado State University (Co-PI)

The emerging crop Camelina sativa has the potential to become an important biofuel oilseed crop for the extensive dryland farming regions of the West. Camelina has a number of properties that make it an already adequate oilseed-based biofuel feedstock for the Great Plains and Western US, including low input requirements. Although the species has received almost no modern breeding, the oil content of Camelina seeds is already nearly 40%, which is twice that of soybean. We focus on development of Camelina as a crop to be grown on marginal farmland with relatively low fertilizer inputs. Despite its numerous positive attributes recommending it as a biofuel feedstock for the Great Plains and Western US, needs for improvement of agronomic and oil quality traits remain. To maximize productivity and sustainability in the face of climate change, Camelina requires improved drought and heat tolerance to fit within dryland and limited irrigation agriculture. The Camelina genome has just been sequenced and assembled, providing the framework for targeted genetic manipulation. Here we seek to leverage this newly available genome, using forward and reverse genetics and natural variation to combine optimal qualities for an oilseed feedstock in the Great Plains and Western US.

Research Objectives: 1) Develop a TILLING population and utilize it to identify lines with preferred oil traits and improved heat tolerance. This resource will also be available for other important traits like germination and herbicide tolerance. 2) To develop germplasm of Camelina, by discovery and selection of loci affecting drought and heat tolerance and oil properties. 3) Leverage the detailed knowledge of lipid biosynthetic pathways in Arabidopsis and target these genes in Camelina. This can largely be achieved though mutagenesis. In addition, Camelina is easily transformed and thus there is an opportunity to radically improve and optimize the oil profile for use in modern diesel engines.

231. Genetic Architecture Of Sorghum Biomass Yield Component Traits Identified Using High-Throughput, Field-Based Phenotyping Technologies

Patrick S. Schnable* (schnable@iastate.edu), Maria G. Salas Fernandez, and Lie Tang

Iowa State University, Ames

Project Goals: A systems approach (Genome---wide Association Studies; GWAS) will be used to identify the genetic control of rates of photosynthesis, photo- protection, and biomass growth, as well as a series of biomass yield---related plant architecture traits (e.g., plant height, stalk diameter, leaf number, leaf width, leaf length, leaf angle, leaf area index) in the C4 grass sorghum [Sorghum bicolor (L.) Moench], a promising and productive biomass crop.

These experiments will identify SNP markers within or closely linked to the genes that control these traits. Using these identified SNPs it will be possible to predict the phenotypes of sorghum lines based on their underlying genotypes and conduct genomic selection experiments designed to improve biomass yields of sorghum hybrids.

The C4 grass sorghum is one of the most promising and productive plant species for biomass production in the US. One of the key breeding objectives for biomass crops is increased yield, which is affected by growth rate. Genome-wide Association Studies (GWAS) are being used to identify the genetic control of rates of photosynthesis, photo-protection, and biomass growth, as well as a series of biomass- related plant architecture traits (e.g., plant height, stalk diameter, leaf number, leaf width, leaf length, leaf angle, leaf area index). Identifying the genetic control of biomass growth rates will allow breeders to genetically stack genes that control maximal growth rates, thereby paving a path to producing higher yielding hybrids.

To date, most automated phenotyping systems have been laboratory-or greenhouse-based. These systems suffer from the limitation that plant performance in laboratories or greenhouses is often only poorly correlated with field performance. To overcome these challenges, in this study, biomass volumes of a large genetically diverse collection of sorghum lines was assayed at multiple time points during the growing season with a unique high-throughput, field-based phenotyping robot. This robotic system is expected to contribute widely to the genetic improvement of biomass crops of importance to the US economy.

This project is supported by Agriculture and Food Research Initiative Competitive Grant no. 2012-67009-19713 from the USDA National Institute of Food and Agriculture.

