

# Facilities Integrating Collaborations for User Science (FICUS) and other joint facility collaborations



- Harvey Bolton, Acting Director, EMSL
- Nigel Mouncey, Director JGI



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# Outline

- **Facility Overviews**
  - Environmental Molecular Sciences Laboratory (Harvey)
  - Joint Genome Institute (Nigel)
- **The FICUS Initiative**
  - Goals, Program and Coordination (Nigel)
  - Metrics of Success (Harvey)
- **Scientific Impact**
  - Michelle O'Malley (Harvey)
  - Kelly Wrighton (Harvey)
- **Expanded FICUS Activities**
  - JGI, NERSC and KBase (Nigel)
  - EMSL and ARM (Harvey)
- **Future FICUS Directions**
  - Structural Biology (Nigel)
  - Biogeochemistry and Data Interactions (Harvey)



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# Environmental Molecular Sciences Laboratory (EMSL)

Harvey Bolton  
Acting Director



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# EMSL at a Glance

- Large multidisciplinary user facility with expert staff - mechanisms and dynamics driving biological and environmental processes and interactions - unique collection of 150 instruments and supercomputer
- 300+ publications (Q3 FY2018), 561 users and 265 active projects
- Becoming a user - 4-page proposal
  - Annual call (Jan., 2 years) – Science Areas, FICUS
  - General (anytime, 1 year)
  - Rapid access (anytime, 1-2 month)
  - Scientific partner (anytime)

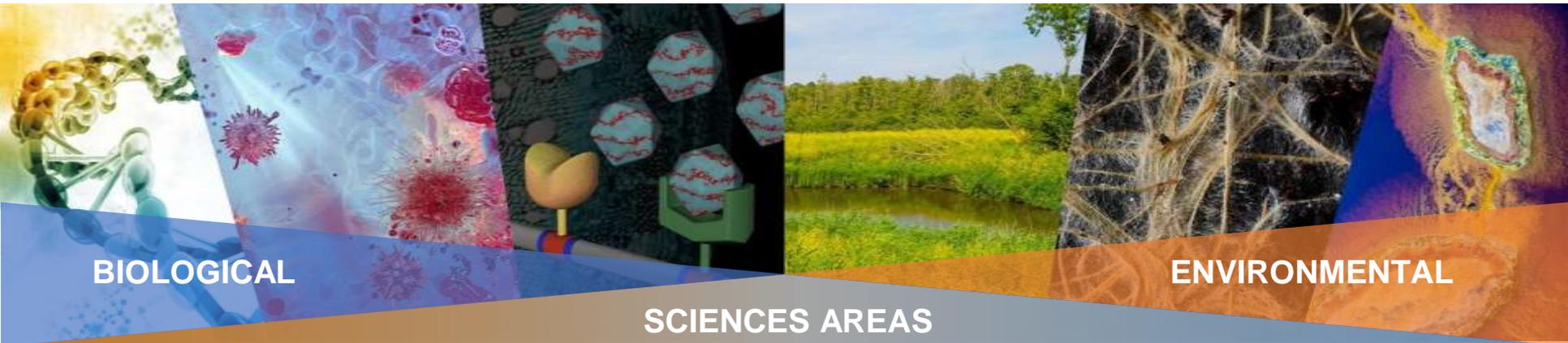


U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# EMSL's Science Areas



**BIOLOGICAL**

**ENVIRONMENTAL**

## SCIENCES AREAS

- Molecular “machines,” biological processes, interactions, and models
- Design plants, fungi, and microbes for bio-based fuels and products
- Complex plant and microbial metabolism impacting carbon, nutrient, and elemental cycles
- Fundamental biogeochemical, plant, microbe, hydrologic, and atmospheric processes
- Mechanistic understanding of these processes
- Interdependencies across scales
- Representation in predictive models

## FICUS

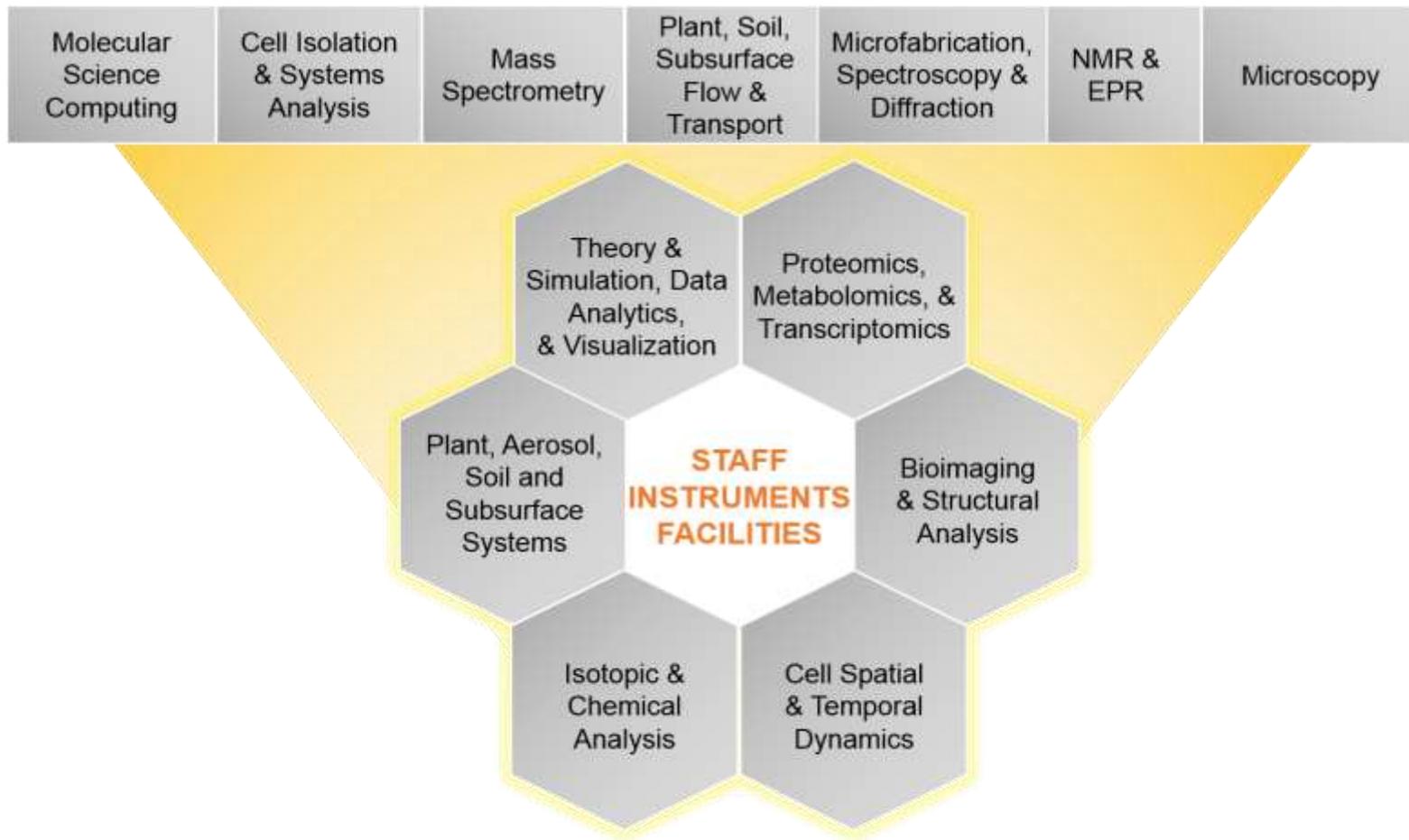


U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# Multiple Capabilities and High-Performance Computing Deliver High-Impact Science



Scientific innovation through integration



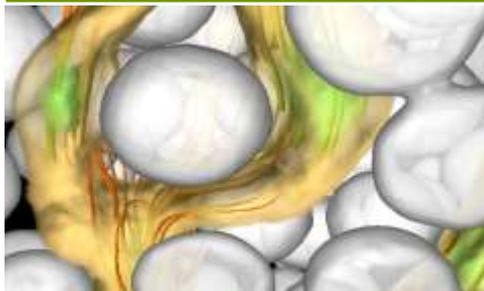
U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



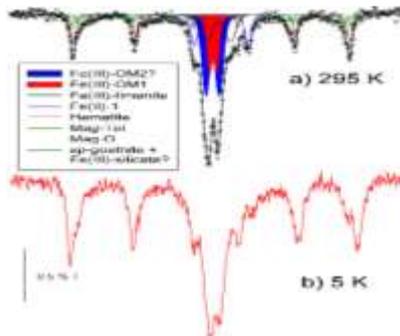
# Scientific Innovation through Integration – Organic-Mineral Interactions

Pore-scale  
Dynamics



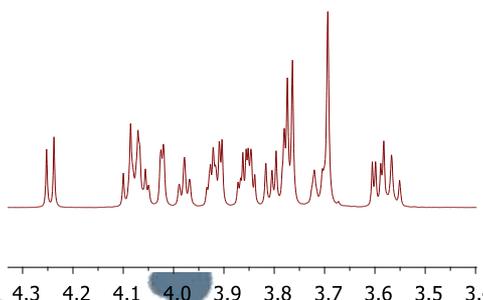
HPC

Fe Speciation

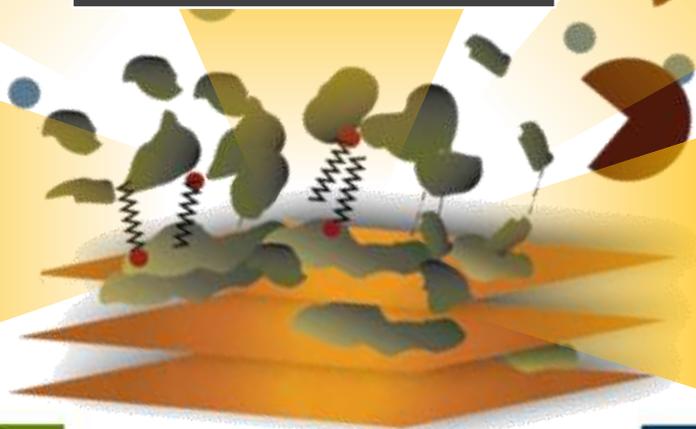


Mössbauer

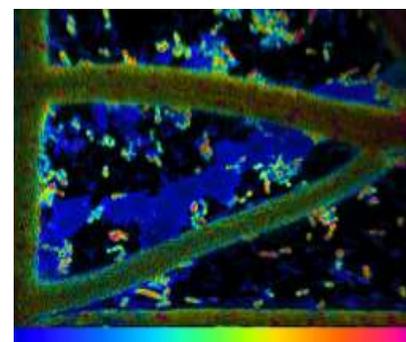
Molecular Structure  
and Dynamics



NMR

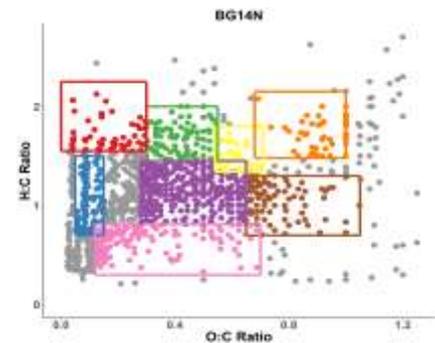


Spatial Mapping



nanoSIMS

Carbon Chemistry



FTICR



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# Joint Genome Institute (JGI)

Nigel J. Mouncey  
Director, DOE JGI



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



The leading Integrative Genome Science User Facility enabling researchers to solve the world's evolving energy and environmental challenges.



