
JGI
JOINT GENOME INSTITUTE
A DOE OFFICE OF SCIENCE USER FACILITY
2016
Progress Report
U.S. Department of Energy
Joint Genome Institute

Cover photo: Mount Diablo at sunrise, Contra Costa County, California. *(Mark Lilly Photography)*

Table of Contents

DOE JGI
Mission

Golden Gate Bridge, San Francisco, California before dawn. *(Peter Burnett)*

The mission of the U.S. Department of Energy Joint Genome Institute (DOE JGI) is to serve the diverse scientific community as a national user facility, enabling the application of large-scale genomics and analysis of plants, microbes, and communities of microbes to address the DOE mission goals of harnessing science and technology to address energy and environmental challenges.

Axel Visel *(Marilyn Chung, Berkeley Lab)*

Director's **Perspective**

Forging Ahead With a Joint Vision

In 2016, the DOE Joint Genome Institute continued on its trajectory of evolving as a Next-Generation Genome Science User Facility that offers its large user community access to cutting-edge genomics capabilities. A continuing trend across our scientific programs is an increased emphasis on approaches that go beyond the assembly of reference genomes. DOE JGI users can now leverage a growing portfolio of advanced experimental and computational techniques to understand the function of genes and genomes, turning sequence data into biological insights.

In March 2016, Eddy Rubin stepped down from his position as the Director of the DOE JGI after 14 years. As the longest-serving Director in the history of the DOE JGI, Dr. Rubin guided its scientific directions from completing DOE's contributions to the Human Genome Project to transforming the Institute into a national user facility enabling the science of thousands of researchers focused on energy and environmental challenges. His leadership and scientific vision were critical in shaping the DOE JGI into the Next-Generation Genome Science User Facility that it is today, with a growing range of sequence-related capabilities including advanced functional genomics, single-cell genomics, DNA synthesis, computational, and metabolomics approaches.

Taking over in April 2016 as the DOE JGI Interim Director while the search for a new permanent Director was under way, I continued to guide the DOE JGI on this successful path. While the DOE JGI user programs and scientific activities kept running on all cylinders, we further refined our strategic directions. In our 2016 Strategic Planning Update we described more than 100 specific goals in support of our vision, to be accomplished over the next two to five years. Through focused internal technology development, new external partnerships, new user program focus areas, and activities aimed at engaging new user communities, the realization of our long-term vision for the DOE JGI as a Next-Generation Genome Science User Facility has gained further momentum throughout 2016.

Our users are critical to our success. In fiscal year 2016, the DOE JGI served 1,391 users from academia, government, and industry from 43 countries across the globe. In support of their projects, we produced nine plant, 186 fungal, and 1,895 prokaryotic *de novo*-assembled genomes, as well as 1,009 metagenomic data sets. Consistent with an increased emphasis on functional genomics, we produced 550 transcriptomics and metatranscriptomics datasets. As in past years, a much larger swath of the scientific community made use of the DOE JGI's Genome Portals and the comparative genomics platforms Phytozome, MycoCosm, and Integrated Microbial Genomes (IMG), which together drew 800,000 visitors for data access and analysis.

Demand for our user programs is ever increasing. In 2016, we received 96 proposals submitted to the large-scale Community Science Program (CSP), as well as 111 proposals in response to the small-scale microbial/metagenome and DNA synthesis calls, and 29 proposals to the joint Facilities Integrating Collaborations for User Science (FICUS) call, a user program collaboratively offered by the DOE JGI and the Environmental Molecular Sciences Laboratory (EMSL) at Pacific Northwest National Laboratory. Across all of these programs, a total of 66 proposals were approved and accepted after peer review, and are now in progress.

The impact of the research performed through these proposals is reflected by the large number of scientific publications originating from the DOE JGI and its users. In 2016, the DOE JGI contributed to 167 peer-reviewed publications, including six in the most selective journals, *Science* and *Nature*. Six of our scientists were included in the annual "Highly Cited Researchers" list by Clarivate Analytics (formerly Thomson Reuters). Selected scientific highlights published in 2016 included:

- **Biomass-Converting Capabilities of Anaerobic Gut Fungi.** In the first study resulting from the FICUS user program, UC Santa Barbara researchers reported that anaerobic gut fungi isolated from ruminants perform as well as the best fungi engineered by industry in their ability to convert plant material into sugars that are easily transformed into fuel and other products. (*Science*, February 18, 2016)

