Office of Biological and Environmental Research Biological Systems Science Division Update

Todd Anderson, Ph.D.

Director, Biological Systems Science Division, Department of Energy, Office of Biological & Environmental Research

October 24, 2019

Update on Programmatic Activities

Completed Reviews/Activities

- ✓ ANL Structural Biology Center (SBC) Review May 29-30
- ✓ Panel Review of Early Career FOA Applications June 6
- ✓ Panel Review of Plant Genomics FOA Applications June 25-27
- ✓ Panel Review of Environmental Microbiome FOA Applications June 25-27
- ✓ Panel Review of Bioimaging FOA Applications June 26-27
- ✓ LBNL mCAFEs Reverse Site Review Aug 20th

Upcoming Reviews/Activities

- ❑ Bioenergy Research Center (BRC) Annual Progress Reviews
  - CBI Nov 6-7, JBEI Dec 3-4, GLBRC Jan 15-16, CABBI Feb 4-5
- ❑ 2020 Genomic Science Annual PI Meeting Feb 24-26
- ❑ 2020 Bioimaging Science Annual PI Meeting Feb 26-27
- ❑ BNL QPSI Reverse Site Review Mar 17
- ❑ QIS Workshop Mar 12-13
- ❑ JGI User Meeting Mar 22-26

MICROBIOME TO FUNCTION

NEXT GENERATION · ECO-MICROBIOLOGY

Workshop and Celebration
Santa Fe · September 15-17, 2019

LANL Microbiome Workshop In Honor of Dr. Cheryl Kuske

- ➤ Attended by all the BSSD DOE Laboratory SFA programs
- ➤ Excellent forum for sharing information among the Lab programs
- ➤ Annual Meeting (next year hosted by LLNL/LBNL)

NEW FY 2020 FUNDING OPPORTUNITIES (TENTATIVE)

Funding Opportunity Announcements (FOAs)

Systems Biology Research to Advance Sustainable Bioenergy Crop Development

Computational Tool Development for Integrative Systems Biology Data Analysis

Likely posting in November
- Subject to availability of funds in FY 2020

Opportunities to the DOE Laboratories

*New Quantum Enabled Bioimaging Approaches for Bioenergy**

*Secure Biosystems Design**

Likely to be communicated to selected Labs in November
- Also, subject to availability of funds in FY 2020

* Mentioned in the FY2020 President's FY 2020 Budget

Genomics-Enabled Plant Biology for Determination of Gene Function (DE-FOA-00002060)

12 new projects

Novel genomics-based approaches and technological innovations that will lead to transformative knowledge in validating gene function in energy crops:

- ➤ Structural and functional genomic characterization of conserved unannotated proteins in Arabidopsis and Setaria
- ➤ Functional analyses of oil storage and stability in pennycress
- ➤ Molecular mechanisms of drought and heat resilience in sorghum

PI: Hiroshi Maeda
Univ of WI

PI: Clint Chapple
Purdue Univ

PI: Kranthi Varala
Purdue Univ

PI: James Schnable
Univ of NE

PI: Gail Taylor
UC Davis

PI: Martin Jonikas
Princeton Univ

PI: James Umen
Danforth Center

PI: José Dinneny
Stanford Univ

PI: Andrea Eveland
Danforth Center

PI: Seung Rhee
Carnegie Institution

PI: Ana Alonso
Univ of North TX

- ➤ Development of acetyltransferase toolbox for plant biomass engineering
- ➤ Use of isotope labeling, LC-MS, GWAS to determine metabolites and gene annotations in Arabidopsis and sorghum
- ➤ Characterization of protein-protein interactions and localization of chloroplast associated proteins in *Chlamydomonas* and Camelina
- ➤ Construction of a nitrogen flux map and determine system-level functionality of plant N metabolic network in Arabidopsis, poplar, Medicago, and sorghum

New **Environmental Microbiome Awards**

Systems Biology Enabled Research on the Roles of Microbiomes in Nutrients Cycling Processes (FOA-0002059)

13 New Projects

Systems biology studies on regulatory, metabolic, and signaling networks of microbes, microbial consortia, and microbe-plant interactions involved in biogeochemical cycling of nutrients

- Microbial responses to drought
- Biogeochemcial cycling of C, N, P, and Cu
- mycorrhizal fungi impacts on biogeochemistry
- Cross-kingdom interactions in nutrient cycling
- The impact of fire on soil microbial processes
- Feedback mechanisms and functional traits in response to soil warming
- Organic carbon degradation and greenhouse gas release in permafrost
- Isotope techniques to track single viral particles in complex systems
- Organic nutrients and the role of communities in producing shared resources
- The impact of viruses and protists in soil microbial ecology and elemental turnover
- Integrating single-cell measurement with ecosystem-scale biogeochemical fluxes


PI: Mari Winkler
U Washington Seattle

PI: Jennifer Bhatnagar
Boston U

PI: Jeremy Semrau
U Michigan

PI: Thea Whitman
U Wisconsin Madison

PI: Michi Taga
UC Berkeley

PI: Mary Firestone
UC Berkeley

PI: Victoria Orphan
Caltech

PI: Matt Sullivan
Ohio State U

PI: John Coates
UC Berkeley

PI: Karen Lloyd
U Tennessee

PI: John Gibert
Duke U

PI: Steven Allison
UC Irvine

PI: Bruce Hungate
Northern Arizona U

New Bioimaging Research Awards

New Bioimaging Approaches for Bioenergy (DE-FOA-00002041)

6 New Projects

New approaches to enable imaging of biological targets non-destructively and in real time, to dramatically enhance our ability to measure biological processes in and among living cells

PI: Prashant Nagpal
Univ Colorado

PI: Peng Chen
Cornell Univ

PI: Shimon Weiss
UCLA

PI: Gary Stacey
Univ Missouri

PI: Peng Chen
Univ Delaware

PI: Michelle O'Malley
UCSB

➤ Quantum dots (QDs) as sensors tuned and functionalized to interact with specific cellular targets for imaging

➤ Approaches to measure and manipulate biological processes in living cells non-destructively in real time at super-resolution

➤ New approaches envisioned to overcome current challenges of suboptimal stability and photo-bleaching of synthetic fluorescent proteins to enable prolonged imaging and detection in living systems