232. HT-CRISPRi studies of gene regulation and function in E. Coli

Harneet S. Rishi1,2*+ (HSRishi@lbl.gov), Esteban Toro,3+ Marcin P. Joachimiak,2 David Chen,4,5 Guillaume Cambray,6 Xiaowo Wang,7,8 Honglei Liu,8 Lei S. Qi,9,10 and Adam P. Arkin2,5

*Presenting Author
+Equal Contributions

1Biophysics Graduate Program, University of California Berkeley, Berkeley, CA; 2Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA; 3Life Sciences Research Foundation, University of California Berkeley, Berkeley, CA; 4UC Berkeley-UCSF Graduate Program in Bioengineering, Berkeley, CA; 5Department of Bioengineering, University of California Berkeley, Berkeley, CA; 6California Institute for Quantitative Biosciences, University of California Berkeley, Berkeley, CA; 7Department of Automation, Tsinghua University, Beijing, China; 8Bioinformatics Division, Center for Synthetic and Systems Biology, Tsinghua National Laboratory for Information Science and Technology, Beijing, China; 9Department of Cellular and Molecular Pharmacology, University of California San Francisco, San Francisco, CA; and 10California Institute for Quantitative Biomedical Research, San Francisco, CA

Research Specialization: Microbial Systems Design for Biofuels

Grant Title: A platform for genome-scale design, redesign, and optimization of bacterial systems URL: http://genomicscience.energy.gov/biosystemsdesign/2012awards.shtml

Project Goals: Synthetic biology aims to leverage the engineering principles of modularity, standardization, and reliability with the design-build-test-learn cycle to rapidly engineer novel biological functions. One of the key hurdles in adopting this strategy is that the innate complexity of biological systems (ex: unmapped interaction networks, context dependence, temporal & spatial population variations) makes it difficult to understand first principles, which in turn makes it difficult to predictably build scalable systems. High- throughput technologies to quantitatively characterize function-phenotype landscapes can help overcome such barriers. As such, we aim to (1) develop CRISPRi into a scalable, high- throughput platform for rapid profiling of genomic features in E. coli across a number of conditions and (2) extend the functionality of the HT-CRISPRi platform to perform double transcriptional knockdowns for investigating the epistatic landscape of E. coli and revealing the basis of complex traits.

The CRISPR-associated protein Cas9 has been adapted as a versatile tool for transcriptional regulation, genome editing, and imaging in a number of organisms. Here, we apply the catalytically inactive dCas9 to conduct high-throughput transcriptional and regulatory studies in coli. Using an Agilent OLS library of 32992 unique sgRNAs, we targeted 4500 genes, 5400 promoters, 640 transcription factor binding sites, and 106 sRNAs in the E. coli genome. These genomic targets cover a wide range of functionalities such as metabolism, stress response, transport, and cell division. By leveraging CRISPRi with next-generation sequencing, we were able to interrogate the fitness effect of transcriptional knockdown for each of the aforementioned genomic features in a single-pot experiment both aerobically and anaerobically. Our fitness results agreed well with current knockout databases, and our ability to induce transcriptional knockdown at any point during an experiment has allowed us to explore target essentiality under different conditions (aerobic, anaerobic, etc.) with great ease. We demonstrated this by showing that although nrdA and nrdB are essential under aerobic conditions – and are annotated as such in databases – they are dispensable anaerobically. We are also extending the HT-CRISPRi platform to scalably perform double transcriptional knockdowns in E. coli and are interrogating libraries of knockdown pairs for epistatic genetic interactions. Overall, HT-CRISPRi enables single-pot, precise measurements of fitness for a large

set of genomic features and will prove useful in genomic studies of model and non-model organisms.

This work is supported by the Genome Science program within the Office of Biological and Environmental Research (Project grant number DE-SC008812, Funding Opportunity Announcement DE-FOA-0000640).

233. 'Species Filter' Effects on Sediment Biofilms and Groundwater Source Diversity

A. Zelaya1,2, K. Bailey3, P. Zhang4, S.P. Preheim5, J. Van Nostrand4, D.A. Elias3,7, E.J. Alm5,7, J. Zhou4,7, P. Adams6,7, A.P. Arkin6,7, and M. W. Fields1,2,7 (matthew.fields@biofilm.montana.edu)

1Center for Biofilm Engineering, Bozeman, MT; 2Department of Microbiology & Immunology, Montana State University, Bozeman, MT; 3Division of Environmental Sciences, Oak Ridge National Laboratory, Oak Ridge, TN; 4Institute for Environmental Genomics, University of Oklahoma, Norman, OK;

5Department of Biological Engineering, MIT, Boston, MA,

6Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA; 7ENIGMA (http://enigma.lbl.gov/)