- **Salt Tolerance Adaptations by Marine Flowering Plants.** Coastal seagrass ecosystems account for an estimated 15 percent of carbon fixed in oceans, and also affect sulfur and nitrogen cycles. A team including DOE JGI researchers sequenced the genome of a seagrass, *Zostera marina*, providing insight into the adaptations of these plants to saltwater environments. (*Nature*, February 18, 2016)

- **Plumbing Earth's Viral Diversity.** Viruses are increasingly recognized to affect environmental processes mediated by microbes, but there are currently less than 2,200 sequenced DNA virus genomes, compared with approximately 50,000 bacterial genomes. DOE JGI researchers assembled metagenomic datasets from around the world to uncover over 125,000 partial and complete viral genomes, the majority associated with microbial hosts. This single effort increased the number of known viral genes by a factor of 16, and provides researchers with a unique resource of viral sequence information. (*Nature*, August 17, 2016)

- **Hydraulic Fracturing Impacts on Microbial Communities.** To answer the question of how subsurface microbial communities are affected by extracting natural gas from shale formations, a team including DOE JGI researchers reconstructed 31 microbial genomes from fractured shale and studied their metabolic interactions for nearly a full year, revealing dramatic changes in the community over that time. The work, led by Ohio State University researchers, was accomplished in part with the DOE JGI and EMSL through the FICUS program. The study appeared online and was selected as one of *Discover* magazine's Top 100 Science Stories of 2016. (*Nature Microbiology*, September 6, 2016)

- **A Novel Bacterial Phylum.** By analyzing more than five trillion basepairs of metagenomic sequence information, a team led by DOE JGI researchers uncovered a novel bacterial phylum dubbed "Kryptonia." The work highlights the utility of metagenomic data analysis in the exploration of taxonomic "blind spots" missed by conventional primer-based ribosomal gene surveys. (*Nature Communications*, January 27, 2016)

- **Uncovering More Microbial Lineages.** With help from the National Energy Research Scientific Computing Center (NERSC), the team analyzed more than 50,000 gene sequences that are found in every microbe. The sequences can serve as an identifying marker from metagenomic datasets. The PCR primers most commonly used in these kinds of analyses can miss as much as 10 percent of those gene sequences, the team found. Identifying biases in current sequencing technologies could help researchers find methods for describing additional phylogenetic lineages. (*Nature Microbiology,* February 1, 2016)

- **Redefining SAR11's Ecological Niche.** Though SAR11 bacteria are estimated to make up as much as half of the total microbial community in the ocean's oxygen-rich surface layers, they have also been found to be abundant in oxygen minimum zones (OMZs). Through a series of single-cell sorting and synthetic biology experiments, DOE JGI researchers helped to demonstrate how SAR11 cells adapted to these anoxic environments by acquiring the biochemical machinery to respire nitrate, in effect initiating oceanic nitrogen loss. (*Nature*, August 3, 2016)

- **Cassava's Genetic Diversity.** Cassava roots represent a strategic source of renewable energy. They contain 20-40 percent starch that costs 15-30 percent less to produce per hectare than starch from corn. To help improve breeding strategies for cassava, a team led by researchers from the University of California, Berkeley and including DOE JGI staff analyzed the genomes of wild and cultivated cassava, revealing extensive interspecific hybridization and genetic diversity. (*Nature Biotechnology*, April 18, 2016)

- **A Technique for Identifying Active Microbes.** Researchers at the California Institute of Technology (Caltech) and the DOE JGI utilized a recently refined technique known as Bio-Orthogonal Non-Canonical Amino Acid Tagging (BONCAT) to identify individual active cells, as well as clusters of active bacteria and archaea within microbial communities collected from deep-sea methane seep sediments. Using BONCAT, the team demonstrated that methane-oxidizing archaea and sulfate-reducing bacteria live together symbiotically in these environments and help to remove some 80 percent of the methane released from ocean sediments. (*Proceedings of the National Academy of Sciences*, June 28, 2016)

- **Expanding the Suite of Workhorse Yeasts.** To help boost the use of a wider range of yeasts and to explore the use of genes and pathways encoded in their genomes, a team led by DOE JGI researchers conducted a comparative genomic analysis of 29 yeasts, including 16 whose genomes were newly sequenced and annotated. (*Proceedings of the National Academy of Sciences*, August 15, 2016)