➤ New approaches to track dynamics of native and engineered lignocellulosic biomass, and to image rhizosphere chemistry in situ across spatiotemporal scales in native environments
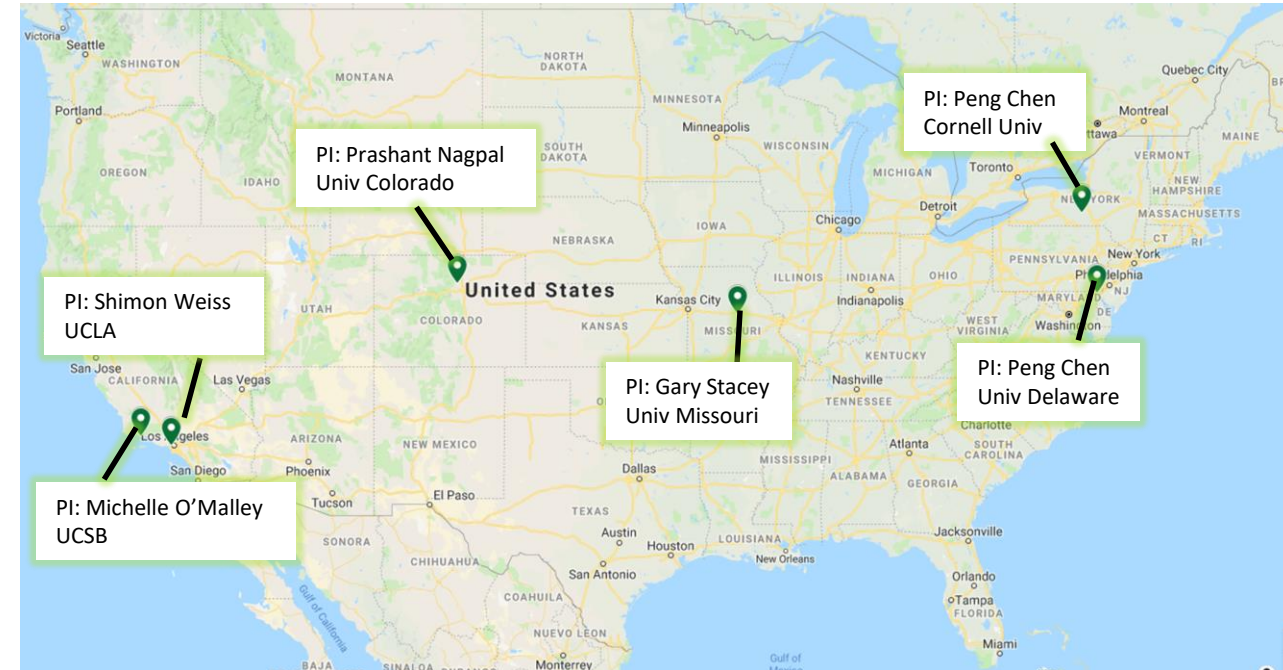
New Lab Funding Biosecurity Awards and Computational Biology Awards

4 New Biosecurity Pilot Projects
(12 months)

Machine Learning-Guided Design of Efficient Safeguard Systems that Operate Under Different Bacterial Physiologies **ANL [PI: Noirot]**

Lock and Key: Securing Genomes Against the CRISPR/CAS9 Gene-Editing Tool **ORNL [PI: Abraham]**

Genome remodeling to control the persistence of engineered functions in soil microbes **PNNL [PI: Egbert]**

Fail-safe Genome Engineering **SLAC [PI: Salit]**

5 New Computational Biology Awards
2-3 years of funding

Optimal experimental design of biological systems **BNL [PI: Alexander]** 2 yrs of funding

Genomes to predictive biology: machine learning for the integration of inter-species functional genomics data **BNL [PI: Yoo]** 2 yrs of funding

Accelerate JGI-KBase Codevelopment Effort on Homology Microservices **LBNL [PI: Fagnan]** 2 yrs of funding

Exascale & Petascale Networks for KBase **ORNL [PI: Jacobson]** 2 yrs of funding

Elucidating Principles of Bacterial-Fungal Interactions **PNNL [PI: Cannon]** 3 yrs of funding

PNNL is building visual proteomics capabilities to enable gene function understanding and prediction

Increase cryo-EM accessibility for BER researchers through providing sample screening as well as data collection capability.

New cryo-EM at EMSL

- BSSD provided 50% funds to EMSL for purchase of a 300 kV cryogenic electron microscope
- Leverages infrastructure PNNL has implemented for the Pacific Northwest Cryo-EM Center

Access & Impacts

- Through the proposal process, BSSD researchers will have dedicated access to 50% instrument time; BER researchers will have priority access.
- Will enhance multimodal visual proteomics pipeline for insights into gene function

New

Thermo Fisher Krios G3i

The National Microbiome Data Collaborative
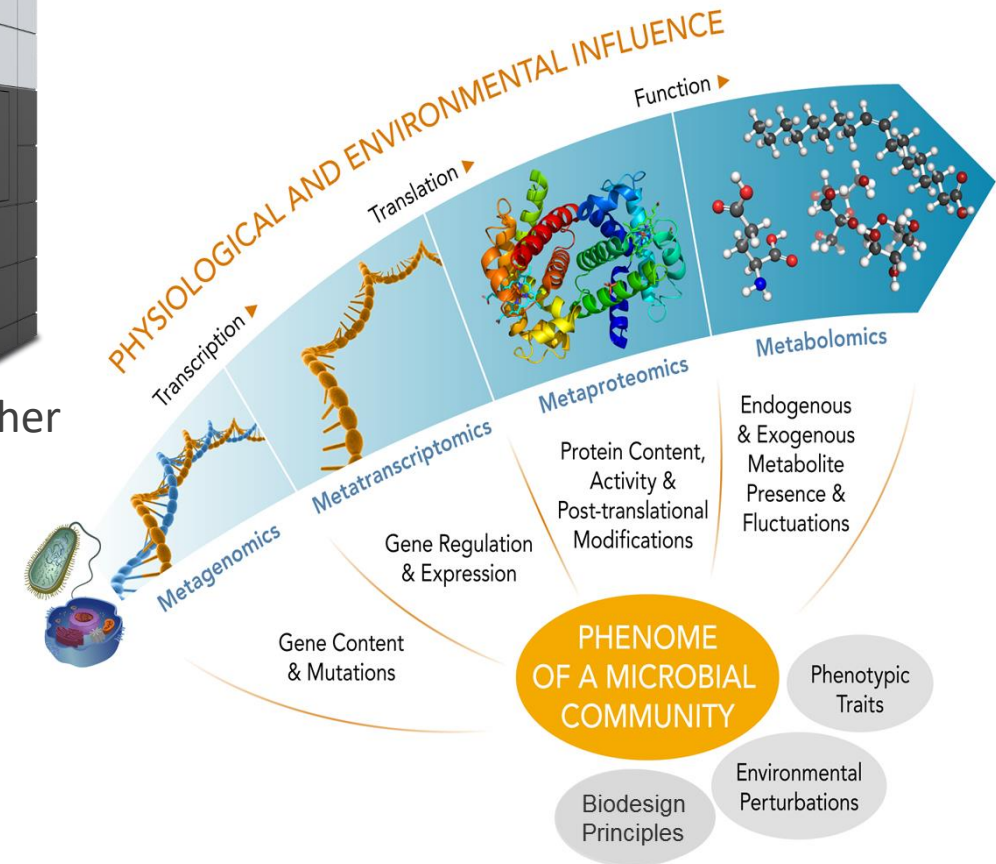
Initiated in July