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# The Overarching Mission of JGI

Sequence



Functional Understanding



Mission Areas

**Bioenergy** **Carbon Cycling** **Biogeochemistry**

Programs

**Plants** **Fungi** **Metagenomes** **Microbes** **DNA Synthesis Science**



U.S. DEPARTMENT OF **ENERGY**

Office of Science





## DNA Sequencing

- Illumina platforms:  
NovaSeq, HiSeq, MiSeq,  
Next Seq
- PacBio platforms:  
RSII, Sequel
- Oxford Nanopore  
platforms:  
MinION, PromethION



## Advanced Genomic Technologies

- Single-cell Seq
- Metabolomics
- Methylation
- Epigenomics
- Transcriptomics
- DAP-seq, Drop-seq,  
HiC



## DNA Synthesis

- Construct  
Design
- Pathway  
Assembly
- Host  
Engineering



## Computational Analysis

- Assembly/Annotation/Analysis
- High-Performance Computing
- Big Data Integration & Analysis Systems
- Microbiome Data Science

# User Access to the JGI

## Community Science Program



- 50% of JGI capacity
- Large-scale: annual Call for Proposals
- New Investigator: bi-annual call
- Letters of Intent and Proposals Solicited
- Technical Review
- Peer Review
- JGI Capacity Planning
- BER Approval

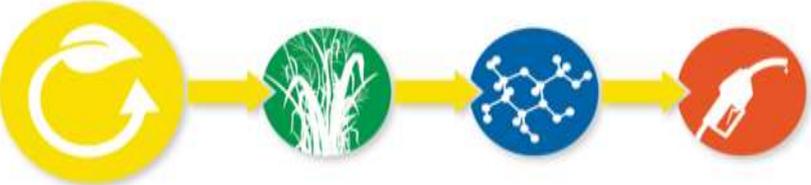
## Facilities for Integrating User Science

- Annual Call for Proposals
- Letters of Intent and Proposals Solicited
- Technical Review
- Peer Review
- JGI Capacity Planning
- BER Approval



## Bioenergy Research Centers

- Direct support at 30% JGI Capacity
- Internal BRC Review
- Technical Review at JGI
- Projects get priority in the queue
- Semi-annual capacity planning
- Monthly project check-ins



## Emerging Technologies Opportunity Program

- Call developed by JLT in collaboration with LBNL Procurement
- Letters of Intent and Proposals Solicited
- Technical/Use Review
- Final decision by JGI Director & JLT



# Outline

- **Facility Overviews**
  - Environmental Molecular Sciences Laboratory (Harvey)
  - Joint Genome Institute (Nigel)
- **The FICUS Initiative**
  - Goals, Program and Coordination (Nigel)
  - Metrics of Success (Harvey)
- **Scientific Impact**
  - Michelle O'Malley (Harvey)
  - Kelly Wrighton (Harvey)
- **Expanded FICUS Activities**
  - JGI, NERSC and KBase (Nigel)
  - EMSL and ARM (Harvey)
- **Future FICUS Directions**
  - Structural Biology (Nigel)
  - Biogeochemistry and Data Interactions (Harvey)



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# FICUS Initiative Purpose and Goals



- Launched 2013 to facilitate access to capabilities at JGI and EMSL (originally JGI-EMSL Collaborative Science Initiative)

- Enable high-impact DOE science combining cutting edge capabilities by reducing administrative barriers

- Expanded 2016 with EMSL-ARM pilot

- Expanded 2017 with JGI-NERSC Microbiome Data Science call supporting computationally intensive data analysis efforts

- In 2018, new calls between EMSL and ARM and JGI, NERSC and KBase initiated



# FICUS Communication and Coordination



- Calls are developed collaboratively with partner facilities and DOE, with input from program and capability leads
- Review panel drawn from both facilities' users / reviewers
- Each facility funds work out of existing user program funding; paylines are jointly decided on
- Monthly coordination calls with partner facilities track individual project progress on a shared google document
- Joint booths at conferences reach new user communities



Terry Law  
EMSL



Susannah Tringe  
JGI



Mary Lipton  
EMSL



Kjersten Fagnan  
JGI



Nikos Kyrpides  
JGI



Emiley Eloie-Fadrosh  
JGI



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



## PI countries – based on submitted Letters of Intent



**PIs from 18 foreign countries:** Australia, Austria, Canada, Chile, Czech Republic, Denmark, Finland, France, Germany, Israel, Mexico, Netherlands, Slovenia, South Africa, Taiwan, United Kingdom

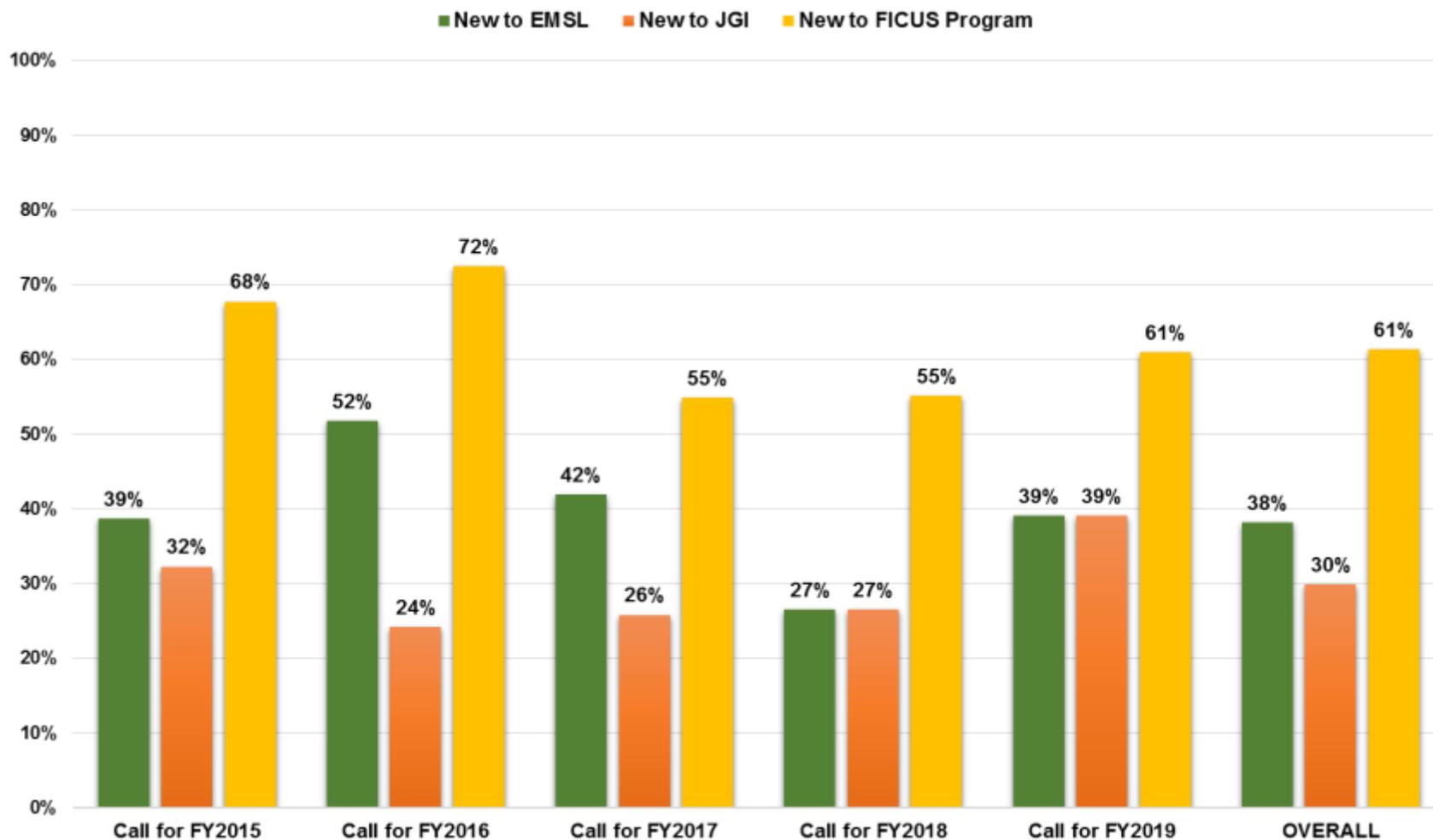


U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# FICUS Attracts New PIs to the Initiative and to Each Facility