- **A Fungus That Helps Its Host Resist Drought Stress.** *Cenococcum geophilum* is the only known mycorrhizal fungus in the Dothideomycetes, a large class comprised of some 19,000 fungal species, many of them plant pathogens. A team including DOE JGI researchers sequenced and compared its genome with those of close relatives with different lifestyles, revealing *C. geophilum* adaptations through which it could help its hosts be more resistant to drought stress. (*Nature Communications*, September 7, 2016)

To enable these types of exciting and impactful scientific discoveries in the future, a major continued thrust at the DOE JGI is providing users with access to state of the art genomic technologies and highly skilled staff. During the past year, we have adopted even higher throughput single-molecule long-read sequencing platforms and continued to explore increasingly robust nanopore technologies. Through these efforts, *de novo* assemblies of genomes continue to improve in quality, enabling our users to probe the organization and content of genomes with increasing ease and efficiency.

In addition, as our ability to rapidly sequence and assemble high-quality reference genomes matures, the next big challenge we continue to face is the interpretation of the information encoded within them. To bridge this sequence-to-function gap, we are focused on developing new genomic and epigenomic tools to support members of our user community in their efforts to characterize the genes and gene networks underlying important functions relevant to energy and the environment. Building upon our previous success in introducing cutting-edge epigenomic tools, we continue to develop methods to investigate the *in vivo* properties of DNA, such as chemical modifications of individual bases and three-dimensional packaging.

As we will ultimately need to understand how gene networks function at the single-cell level, we are also working to leverage the JGI's existing expertise in single-cell microbial analysis to develop single-cell genomic and epigenomic techniques for multicellular plants and algae. By annotating the functional roles of gene networks at the single-cell level, we can begin to better interpret how natural genomic diversity influences traits important to DOE mission science.

Our investment in metabolomics capabilities over the past few years has now developed to where we routinely offer these techniques to users, and a first set of user projects is well under way. It is anticipated that linking genomes and genes to orthogonal metabolomics datasets will provide critical insights linking sequence to organism function. Likewise, our DNA Synthesis Science Program continues to expand year to year in DNA construct output, external user access, and scientific impact, which included the first *Nature* and *Science* publications supported by the program in 2016. Additional focus areas in 2016 were the development and implementation of new software tools to support pathway design, as well as strain engineering to expand the portfolio of hosts available for DNA synthesis studies.

To further broaden our technological portfolio beyond in-house development efforts, we launched three new Emerging Technologies Opportunity Program (ETOP) projects in collaboration with external partners in 2016, all aimed at deeper characterization of uncultivated organisms in the environment. One will develop a pipeline for targeting functionally active populations within a microbial community using stable isotope probing combined with metagenomics and metatranscriptomics. Another will enable the physical isolation and sequencing of cells from specific phylogenetic groups of interest, and the third will enhance the sensitivity of flow cytometers to enable improved analysis and sorting of single microbial cells. In 2016, we also saw earlier ETOP projects come to fruition in providing capabilities in functional targeting of cells and populations for sequencing, microfluidic library generation, informatic binning of genomes from metagenomes, and supplying critical nucleic acids for eukaryotic genome and transcriptome sequencing.

Beyond our user programs, the DOE JGI engages with its user community through a variety of meetings and workshops. The largest of these meetings, the 11th Annual Genomics of Energy and the Environment Meeting drew 497 users, prospective users, and interested scientists, with registrations filling to capacity weeks in advance. An exciting lineup of speakers included Dianne Newman of Caltech, who shortly after became a MacArthur Fellow; Hamilton Smith of JCVI, a Nobel Laureate; and Jeanine Olsen, PI on the *Zostera marina* (seagrass) genome project, later named one of *Nature's* "Editor's Choice" articles for 2016. A summary of the meeting can be found on page 62. In fiscal year 2016, the DOE JGI also co-hosted a workshop, "Exploring Diversity of Life," and hosted a meeting on "Microbial and Plant Systems Modulated by Secondary Metabolites," as well as two microbial genomics and metagenomics workshops.

Diversity and inclusion continued to be major institutional guiding principles at the DOE JGI over the past year. Formed in April 2015, our member-supported Diversity and Inclusion Working Group encompasses approximately a dozen passionate DOE JGI employees dedicated to increasing organizational awareness and proactively creating a more diverse and inclusive work environment. To name a few activities, in 2016 the working group developed formal hiring manager processes with human resources to ensure diversity in recruitments and helped facilitate summer internships for undergraduate minority students to cultivate the next generation of genome scientists. Other organizational accomplishments included substantially increased participation of women and minorities in scientific review panels (>45% female), across all DOE JGI scientific advisory boards (>40% female) and as speakers at our Annual User Meeting (>50% female) and at DOE JGI workshops (>40% female).