www.microbiomedata.org
@microbiomedata

Emiley Eloe-Fadrosh (LBNL)
NMDC Lead

Computational Biosciences:KBase Developments
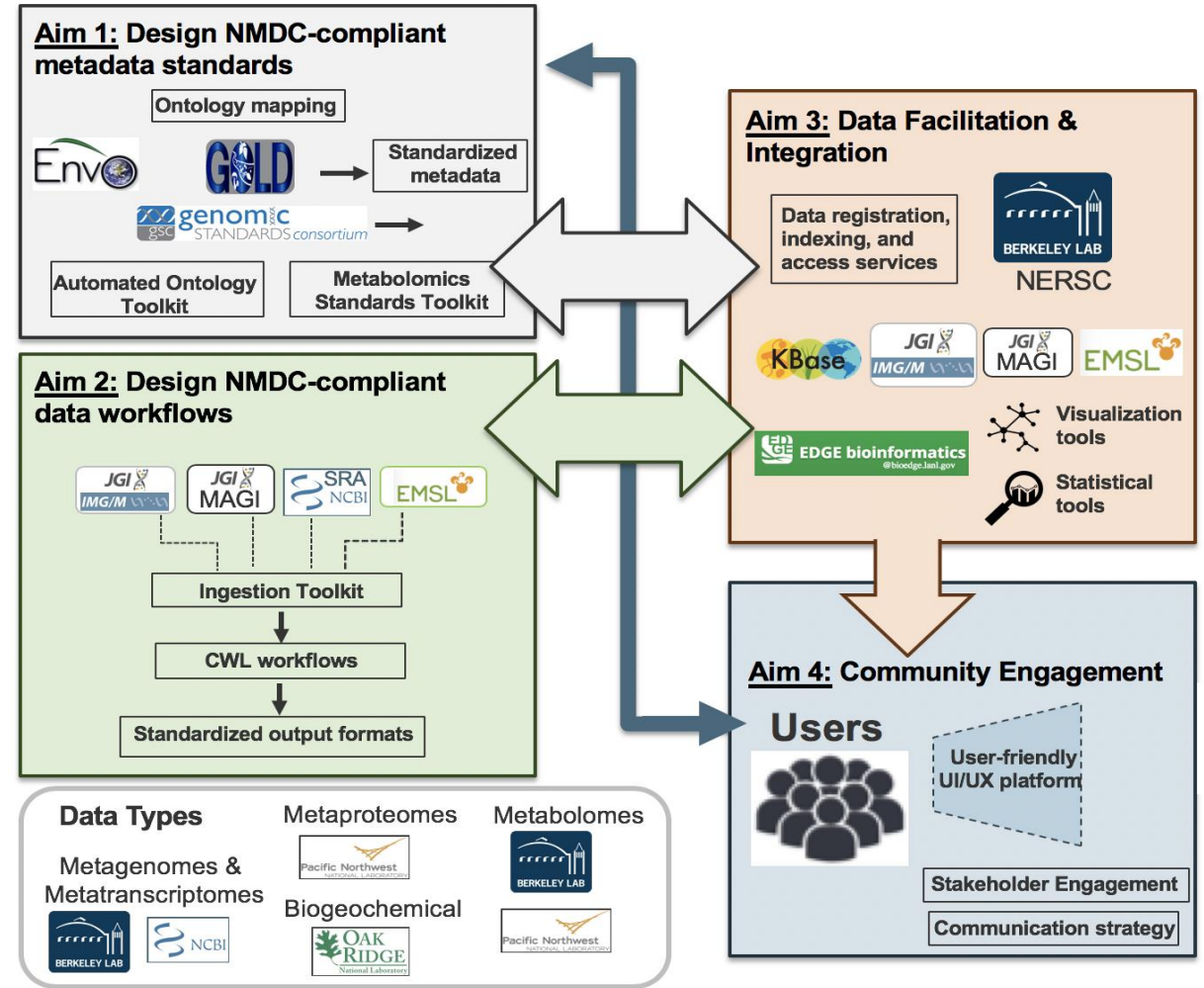
Increasing User Impact

- User data >115 TB
- Narratives > 16K
- Publications > 125
- User growth > 6.5K
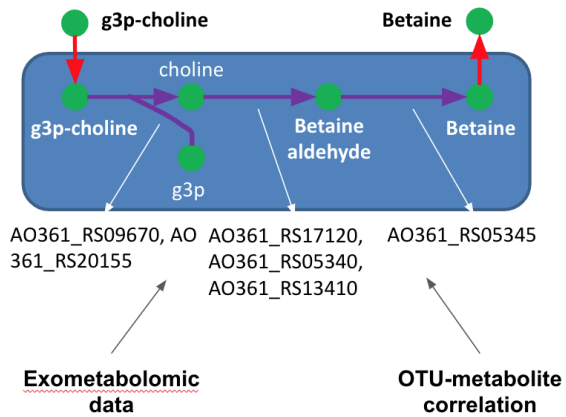- Working with 5+ SFA's to get their data and tools into KBase