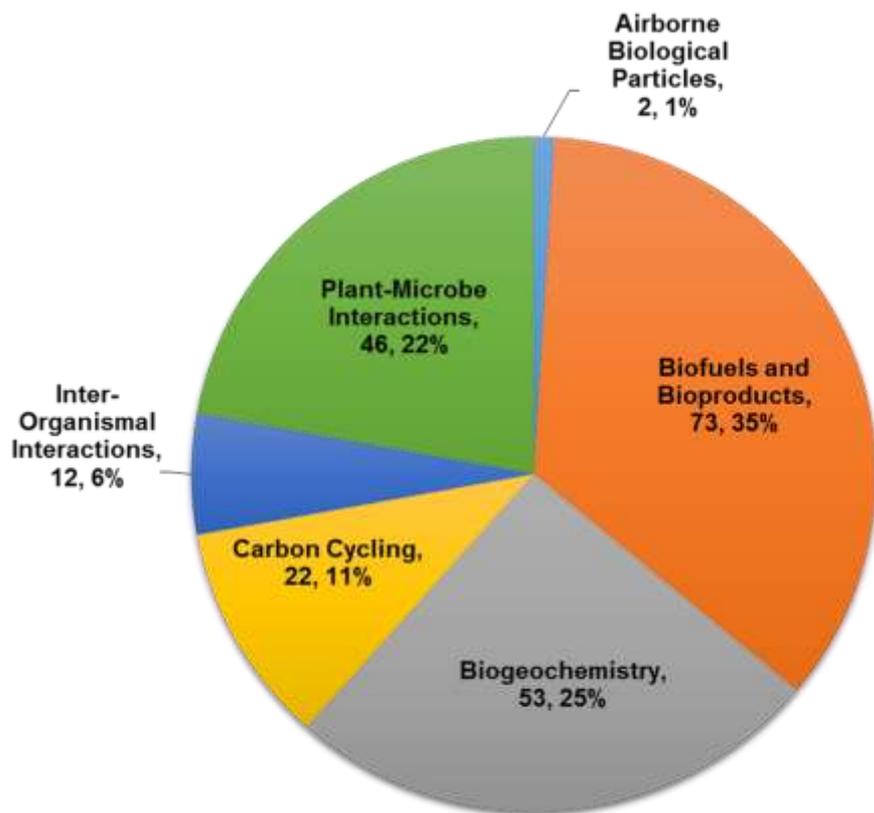
U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science