In 2016, we continued the successful FICUS initiative, through which users can apply — with a single proposal — for access to capabilities located at the DOE JGI, as well as the Environmental Molecular Sciences Laboratory (EMSL), a DOE user facility at Pacific Northwest National Laboratory. This year, the first set of studies from this user program established in 2014 was published, including a study on the microbes found in fracking wells that was named one of *Discover* magazine's top 100 discoveries of 2016, as well as a report in *Science* about biomass-degrading gut fungi (see Science Highlights above). In 2016, 10 new proposals were accepted through our collaborative FICUS call, pursuing new combinations of omics and molecular analysis capabilities across the DOE JGI and EMSL. Beyond this existing call, we laid the groundwork for new collaborative FICUS efforts with other DOE User Facilities, including a FICUS call for microbiome data analysis jointly offered with NERSC to be launched in January 2017.

Over the last years, the DOE Systems Biology Knowledgebase (KBase) has become an increasingly important partner for the DOE JGI, and we further expanded this relationship in 2016. Enabling researchers to collaboratively develop, share, and apply advanced tools for the analysis of biological function, the KBase platform lowers the bar for the integration of diverse biological data and results, often including genomic data sets produced by the DOE JGI. Both the DOE JGI and KBase aim to enable seamless interactions between their respective resources, with the goal of providing JGI users with access to KBase tools and data that go beyond what is offered through DOE JGI resources. It is already possible to directly transfer several types of datasets from DOE JGI data portals into KBase for further analysis with the single click of a "Push-To-KBase" button. With the planned colocation of KBase and DOE JGI, further integration of resources is particularly important and has been a focus of joint coordination activities in 2016.

The DOE JGI has generated petabytes of high-quality sequence data and analysis. In order to support this workload, we have invested significant resources in our high-performance compute cluster, Genepool, as well as its storage and web infrastructure. The success of our computing projects is in part due to our ongoing partnership with NERSC, one of the nation's foremost centers for high-performance computing. In 2010 all of the DOE JGI's computational resources were moved to NERSC, and both sides have learned a great deal through this partnership. The infrastructure advancements to Genepool and other DOE JGI portals mean rapid and smooth access for users across the globe. Our partnership with NERSC enables DOE JGI researchers and users to devote more of their time to cutting-edge genomics research. In 2016, the DOE JGI used 15 million central processing unit hours on NERSC's petascale supercomputers, Edison and Cori. Many of these calculations could not have been completed on the Genepool cluster. NERSC also deployed new software, called Shifter, to allow Docker containers to run on its supercomputers, a work that grew out of a collaboration with the DOE JGI staff. In 2015 NERSC moved to a new, state of the art facility on the main Lawrence Berkeley National Laboratory (Berkeley Lab) site. The minimally disruptive move of the DOE JGI's computational infrastructure was completed in early 2016, providing nearly uninterrupted access to compute capabilities for DOE JGI scientists.

A long term goal and strategic initiative of the Berkeley Lab and the Biosciences Area management team has been to relocate and integrate the various off-site Biosciences organizations at the main Berkeley Lab site. I am thrilled to report that the first steps of this strategic initiative are well under way with the completion of the design phase and beginning of construction in November 2016 of the DOE JGI's new home, the Integrative Genomics Building. We anticipate the DOE JGI will relocate to this building at the Berkeley Lab site in the spring of 2019. We are very excited about this move, as it is expected to provide significant gains in scientific and operational productivity, efficiency, and competitiveness of the DOE JGI and KBase, as well as related research programs in the Biosciences Area.

With these significant achievements, 2016 has been another productive year for the DOE JGI. As we look forward to a new permanent Director coming on board in 2017, our efforts have not slowed down during this period of transition. We continued to forge ahead with the implementation of our strategic vision, and we have again enabled our global user community to produce a large number of insightful and impactful studies in energy and environmental genomics. I hope you agree that these accomplishments bode well for the future success of the DOE JGI.