26%

Cumulative Narratives

Months (1/2016 – 9/2019
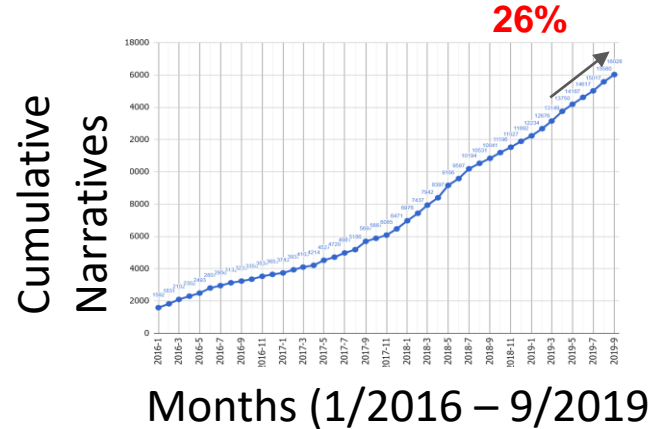
Running Total of KBase Publications and Mentions (September 25, 2019)

47%

- Mentions of KBase
- Pubs Citing KBase in Methods

Year
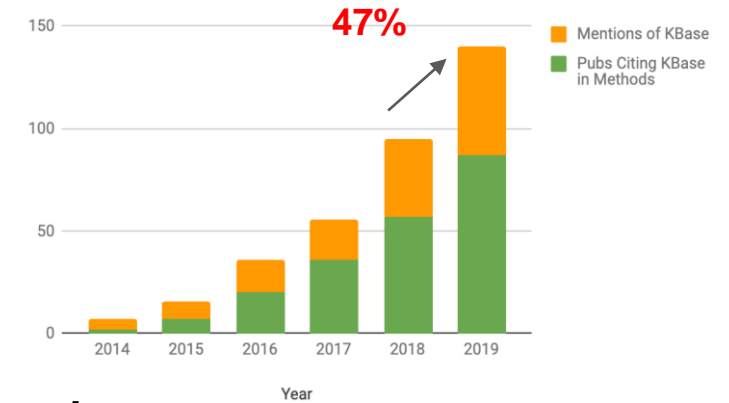
Examples of Results from User Working Group Collaborations

Integrating multi-omics data to highlight ecological mechanisms in ENIGMA

g3p-choline

Betaine

choline

g3p-choline → Betaine aldehyde → Betaine

g3p

AO361_RS09670, AO 361_RS20155

AO361_RS17120, AO361_RS05340, AO361_RS13410

AO361_RS05345

Exometabolomic data

OTU-metabolite correlation

LLNL SFA: building workflow for loading and reconciling multiple annotation sources

Genes with Annotated Function

Unique ModelSeed (rxns)

Annotation Tools
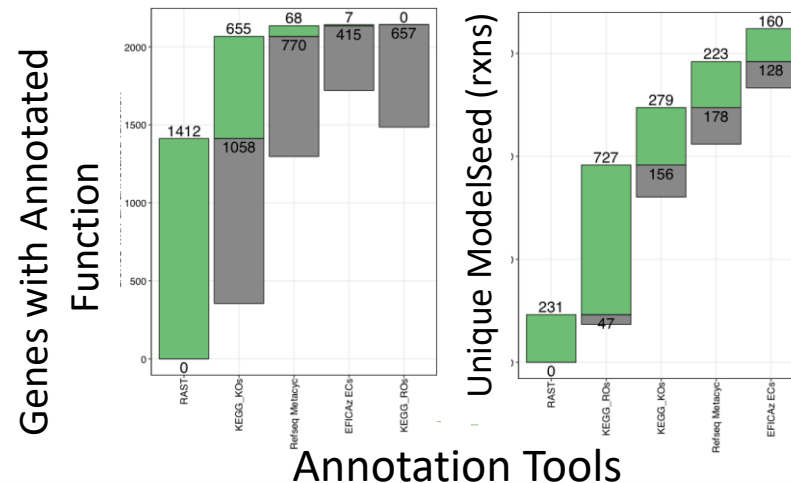
ORNL SFA (with Protein Data Bank and EMSL) Protein-Ligand Interaction

LLNL Soils SFA (with Ohio State, JGI) Annotation and Ecology of Phage

LLNL Soils SFA, ENIGMA, JGI Extraction and Comparison of MAGs Across Diverse Environments

Microbial Carbon Use Efficiency Predicted from Genome-scale Metabolic Models
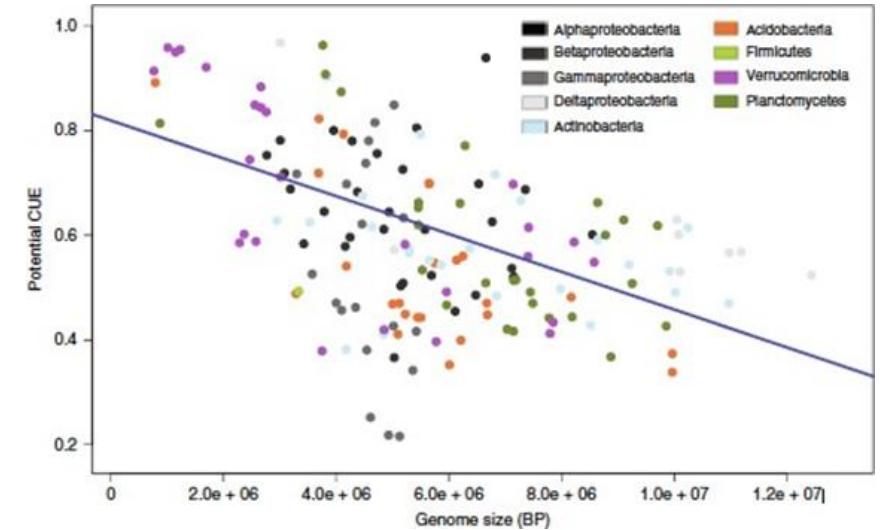
Objective
Characterizing variation in microbial carbon use efficiency (CUE) across taxa and substrate types in soils.

Approach

- Generated theoretical predictions of CUE for more than 200 taxa using genome-scale constraint-based metabolic modeling.
- Used genome-scale metabolic models to predict microbial physiology based on genes.
- Developed hypotheses on CUE structure across taxa and substrate types.

Potential CUE versus genome size. Potential CUE regressed against genome size (bp).
Blue lines show GLS fit. Points are colored by phylum.

Results/Impacts

- A wide range in average potential CUE (0.62 ± 0.17) across taxa with structuring at subphylum levels.
- CUE declines with increasing genome size, while larger genomes are able to access a wider range of C substrate types.
- These findings provide a framework for predicting CUE from genomic traits and for inferring potential impacts of shifts in bacterial community composition on C cycling.
- Phylogenetic variation in microbial CUE can dramatically alter patterns in soil C cycling over time.