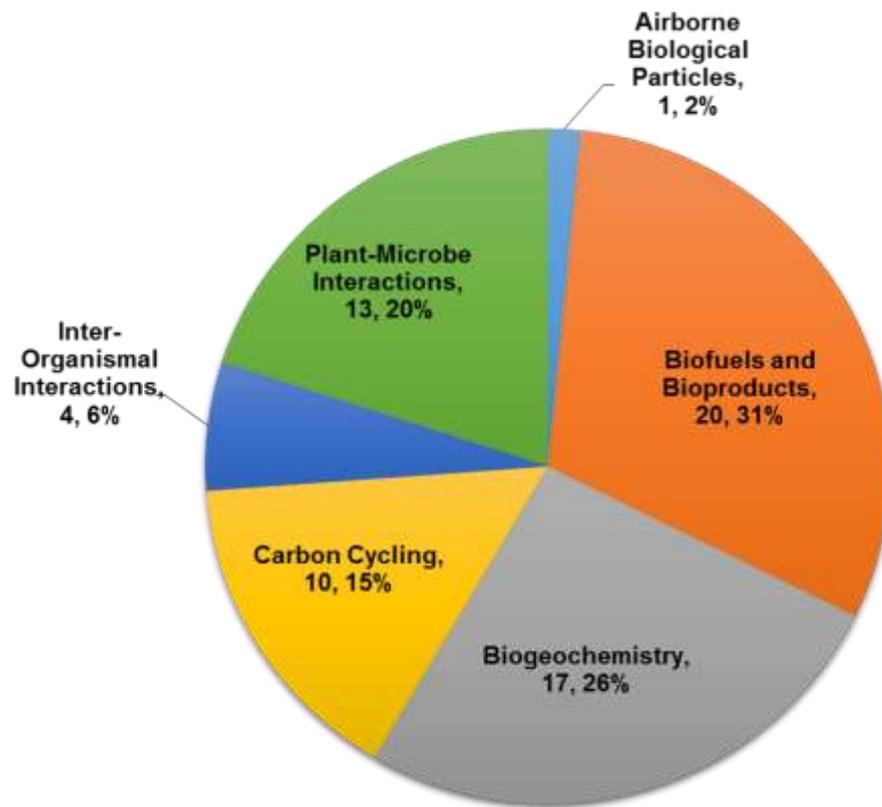


# FICUS Topics Advancing DOE Goals

**Submitted Proposals**  
(208 total)

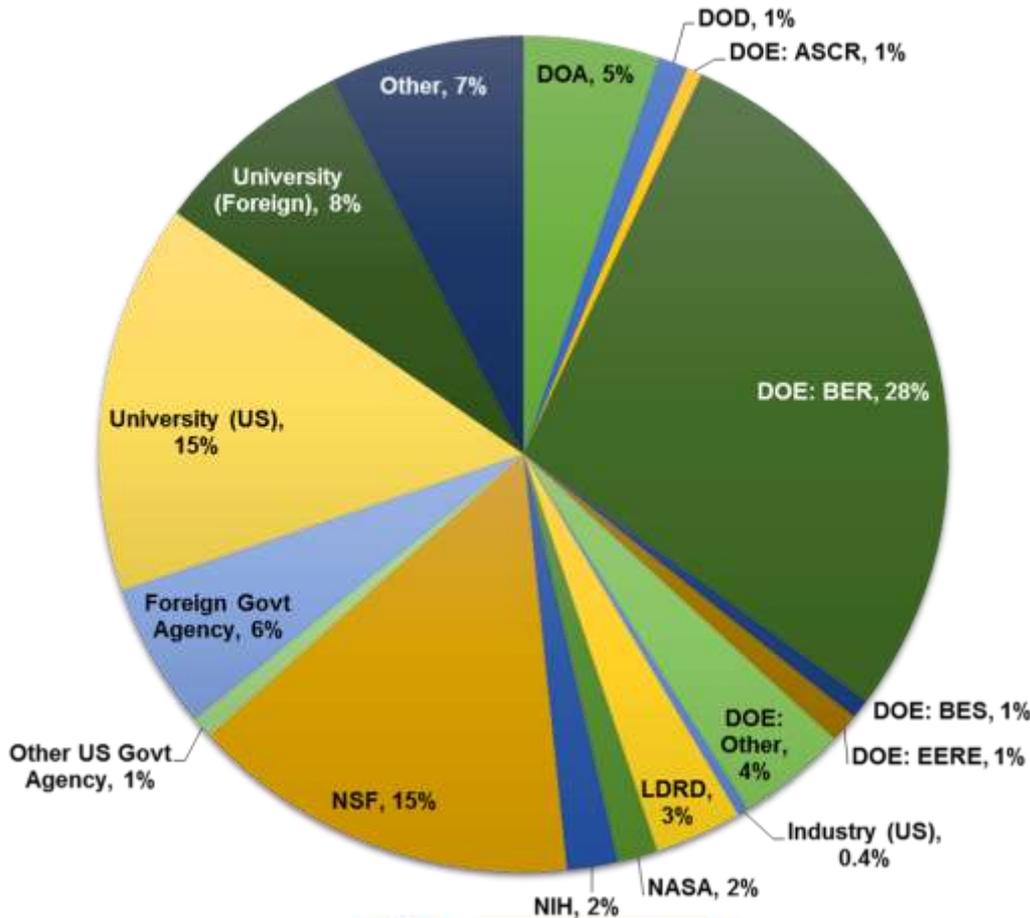


**Accepted Proposals**  
(65 total)

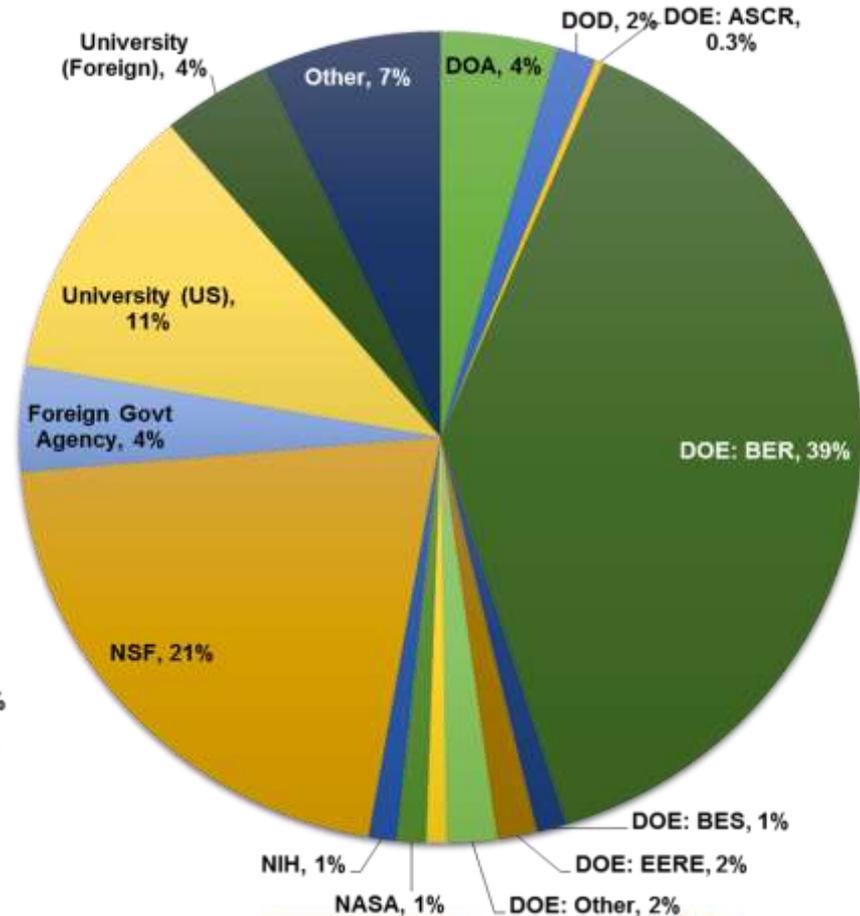


# FICUS Supports Research for Multiple Funding Agencies

## Submitted Proposals



## Accepted Proposals



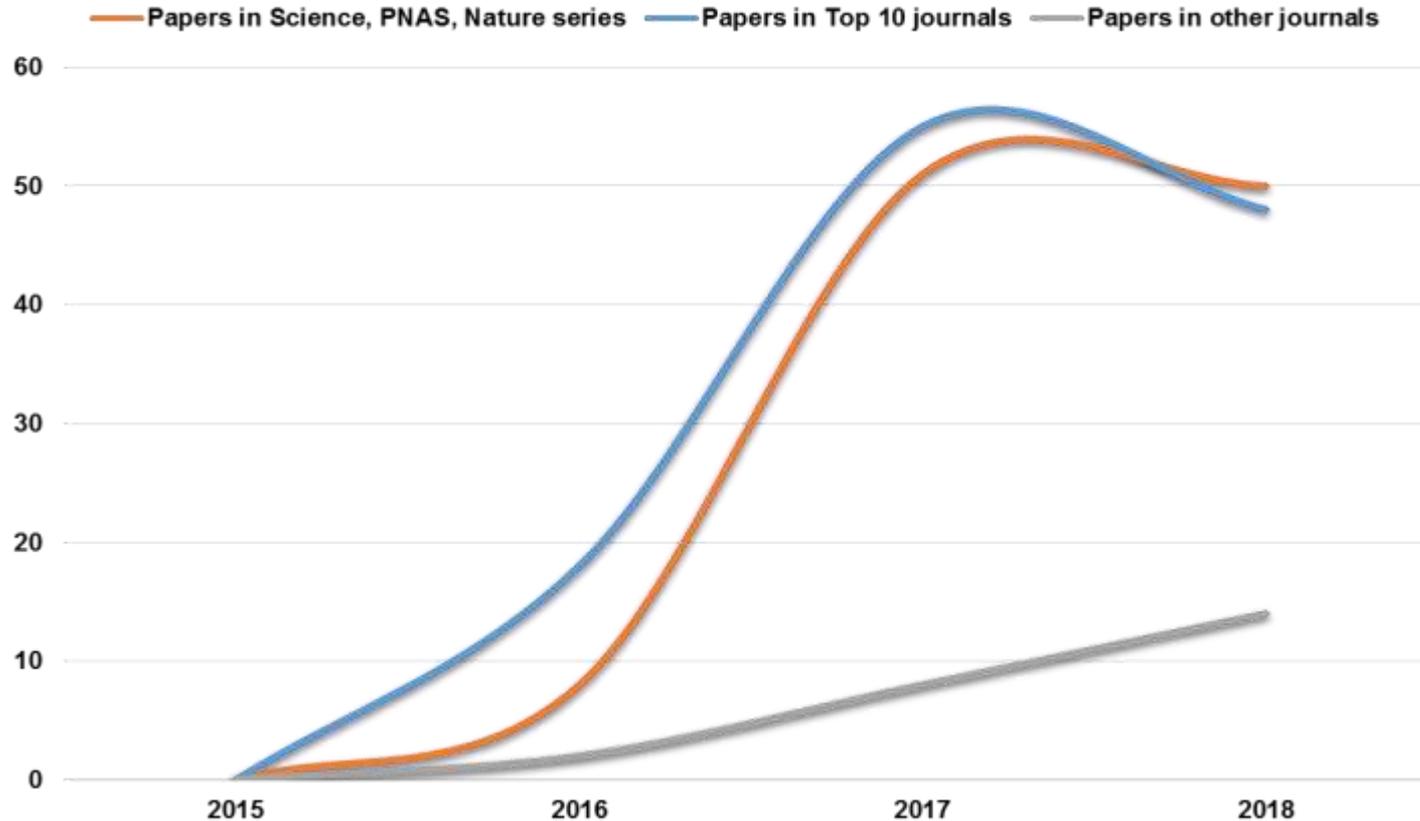
U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# FICUS Research Cited 254 Times, Average Impact Factor of 8.15

# Citations by Year for FICUS Publications



# Pubs:

	2015	2016	2017	2018
Top 3	0	2	2	3
Top 10	0	3	4	1
Other	0	1	4	5



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# Outline

- **Facility Overviews**
  - Environmental Molecular Sciences Laboratory (Harvey)
  - Joint Genome Institute (Nigel)
- **The FICUS Initiative**
  - Goals, Program and Coordination (Nigel)
  - Metrics of Success (Harvey)
- **Scientific Impact**
  - Michelle O'Malley (Harvey)
  - Kelly Wrighton (Harvey)
- **Expanded FICUS Activities**
  - JGI, NERSC and KBase (Nigel)
  - EMSL and ARM (Harvey)
- **Future FICUS Directions**
  - Structural Biology (Nigel)
  - Biogeochemistry and Data Interactions (Harvey)
  - Mechanisms for expanded efforts (Nigel)

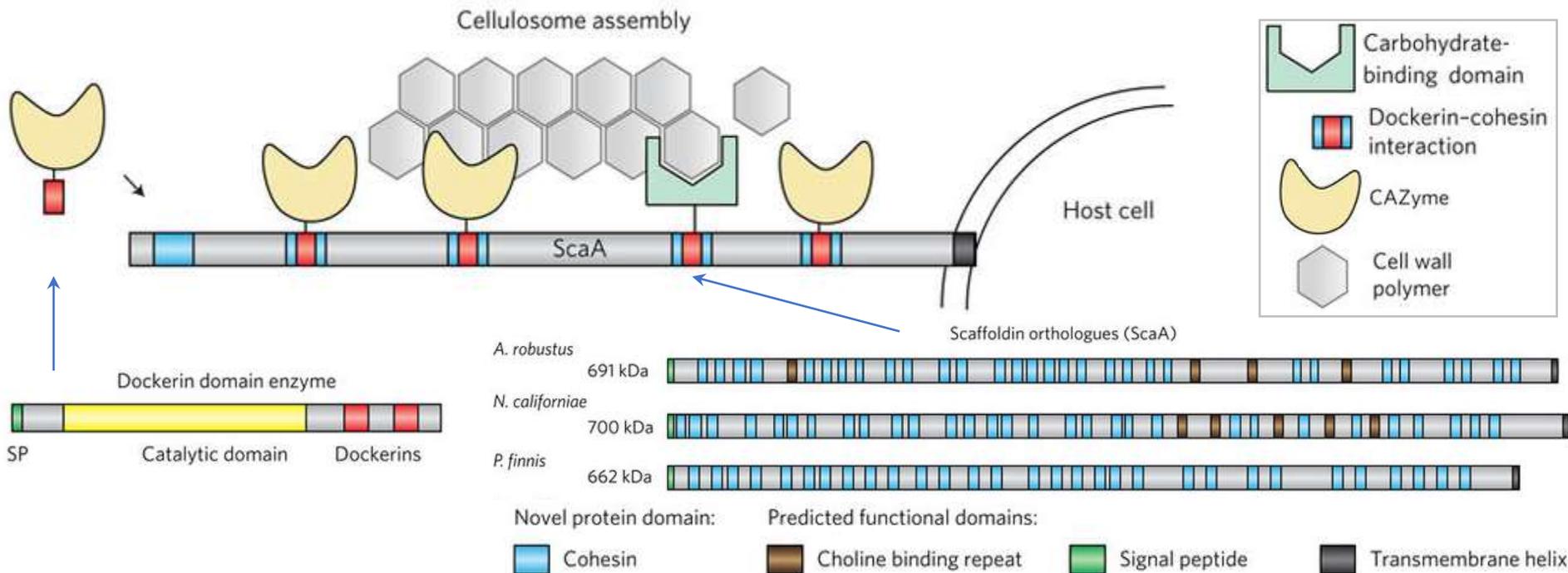


U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# A Cellulosome from Anaerobic Gut Fungi



- Neocallimastigomycota harbor a vast array of CAZymes for biomass breakdown along with unique scaffoldins and dockerins for improving degradation efficiency (JGI)
- Protein expression and interactions validated via proteomics (EMSL)

Solomon *et al.*, Science 2016

Haitjema *et al.*, Nature Microbiol 2017



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# Microbial Interactions in Hydraulically Fractured Shales

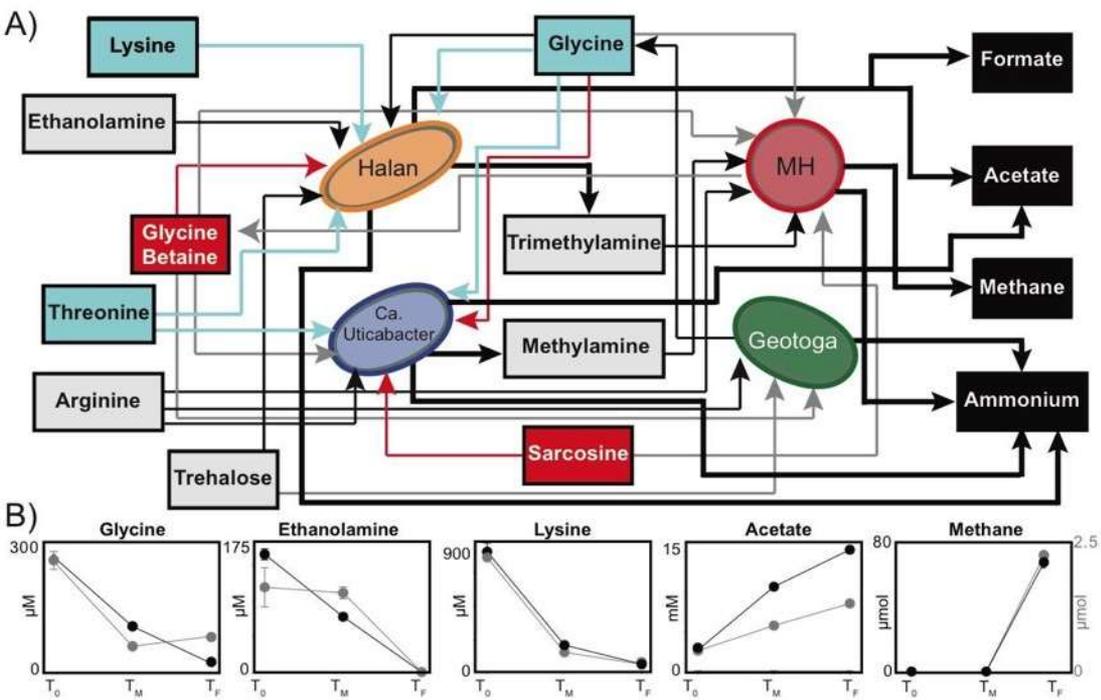


**Coupled laboratory and field investigations resolve microbial interactions that underpin persistence in hydraulically fractured shales**