Project Goals: The goal of Ecosystems and Networks Integrated with Genes and Molecular Assemblies (ENIGMA) is to push the frontier of knowledge on the impact of microbial communities on ecosystems and to elucidate microbial assembly rules. As part of this, we must link genetic capacities to ecological function (e.g., heavy-metal- and radionuclide-contaminated groundwater in DOE sites of interest). In these sites, the microbial communities are often predominantly attached to sediment particles as a biofilm. We have characterized temporal dynamics of bacterial groundwater communities and compared to the establishment of particle- associated bacterial communities on native sediments incubated down-well. In pursuing this goal, inter-laboratory collaboration facilitated by ENIGMA has led to the identification of aquifer population distributions between the liquid/solid boundaries in situ.

Understanding the factors that determine microbial assembly, composition, and function in subsurface environments are critical to assessing contributions to biogeochemical processes such as carbon cycling and bioremediation. However, these factors are still not fully understood. In this study, surrogate sediment samples were incubated for 3 months in 3 wells (FW301, FW303, FW305) within the background site of the Oak Ridge Field Research Center in Oak Ridge, TN. Local sediment biofilm communities were compared to those of the groundwater (source diversity). Groundwater samples from each well were collected approximately 3 times a week. Multiple sediment samples (n=12) were used per well to determine inter- and intra-well variation. Spatial and temporal community analysis of local and source samples via ss-rRNA paired-end sequencing and distribution-based clustering revealed higher richness, diversity, and variability in source groundwater communities compared to sediment---associated communities. Ordination analysis grouped the newly formed local communities as more similar to those of the groundwater than to those of the original parent sediment. The predominant groundwater sequences per well were Curvibacter, Delftia, and Acidovorax for FW- 301, Aquabacterium, Oleiphilus, and Bradyrhizobium for FW---303, and Acidovorax, Curvibacter, Caulobacter, and Elusimibrobium for FW-305. Other sequences displayed transitory predominance for different wells. The community composition was different between wells, and FW-305 (a younger well) showed more diversity over time. In sediment samples, 20-40% of the communities consisted of populations that were abundant at less than 5% of the total sampled diversity. The sediment biofilms from each well were also distinct from each other. Intra-well sediment biofilms showed much less variability, with the exception of FW-305. Sediment biofilm communities were distinct from corresponding groundwater communities, with some populations becoming predominant in the biofilm (e.g. Aquabacterium, Perlucidibaca, and Paraperlucidibaca); however, different OTUs were respective to each well. These results indicate a shift in local community structure that is influenced by the available source community as well as hydrology.

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.

234. A Platform E. coli Strain for Optimizing Biosynthetic Pathways via High-throughput Genome Engineering and Screening

Robert Egbert1,2 (rgegbert@lbl.gov), Harneet Rishi3, Esteban Toro2, Adam Arkin1,2

1Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA;

2Department of Bioengineering, University of California Berkeley; 3Biophysics Graduate Program, University of California Berkeley

Project Goals: Efforts to optimize biofuel synthesis or other engineered functions in microbes is complicated by poor understanding of interactions between engineered components and host functions and among engineered components. Developments in regulated gene expression, genome engineering, high-throughput screening and selection have each made the tasks of understanding host context and engineering novel behaviors easier. However, these developments have not been integrated for use in one strain for an integrated approach to the design, build, test, and learn cycle. This project aims to develop a general-use host platform to perform high-throughput studies to improve biofuel production and engineer other DOE-relevant phenotypes. We plan to extend the strain engineering methodology to other organisms known to utilize unique feedstocks for growth and produce useful chemicals and polymers.

Precision genome engineering is an important tool used to discover the genetic determinants of phenotype and to engineer synthetic gene circuits. Major advances in phage-derived, short-homology recombination systems and high-throughput DNA synthesis have accelerated the integration rate of biosynthetic pathways and mutational diversity to bacterial chromosomes. However, the highest directed mutation rates are currently available in only persistent mutator strains, which could generate mutations on introduced biosynthetic pathways, thus limiting efforts towards multi-stage genome engineering and directed evolution. These strains also lack a collection of inducible regulators to concurrently control the expression of genes that mediate genome engineering along with integrated gene expression cassettes.