Methyl ketone production by *Pseudomonas putida* is enhanced by plant-derived amino acids
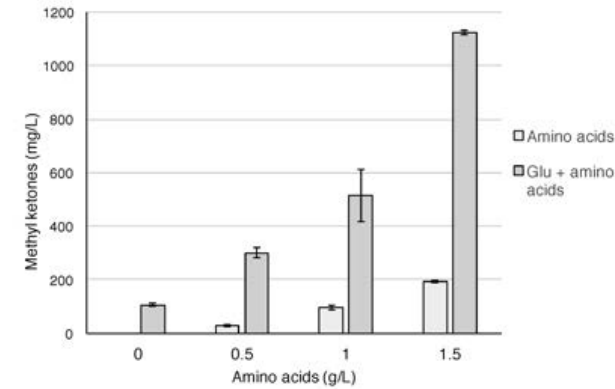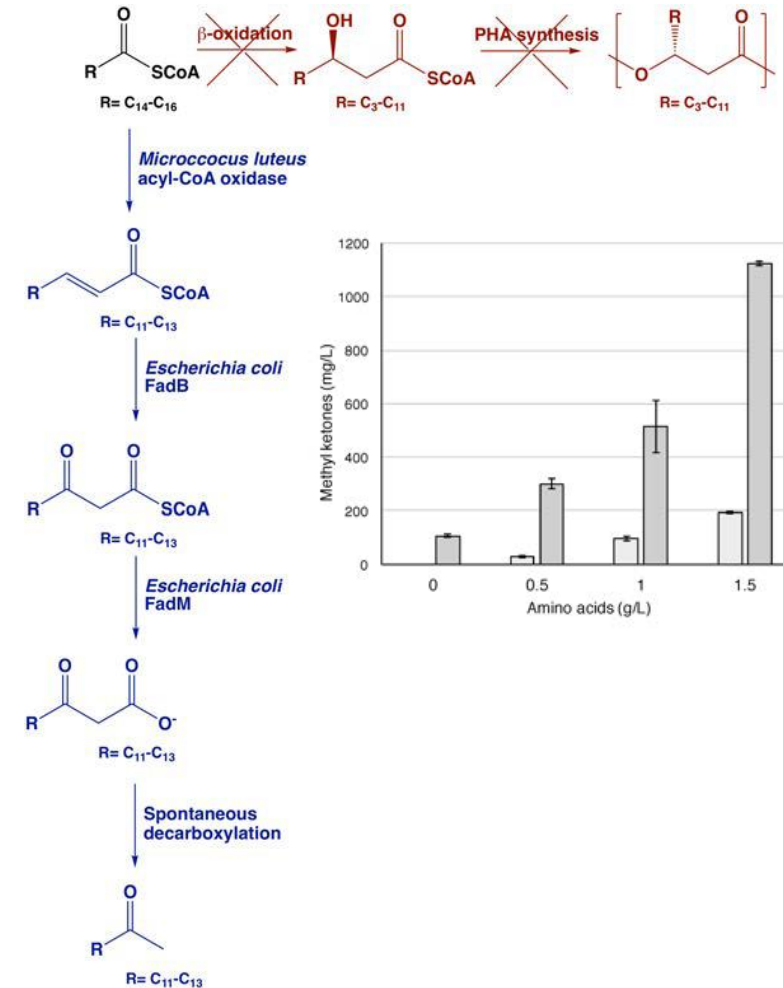
Objective

- Methyl ketones are a promising diesel blendstock and have been produced by multiple microbial platforms.
- Test use of plant-derived aromatics and amino acids to improve the efficiency of overall biomass conversion to methyl ketones.

Approach

- Engineered *Pseudomonas putida* to produce methyl ketones by introducing a truncated beta-oxidation pathway.
- Tested the strain on pure substrates and plant hydrolysates containing significant amounts of plant-derived aromatics.

Results/Impacts

- Engineered *P. putida* produced methyl ketones from glucose and lignin-related aromatics.
- Methyl ketone production by *P. putida* increased with plant hydrolysates as substrates in comparison to sugar controls.
- Metabolomics and proteomics demonstrated that the increased production was due to plant derived-amino acids.
- Inclusion of a defined amino acid mixture increased methyl ketone production 9-fold relative to a sugar-only control.

Dong et al. (2019) *Biotechnology and Bioengineering*, doi: 10.1002/bit.26995.

Transformation of *Clostridium thermocellum* via complete methylome analysis and directed DNA methylation

Objective

- *Transform Clostridium thermocellum.*
- *C. thermocellum* is an anaerobic, thermophilic bacterium that efficiently deconstructs lignocellulosic biomass, but can not be genetically manipulated.
- Test if directed methylation of DNA will protect it from *C. thermocellum* restriction-modification systems.
- Characterize methylome of *C. thermocellum* genome.

Approach

- Sequence *C. thermocellum* ATCC 27405 using PacBio single-molecule real-time (SMRT) sequencing to detect 6-methyladenine and 4-methylcytosine, and the rarely used whole-genome bisulfite sequencing to detect 5-methylcytosine.
- Clone and express *C. thermocellum* ATCC 27405 DNA methyltransferases genes in *E. coli*.
- Pass transformation DNA through *E. coli* expressing methyltransferases.

Results/Impacts

- No transformation was detected without methylation of DNA, but up to 80 transformants per μg DNA with methylation.
- Specifically provides for transformation and genetic manipulation of this important microbe
- More broadly provides tools for genetic transformation and engineering of other non-model microbes