Mikayla A. Borton, David W. Hoyt, Simon Roux, Rebecca A. Daly, Susan A. Welch, Carrie D. Nicora, Samuel Purvine, Elizabeth K. Eder, Andrea J. Hanson, Julie M. Sheets, David M. Morgan, Richard A. Wolfe, Shikha Sharma, Timothy R. Carr, David R. Cole, Paula J. Mouser, Mary S. Lipton, Michael J. Wilkins, and Kelly C. Wrighton



- Hydraulic fracturing creates a new rock-hosted ecosystem 2,500 meters below the surface
- Here we used JGI and EMSL resources to test our field derived hypothesis that amino acid metabolisms are essential to microbial persistence



- Genome-resolved proteomics network shows the metabolic exchanges and competitions that co-occur in this community
- Metabolite stoichiometry over time supports the metabolisms inferred from proteomics
- This laboratory based model predicted carbon and nitrogen cycling from over 40 field derived metagenome and metabolomes



Kelly Wrighton  
Colorado State University



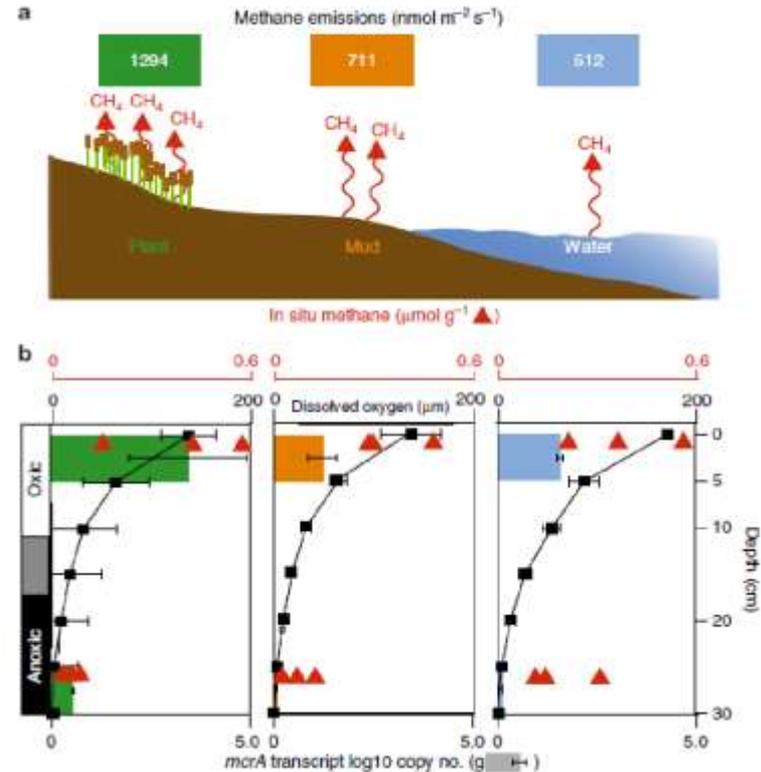
# Aerobic Wetlands Emit High Levels of Methane

## Methanogenesis in oxygenated soils is a substantial fraction of wetland methane emissions

J.C. Angle, T.H. Morin, L.M. Solden, A.B. Narrowe, G.J. Smith, M.A. Borton, C. Rey-Sanchez, R.A. Daly, G. Mirfenderesgi, D.W. Hoyt, W.J. Riley, C.S. Miller, G. Bohrer, and K.C. Wrighton



- Dogma that methanogenesis only occurs in anaerobic soils and sediments
- Here we used EMSL and JGI resources to identify porewater metabolites and DNA and RNA sequencing for novel methanogen species
- This work also links geochemical and biological evidence for methane production in well-oxygenated soils of a freshwater wetland



- Methane production does occur in aerobic soils (colored bars in a)
- Dissolved oxygen in the black boxes, pore water methane by the red triangles and transcripts for mcrA [methyl-coenzyme reductase subunit A] in the colored bars (b).
- Genomes of novel methanogen species *Candidatus* identified
- Results go against paradigm that microbial methanogenesis can only occur in anoxic habitats.

# Outline

- **Facility Overviews**
  - Environmental Molecular Sciences Laboratory (Harvey)
  - Joint Genome Institute (Nigel)
- **The FICUS Initiative**
  - Goals, Program and Coordination (Nigel)
  - Metrics of Success (Harvey)
- **Scientific Impact**
  - Michelle O'Malley (Harvey)
  - Kelly Wrighton (Harvey)
- **Expanded FICUS Activities**
  - JGI, NERSC and KBase (Nigel)
  - EMSL and ARM (Harvey)
- **Future FICUS Directions**
  - Structural Biology (Nigel)
  - Biogeochemistry and Data Interactions (Harvey)



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# JGI-NERSC FICUS



Enabling the scientific community to access large-scale compute capacity and expertise in microbiome data science



Patsy Babbitt  
UCSF



Ed DeLong  
University of Hawaii



Steve Hallam  
UBC

First call Spring 2017:  
6 of 14 proposals approved

- JGI/NERSC monthly coordination meetings to review pilot projects and progress
- Projects completed in FY19Q1



David Baker  
University of Washington



Kostas Konstantinidis  
Georgia Tech



Phillip Brooks  
UC Davis



U.S. DEPARTMENT OF ENERGY

Office of Science



# JGI-NERSC FICUS Hallam Lab

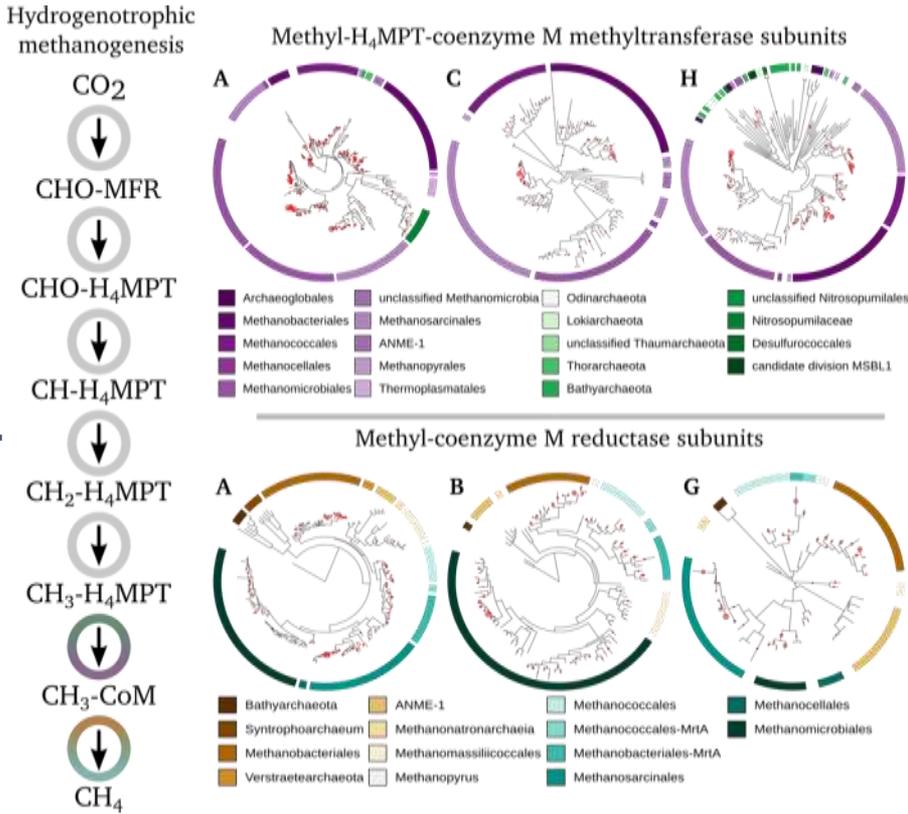


**Goal: Enhance TreeSAPP software for functional annotation of marker genes using phylogenetics**

- TreeSAPP
  - Identifies novel homologues and estimate taxonomic-relatedness
  - Annotates genes with information such as geographic location, substrate specificity, or metabolic life-style association

## Benefits of FICUS Participation

- NERSC access enabled scalability
  - Build, bootstrap and map to reference trees with >10,000 sequences in less than 2 hours
- IMG access allowed us to easily update reference trees with *novel sequences*
- Gained input from a larger panel of experts with diverse backgrounds and interests
- Exploring the possibility of making TreeSAPP available through KBase



Software available at <https://github.com/hallamlab/TreeSAPP>



U.S. DEPARTMENT OF ENERGY

Office of Science

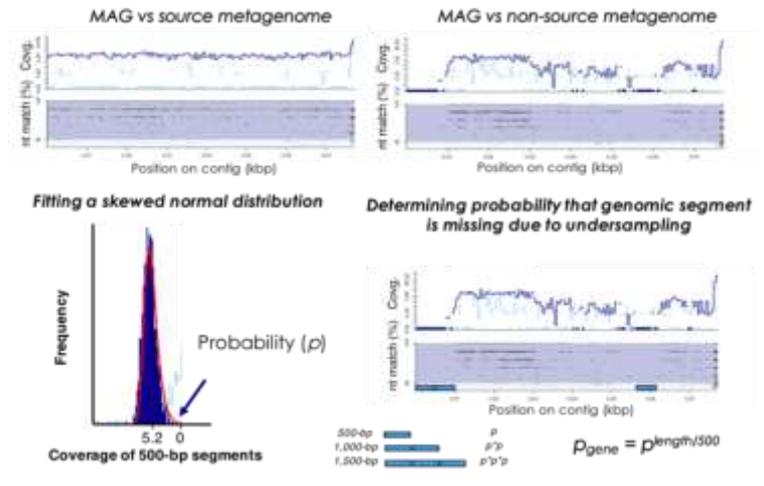


# JGI-NERSC FICUS Konstantinidis



## Assessing microbiomes at the individual population level: tool development and applications to soil carbon cycling.

- Analyzed Metagenome Assembled Genomes (MAGs) from JGI on Cori
  - Discovered some MAGs are widespread (>8,000 km distances) in different soil ecosystems with similar latitudes or habitats
  - Identified potential keystone species that must be interrogated further
- Utilized high-performance computing to process O(1,000) JGI data sets
  - Optimized pipeline for Cori infrastructure including use of the Burst Buffer data resource
- Applied for continued support through the 2019 NERSC Allocation Request Form (ERCAP)