Here we engineer a platform strain of E. coli that includes multiple inducible regulators and a high-efficiency recombination system that can act as a chassis for engineering and evolving complex gene networks for biofuel production and other uses. We include engineered induction systems for lactose, cumate and arabinose to allow multi-input control of genome engineering functions and biosynthetic pathways. We have re-engineered the lambda Red recombination cassette to reduce toxicity and recombination cycle time. We have acheived single-cycle mutation rates of nearly 20% in a transient mutator background by introducing multiple mutations to increase single-stranded DNA half-life in cells and its accessibility to the chromosome. We also characterize gene expression, dsDNA recombination efficiencies and ssDNA recombination rates for a collection of context-neutral genomic loci, termed SafeSites. This strain will be a resource for engineering, characterizing and optimizing complex synthetic gene networks via iterative cycles of metabolic modeling and high-throughput screening and will be used to model the development of similar induction and recombination systems in other DOE-relevant target organisms.

This work is supported by the Genome Science program within the Office of Biological and Environmental Research (Project grant number DE-SC008812, Funding Opportunity Announcement DE-FOA-0000640)

235. Expanding the chemical diversity of E.coli derived products

Robert W. Haushalter, Leonard Katz, and Jay D. Keasling

While the majority of crude oil refined in the world is used form the production of fuels, many other important industrial compounds derived from oil are used in the production of fabrrics, plastics, and other materials. Engineering microbes to produce these compounds from renewable sources will be crucial to decreasing our dependence on petroleum. We are therefore focusing efforts on producing long chain dicarboxylic acids, hydroxy fatty acids, and diols from glucose in E. coli. These compounds are used in the production of polyesters and nylons. We have established a system to produce long chain dicarboxylic acids in E. coli by engineering the biotin and fatty acids biosynthetic pathways, and are working towards production strains for hydroxy fatty acids and diols. The production of these relatively high-value compounds has a greater potential for economic viability than fuels, and will further decrease our need to produce and import oil.

236. Synthetic Biology-Enabled Biosystems Design and Construction of Radically Recoded Organisms

Julie Norville, Marc Guell, Jun Teramoto, Matthieu Landon, Nili Ostrov, Michael Napolitano, Marc Lajoie, Gleb Kuznetsov, Kerry Singh, Maria Mincheva, Victoria Longe, Ellen Shrock, Kento Tominaga, Mark Moosburner, James DiCarlo, Benjamin Pruitt, Nicholas Conway, Barry Wanner, George Church

Why it is important to radically modify microbial genomes? Four reasons are to develop multi-virus resistant cells, enable efficient use of novel amino acids, and enforce genetic and metabolic isolation. Recently, we successfully developed a biocontainable C321.ΔA strain that lacks the UAG stop codon and functionality, which cannot grow without human provision of a non-standard amino acid (NSAA) and exhibits only very minimal evolutionary escape rates in the absence of the NSAA. We our now extending upon this work to construct a radically recoded organism (RRO). Building upon our experience in removing the stop codon and functionality of UAG in the C321.ΔA E. coli, the testing of multicodon reassignment in E. coli essential genes, and our analysis of E. coli promoter/ribosome binding sites and promoter/N terminal codon sequence combinations, we developed computationally aided design tools for radical recoding to facilitate the in silico reassignment of 7 codons in all E. coli genes (62,733 codons or 5.4% of the genome), in order to free up 3 anticodons for non-standard amino acid usage and leave one codon unused. DNA synthesis of 99.7% of the 3,972,391 base pair radically recoded genome is complete. We have designed and optimized a method for using integrases to replace genomic segments with synthetic recoded sequence, and have made 2,358 codon substitutions (7 of 64 codon types) in 3 sequence verified E. coli strains. Our software package Millstone provides the first software platform that can fully automate the intense computational process required to analyze the genomes of highly modified strains, such as RROs. Once complete, our RROs will provide a safe organism that can incorporate multiple redundant genomic and metabolic safeguards that prevent release or deliberate misuse. Moreover, RROs can be combined with our in vivo sensor selector systems and pathway engineering, allowing us to both sense and select for strains that produce optimal quantities of a desired biomolecule, without interference from phage and within a biocontained organism.

Index