Bioenergy Research Centers Science Highlight

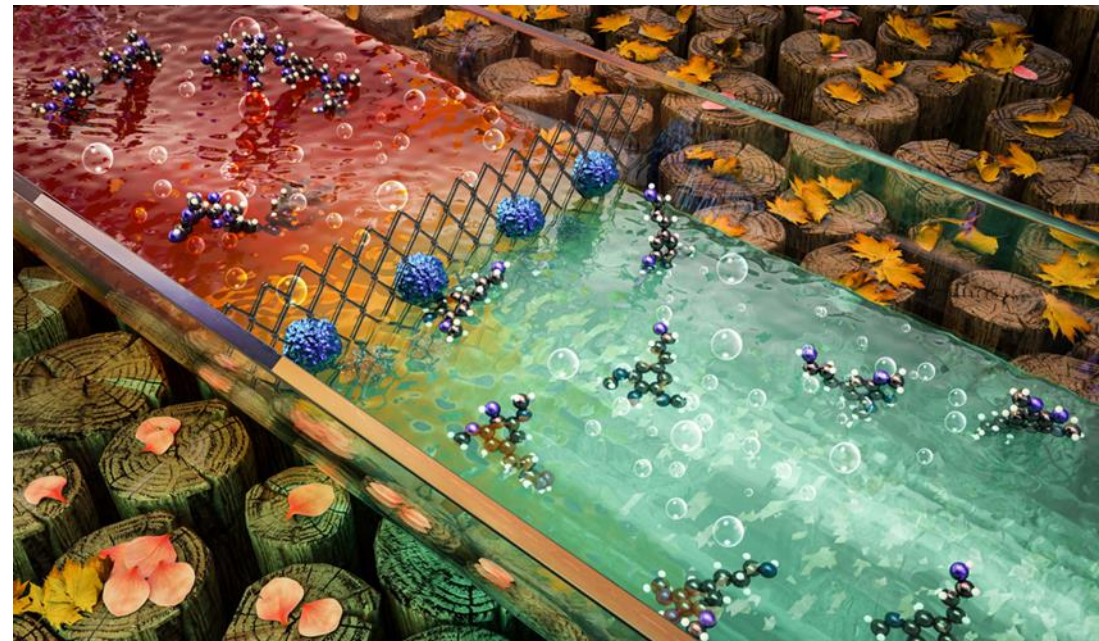## Hydrogenolysis offers an improved method to break down lignin

Objective

Explore lignin hydrogenolysis in a continuous-flow reactor as a method to improve depolymerization.

Approach

➢ Evaluated the kinetics and reaction mechanisms of lignin hydrogenolysis in a continuous-flow reactor.

➢ Optimized reaction parameters to increase monomer yield and suppress formation of unwanted byproducts.

➢ Assessed hydrogenolysis of lignin isolated from poplar.

The article has been selected for the cover of Green Chemistry.

Results/Impacts

➢ Lignin hydrogenolysis in the continuous-flow reactor produced lignin-derived monomer yields around 29 percent, close to the calculated theoretical maximum for lignin from pretreated biomass.

➢ Unreacted lignin retained its native structure, allowing material to be recycled through the reactor to minimize carbon loss into unusable compounds.

➢ Combining a continuous-flow reactor with biomass treatment processes could enhance the yield and value of product streams from a biorefinery.

Y. Li, et al. "Kinetic and mechanistic insights into hydrogenolysis of lignin to monomers in a continuous flow reactor." Green Chemistry (2019). 10.1039/c9gc00986h

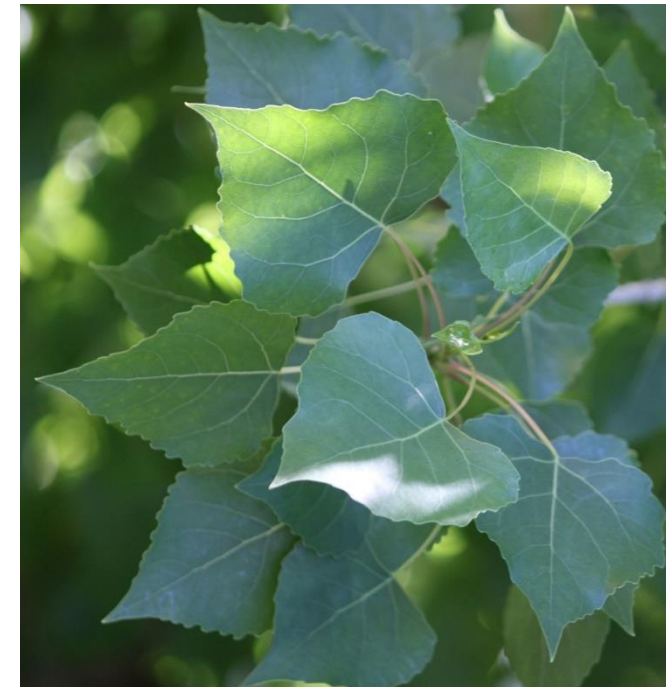## How Many Copies Does it Take to Change a Trait?

Objective

Investigate the impact of differences in gene copy number on yield and other important, hard-to-measure traits.

Approach

- Developed a collection of clonally propagated F1 hybrid poplar trees in which different plants have DNA segments that are repeated (insertions) or deleted
- Performed a genome-wide screen for dosage-sensitive loci affecting quantitative trait variation for numerous bioenergy-relevant traits
- Investigated the correlation between gene dosage at specific chromosomal locations and phenotype to identify dosage quantitative trait loci (dQTL)

Results/Impacts

- Many genomic regions were sensitive to gene dosage changes and were correlated with key bioenergy-relevant traits
- Indicates that copy number variation and gene dosage help explain quantitative trait variation in *Populus*
- Findings should help improve predictions of high-performing trees

TEXAS
The University of Texas at Austin

MICHIGAN STATE
U N I V E R S I T Y

Sustainable Bioenergy Research

Environmental Microbiology

GREAT LAKES BIOENERGY
RESEARCH CENTER

JGI

Developing Switchgrass for Biomass Production: Community Gardens Help Distinguish Genetic Bases of Fitness Traits from Climatic Influence


Objective

Investigate the genetic basis of local adaptation in order to increase biomass yield for sustainable biofuel production.

Approach

- Established switchgrass community gardens in 10 different field sites in US
- Developed and clonally propagated > 400 lowland x upland hybrids
- Samples collected for sequencing and analysis at all sites over two full years
- QTL mapping applied to investigate genes involved in key fitness traits, and how these genes interact with the environment