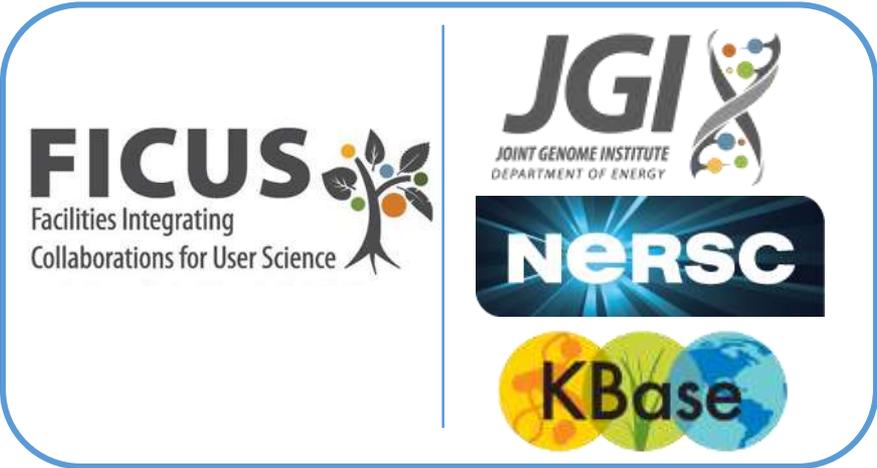
## Assessing gene loss with read recruitment plots



Biogeography and connectivity of select MAGs.



# JGI-KBase-NERSC FICUS Call



FY19 Call went out in September 2018 and responses are under review.

- Supercomputing resources have a steep learning curve that we can mitigate through
  - Training resources (nersc.gov), consulting
  - Access to optimized apps and pipelines through the KBase platform
- Many projects generated data resources that would benefit the broader community
  - JGI and KBase are making these resources available through co-developed infrastructure

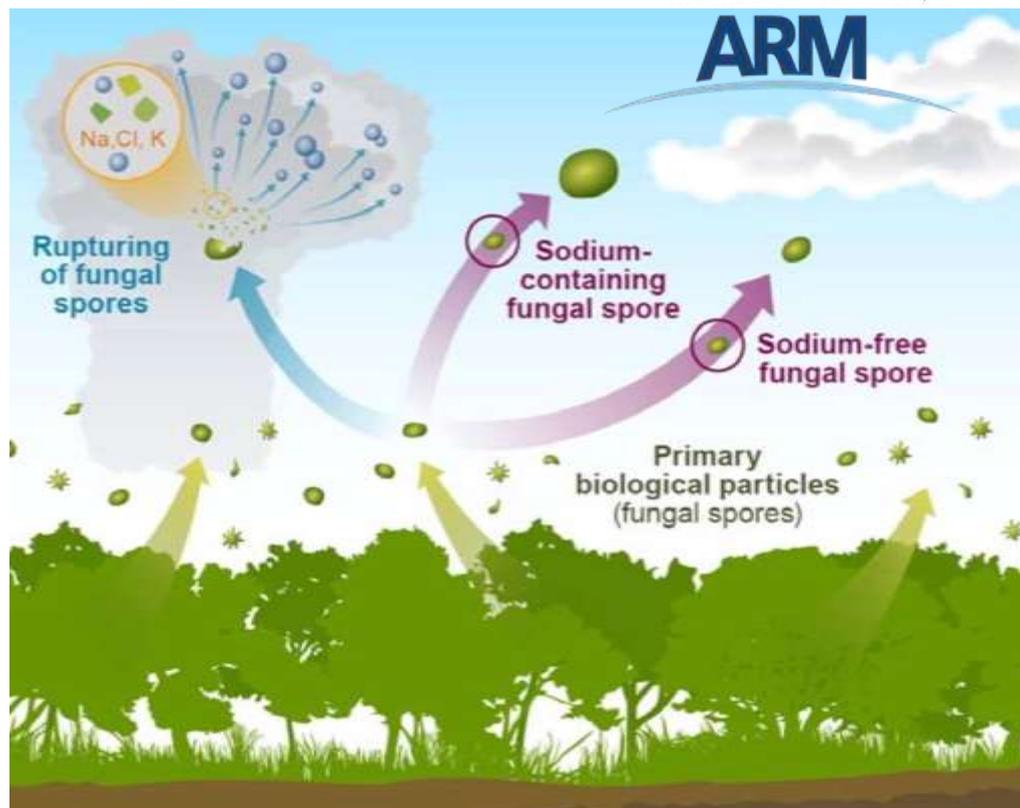
FY19 projects will be coordinated with the NERSC Exascale Application Readiness Program for Data and the Superfacility efforts in Computing Sciences



# EMSL Atmospheric Aerosol Research

Chemistry, physics and molecular-scale dynamics of aerosols to improve accuracy of earth system models

- Had 90 publications with 860 citations (FY2014-FY2017) focused on:
  - Secondary Organic Aerosol (SOA) analysis (55%)
  - Brown carbon (14%)
  - Primary particle analysis (20%)
  - Aerosol modeling (22%)



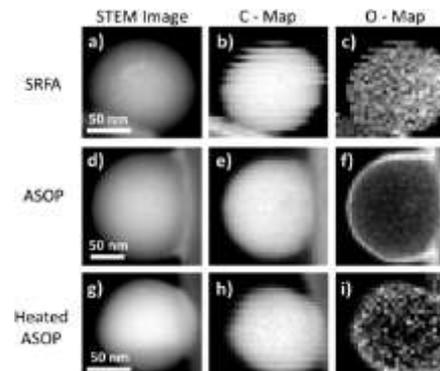
Salt containing biological aerosol previously unrecognized role in cloud formation in Amazon.

**Reference:** China, S., S. Burrows, B. Wang, T. H. Harder, J. Weis, M. Tanarhte, L. V. Rizzo, J. Brito, G. G. Cirino, P. Ma, J. Cliff, M. K. Gilles, and A. Laskin. 2018. "Non marine sources of sodium salt particles in the Amazon basin." Nature Communications, accepted

# EMSL – ARM FICUS FY2016 pilot

ARM pilot focused on collection of samples at the surface or near surface at ARM's Southern Great Plains (SGP) site in Oklahoma with subsequent analysis of those samples at EMSL.

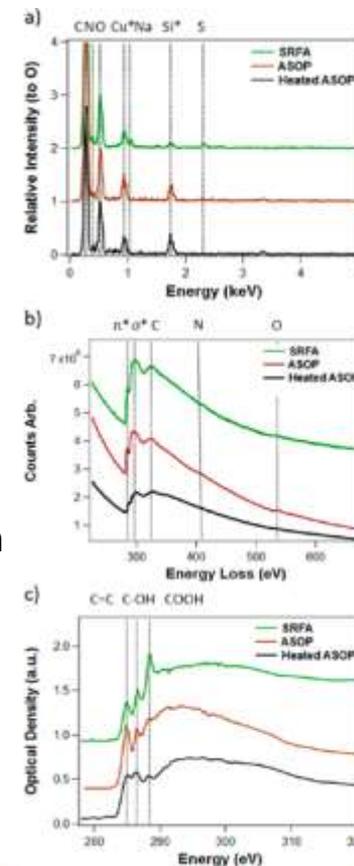
- Limited pilot call
  - Five proposals received, four funded
  - Projects limited in scope, budget, and location
  - Lessons learned - enhance scope and budget, tie to ARM activities versus just ARM site



Optical properties of airborne soil organic particles similar to brown carbon

STEM/EELS – EMSL  
STXM/NEXAFS – ALS, LBNL

**Reference:** D. P. Veghte, S. China, J. Weis, L. Kovarik, M. K. Gilles, and A. Laskin. 2017. "Optical properties of airborne soil organic particles." ACS Earth Space Chem. 1:511-521.



# EMSL – ARM FICUS future



On the horizon:

- EMSL – ARM FICUS call FY2020 - Aerosol particle collection at ARM site and analysis at EMSL
  - Tethered balloon
  - ArcticShark
- Access to EMSL instrumentation and computing



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# Outline

- **Facility Overviews**
  - Environmental Molecular Sciences Laboratory (Harvey)
  - Joint Genome Institute (Nigel)
- **The FICUS Initiative**
  - Goals, Program and Coordination (Nigel)
  - Metrics of Success (Harvey)
- **Scientific Impact**
  - Michelle O'Malley (Harvey)
  - Kelly Wrighton (Harvey)
- **Expanded FICUS Activities**
  - JGI, NERSC and KBase (Nigel)
  - EMSL and ARM (Harvey)
- **Future FICUS Directions**
  - Structural Biology (Nigel)
  - Biogeochemistry and Data Interactions (Harvey)