Results/Impacts

- Spring emergence and flowering time determined by location
- Fewer tradeoffs than expected found in genetic contribution to local adaptation across geographic range
- Large-scale field experiment demonstrates potential for development of "generalist" switchgrass by combining locally advantageous alleles
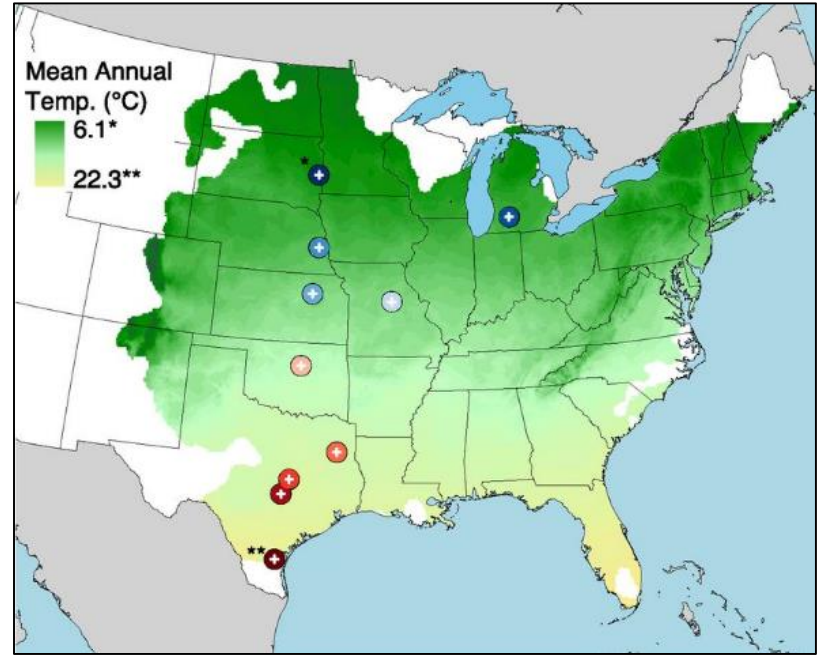
Lowry DB et al. QTL x environment interactions underlie adaptive divergence in switchgrass across a large latitudinal gradient. *Proc Natl Acad Sci* U S A. 2019 June 10. doi: 10.1073/pnas.1821543116

A Viral Gold Rush: New tool to find viruses hiding in genomic data

Objective

Build a genome-based viral identification tool to better capture the viral diversity

Approach

Use a bioinformatical approach for characterizing viral diversity and the complex dynamics of virus-host interactions in nature

vConTACT: automatic network-based taxonomic classifier of prokaryotic viruses

Results/Impacts

✓ Microbiomes contain a myriad of uncultivated viruses, but there is lack of tools to detect them in metagenomic datasets

✓ Developed vConTACT 2.0, an open source, community-available, network-based software application to establish prokaryotic virus taxonomy that scales to thousands of uncultivated virus genomes/fragments, while integrating multiple confidence scores for all taxonomic predictions

✓ Approach can resolve taxa through an integrated distance-based hierarchical algorithm

✓ Classified >1,300 previously unclassified reference viruses

✓ Can be scaled to metagenomic datasets where it could potentially uncover thousands more viral sequences
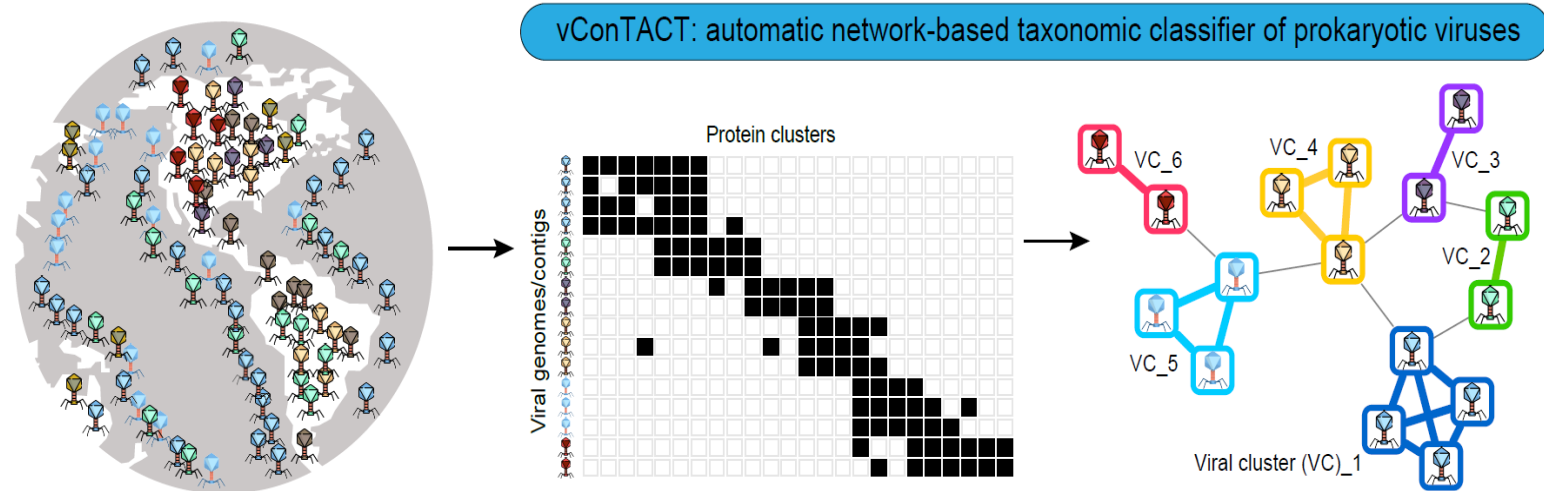
# Molecular mechanisms of nature's fastest CO_2 fixing enzyme

Objective

Understand the molecular mechanisms of nature's fastest CO_2 fixing enzyme, a bacterial enoyl-CoA reductase/carboxylase (ECR), for improving C-cycling energy efficiency

Approach

Elucidate the CO_2 activation mechanism using protein crystallography, quantum mechanical / molecular mechanism (QM/MM) simulations, enzyme activity assay