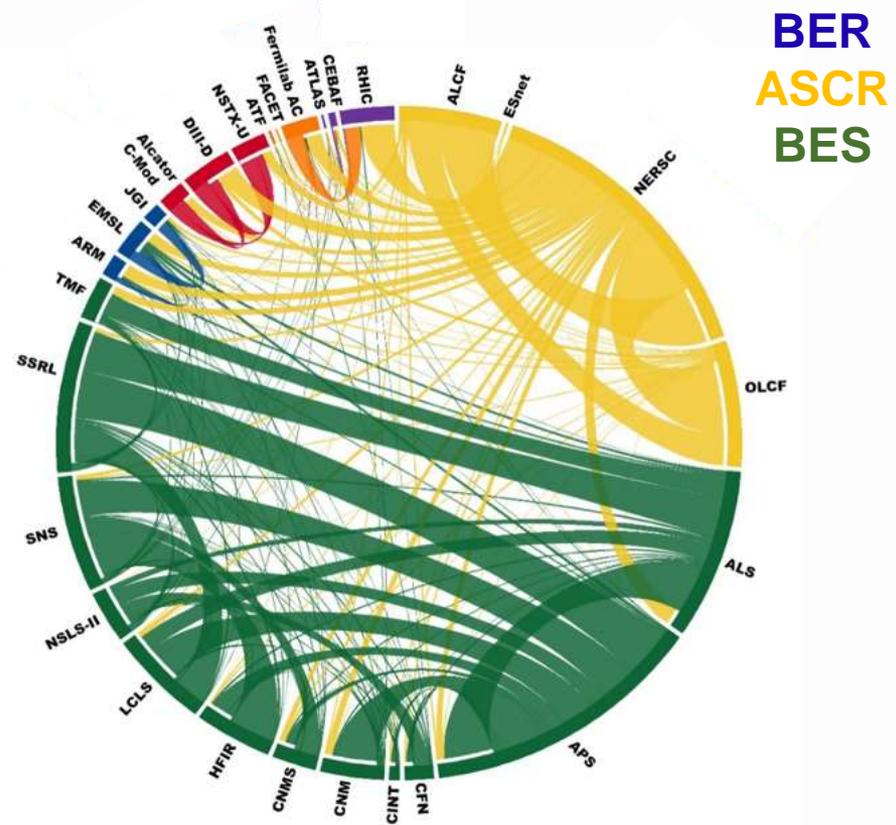
U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



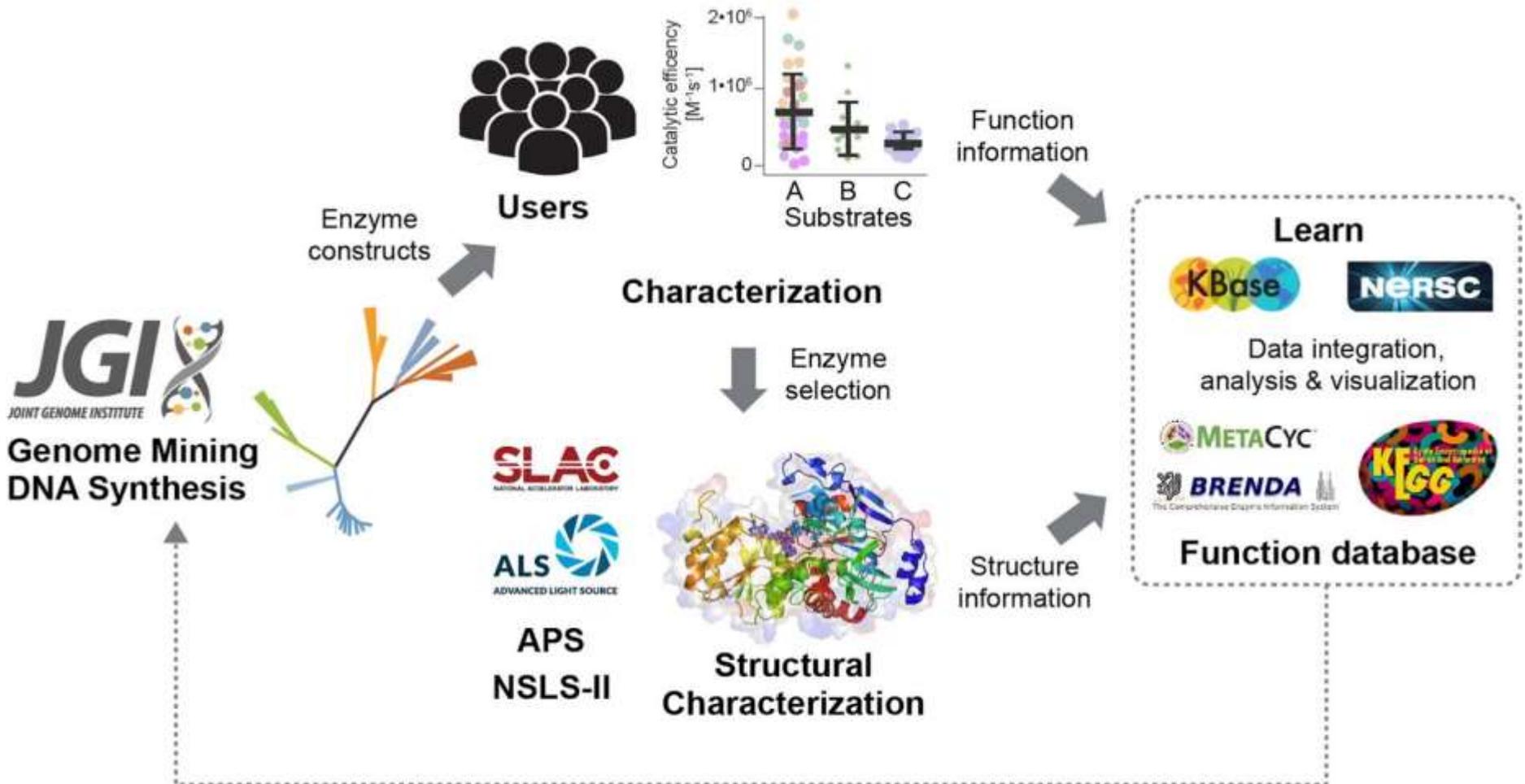
# Expanding Cross-User Facility Users

- **Already, many users work across BES and ASCR User Facilities**
- **How to expand across all User Facilities?**
  - July 2016 meeting with representatives of 8 user facilities exchanged ideas, made plans to collect data on multi-facility users and user interest
  - March 2018 User Facility meeting of JGI, EMSL, NERSC, SLAC and KBase staff, served as basis for forming Cross-Facility Working Group
  - August 2018: EMSL Integration Meeting



FY16 user crossover data  
Ben Brown, DOE

# Genomes to Structure: Cross BER and BES User Facility Initiative



Improve JGI design pipeline for enzymes and pathways



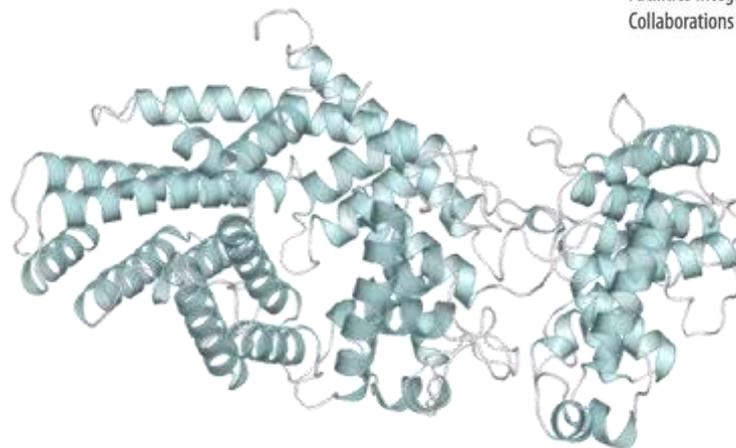
U.S. DEPARTMENT OF **ENERGY**

Office of Science



# Pilot Structural Biology Projects between JGI and ALS through CSP

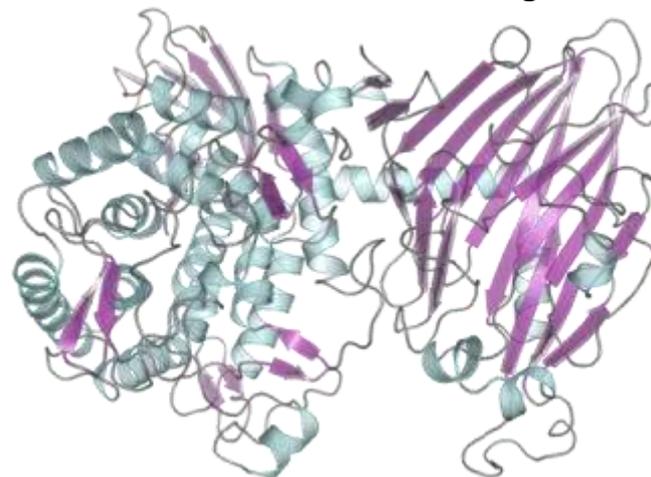
- Potential collaborators identified from JGI CSP
- Candidate proteins selected in collaboration with JGI user
- Genes in appropriate vectors supplied by JGI synthesis group
- Proteins expressed, purified and crystallized by LBNL researchers
- Data collected at the Advanced Light Source (through general user time)
- Structural interpretation and follow up experiments in collaboration with JGI user



Crystal structure of a terpene synthase solved at 2.1 Å resolution using BL 8.2.1



Philipp Zerbe,  
UC Davis



Glycoside phosphorylase crystal structure solved at 1.1 Å resolution using BL 8.2.2

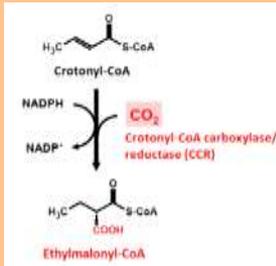


Stephen Withers,  
UBC

# Pilot Projects between JGI and SSRL through CSP Mechanism



**CSP1755: Mining novel CO<sub>2</sub>-fixation biology in the genomes of microbes using synthetic biology**

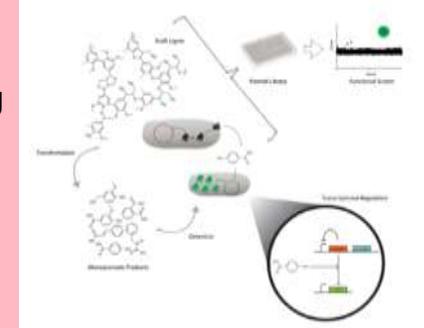


Tobias Erb, MPI

**CSP2579: Towards whole cell biocatalysts for lignocellulose processing**



Beth Sattely, Stanford



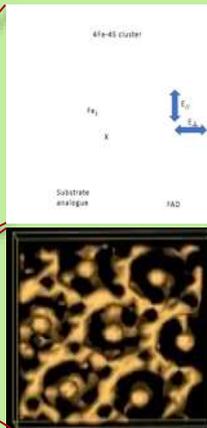
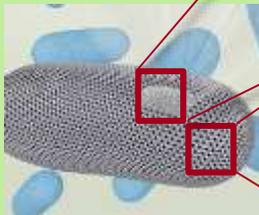
Metalloenzymes:

- Laccase
- Peroxidase

**CSP973: Metagenomic and biogeochemical characterization of marine ammonia-oxidizing archaeal communities in a coastal upwelling system**



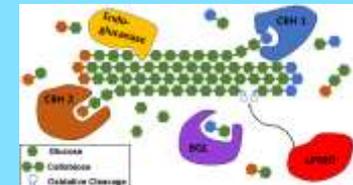
Chris Francis, Stanford



**CSP 719: Synthesis of CAZymes from anaerobic fungi involved in plant biomass degradation in the rumen**



Matthias Hess, UC Davis, and in collaboration with A. Joachimiak (SBC, APS)



Office of Science

Office of Science



# EMSL Integration 2018 - Molecular Structure and Dynamics in Biology and the Environment



- Focus on the structure and dynamics of molecules important in biology and the environment using advanced techniques available at Department of Energy user facilities.
- Science sessions with plenary presentations, invited talks, user facility flash talks, poster session, tutorials from multiple facilities (EMSL, KBase, SSRL, NSLS-II)
- Brainstorming session identified key attributes of expanded FICUS program
  - Common website describing capabilities, example use cases
  - Joint proposal/review process
  - Versatile sample holders
  - Standard data/metadata formats



Molecular Structure and Dynamics in  
Biology and the Environment



Spallation Neutron Source



U.S. DEPARTMENT OF  
**ENERGY**

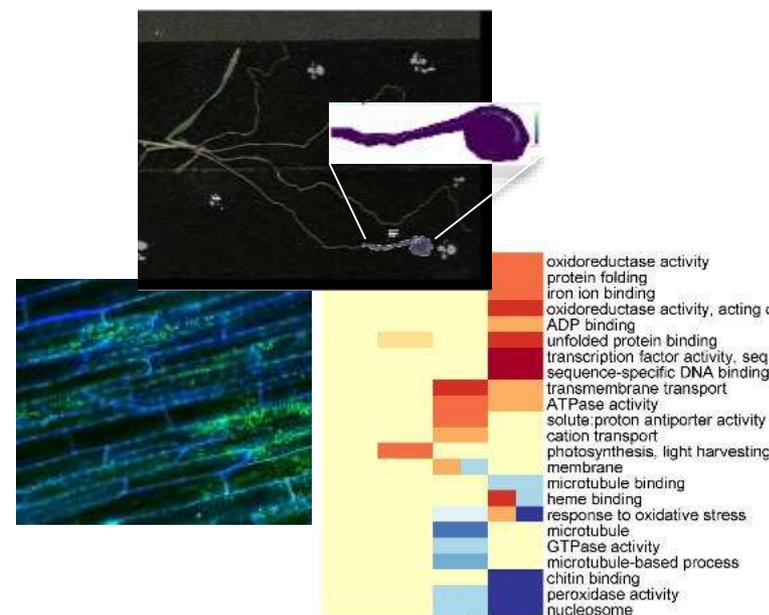
Office of  
Science



# Proposed FY20 FICUS Call with EMSL, JGI and NSLS-II

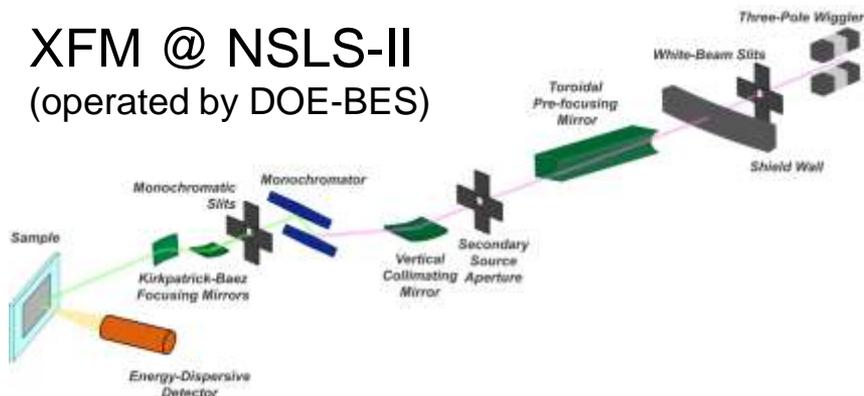


- Upcoming FICUS would include rhizosphere focus area with *bioimaging pilot test* for NSLS-II's X-ray Fluorescence Microscope (XFM)
- Partner User Agreement with NSLS-II being explored
- Anticipate 1-2 projects selected



Pursue a holistic view of nutrient acquisition and exchange in plant-soil-microbe systems by characterizing genomic, metabolic and physical characteristics across scales

## XFM @ NSLS-II (operated by DOE-BES)

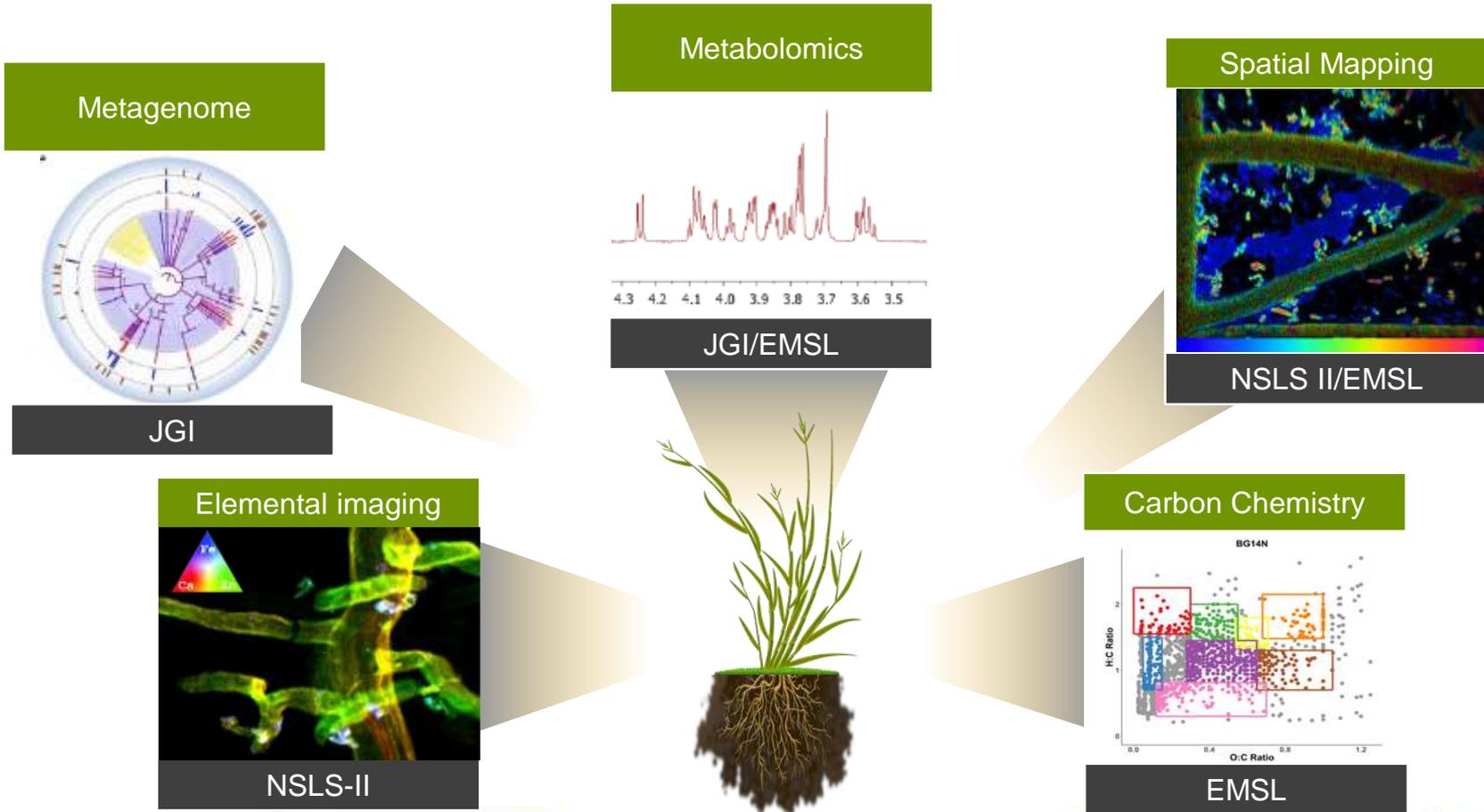


U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# Understanding the influence of micronutrients in the rhizosphere interactions



# EMSL Data Interactions – MyEMSL, ESS-Dive and KBase

## ➤ MyEMSL – Current and Future

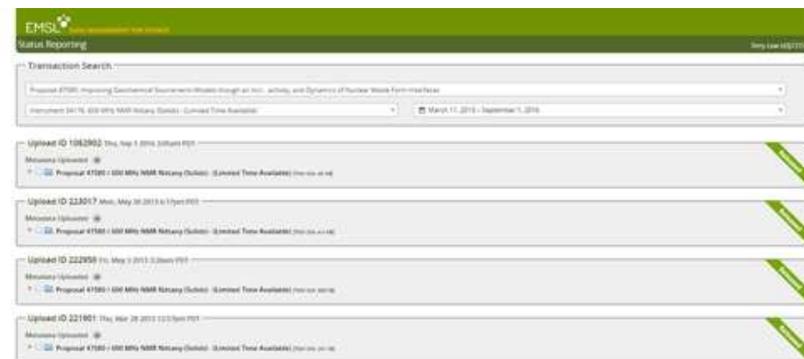
- ▶ Users can now access their data directly through the User Portal
- ▶ Public search and retrieval, DOI minting coming in FY19

## ➤ MyEMSL and ESS-Dive

- ▶ WHONDRS network as a use case for dedicated project spaces and data linkages to MyEMSL
- ▶ ESS-Dive to provide backend interface to automate data uploading
- ▶ Aspiration: Linked search without physically moving data

## ➤ MyEMSL and KBase

- ▶ Implementing EMSL developed MFAPipe fluxomics software as KBase app
- ▶ Discussing potential linkages to NWChem



# Thanks

- BER and BES Program Managers
- DOE SC Leadership
- EMSL and JGI Staff and Management
- Our Users
- Terry Law and Susannah Tringe



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science