Kitasatospora setae

Results/Impacts

➤ Determined high resolution structure of ECR tetramer, identifying 4 residues critical in capturing and activating CO_2 using QM/MM simulations
➤ Showed ping-pong mechanism of the coupled movements of catalytic domains for fast CO_2 fixation
➤ Provided basis for engineering a large class of CO_2 fixing enzymes for improving their rates and promiscuity/specificity
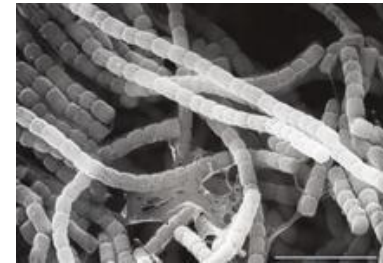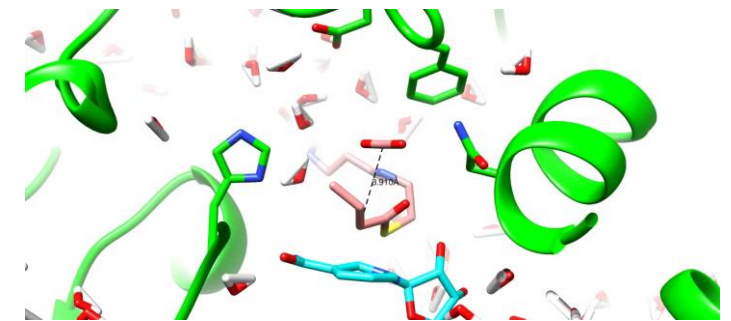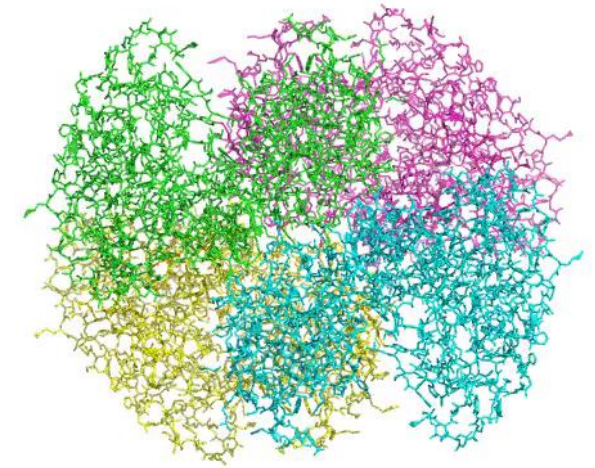
QM/MM simulations of CO2 captured in ECR based on the X-ray structure

G.M.M Stoffel et al., "Four amino acids define the CO2 binding pocket of enoyl-CoA carboxylases/reductases", PNAS 2019, 116 (28) 13964-13969; https://doi.org/10.1073/pnas.1901471116; DeMirci et al., bioRxiv 607101; doi: https://doi.org/10.1101/607101
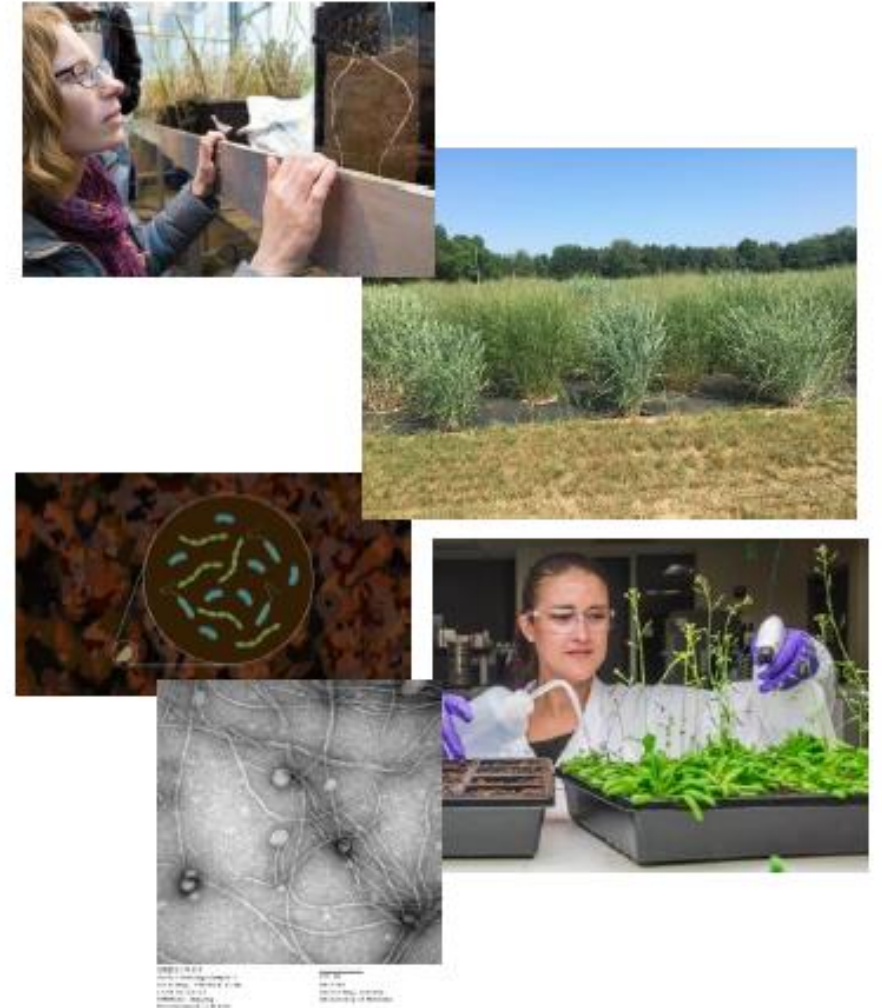
116 JGI Publications since March 27, 2019

Selected Highlights:

- QTL x environment interactions underlie adaptive divergence in switchgrass across a large latitudinal gradient. *PNAS*

- Assembly and seasonality of core phyllosphere microbiota on perennial biofuel crops. *Nature Communications*

- Cryptic inoviruses revealed as pervasive in bacteria and archaea across Earth's biomes. *Nature Microbiology*

- High-throughput single-cell transcriptome profiling of plant cell types. *Cell Reports*

- Probing the active fraction of soil microbiomes using BONCAT-FACS. *Nature Communications*

- Chemical synthesis rewriting of a bacterial genome to achieve design flexibility and biological functionality. *PNAS*

- EcoFABs: advancing microbiome science through standardized fabricated ecosystems. *Nature Methods*

- Towards a genome-based virus taxonomy. *Nature Microbiology*
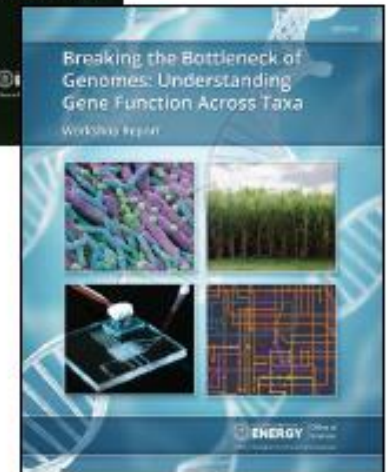
JGI news releases & highlights:
http://jgi.doe.gov/news-publications/
http://jgi.doe.gov/category/science-highlights/

Thank you

Workshop Reports
Now Available at:
http://genomicscience.energy.gov

http://science.energy.gov/ber

U.S. DEPARTMENT OF ENERGY | Office of Science

April 2019 BERAC Meeting   BSSD Updates   Biological and Environmental Research