Axel Visel, PhD
Interim Director, DOE Joint Genome Institute

Achieving the **DOE Mission**

A proposal to study sphagnum, or peat moss, was selected for the DOE JGI Community Science Program in part because it complements *Physcomitrella patens* (earthmoss), a moss model previously sequenced by the DOE JGI. (*Oskar Gran, Flickr, CC BY-NC 2.0)*

The Department of Energy Joint Genome Institute (DOE JGI) is a national user facility funded by the DOE's Office of Biological and Environmental Research (BER) that conducts high-throughput DNA sequencing, synthesis, and analysis aligned with BER's bioenergy and environmental missions. These missions mirror DOE and national priorities to:

- Develop renewable and sustainable sources, exploiting genomic knowledge of plants, microbes, fungi, and microbial communities, of biofuels from plant biomass
- Gain insights into biogeochemical processes controlling the cycling of carbon, nitrogen, and key nutrients in environments and the mobility of heavy metals and radionuclides at contaminated sites for which DOE has stewardship responsibilities

Bioenergy

The United States is the world's largest consumer of petroleum, and most of this energy is used for transportation and industry. This drives the DOE's focus on developing clean, renewable, and sustainable alternative fuel sources from lignocellulosic biomass. Such fuels would ideally offer energy content on par with gasoline while being compatible with the existing fuel distribution infrastructure. Sequencing projects at the DOE JGI that contribute to meeting this goal focus on one of three categories: terrestrial plants that can be used as feedstocks for biofuel production and their associated microbial communities (microbiomes); fungi, microbes, and microbial communities that can break down the lignin and cellulose in plant walls; and organisms that can convert lignocellulose-derived sugars or lignols into biofuels or other bioproducts currently produced from petroleum, such as plastics.

Biogeochemistry

Many DOE-relevant environmental processes are controlled by complex, interconnected biogeochemical reactions. The DOE JGI engages in projects that can couple a genome-enabled understanding of biological processes in the context of the physical, chemical, and geochemical processes controlling the cycling and fate of key elements in environments affecting BER's energy and environmental missions. Microbes and microbial communities of interest to the DOE JGI as targets for sequencing include those involved in terrestrial carbon, nitrogen, phosphorus, sulfur and other macronutrient cycles that impact sustainable bioenergy crop growth or global carbon cycling. Others include those involved in the iron, sulfur and manganese cycles that mediate the transformation of DOE-relevant contaminants, such as heavy metals or radionuclides in soils, freshwater aquatic sediments and the subsurface. As microbes constitute the largest component of Earth's biodiversity and biomass, understanding how they metabolize these elements and how environmental changes affect these processes is crucial.

Organizational **Structure**

Senior Management Team

Axel Visel
DOE JGI Interim Director

Kjiersten Fagnan
Chief Informatics Officer

Susannah Tringe
Deputy, User Programs

Len Pennacchio
Deputy, Genomic Technologies

Axel Visel
Deputy, Science Programs

Ray Turner
Deputy, Operations

Tootie Tatum
Project Management Office

Ronan O'Malley
Sequencing Technologies

Yasuo Yoshikuni
DNA Synthesis Science Program

Atif Shahab
Institutional Informatics

Trent Northen
Metabolomics Group

Dan Rokhsar
Eukaryote Super Program

Igor Grigoriev
Fungal Program

Jeremy Schmutz
Plant Program

Nikos Kyrpides
Prokaryote Super Program

Tanja Woyke
Microbial Program

Emiley Eloe-Fadrosh
Metagenome Program (Interim)

Impact **2016**

Primary Users **Fiscal Year 2016**

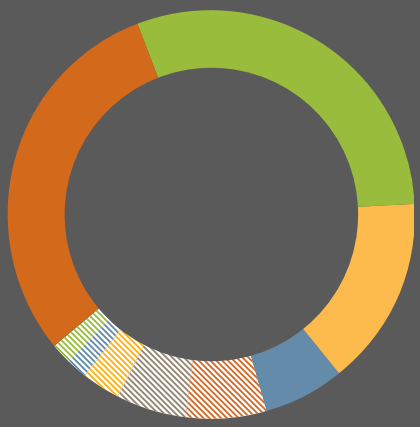
This category captures the primary users of the DOE JGI, which include PIs and their collaborators on all user projects that were active during FY 2016. Each user is uniquely identifiable and is counted once per year regardless of the number of active projects in which he/she may be involved. This count does not include collaborators who are employed by the DOE JGI or funded through the DOE JGI's partner subcontracts.

Users **by Institution Type**

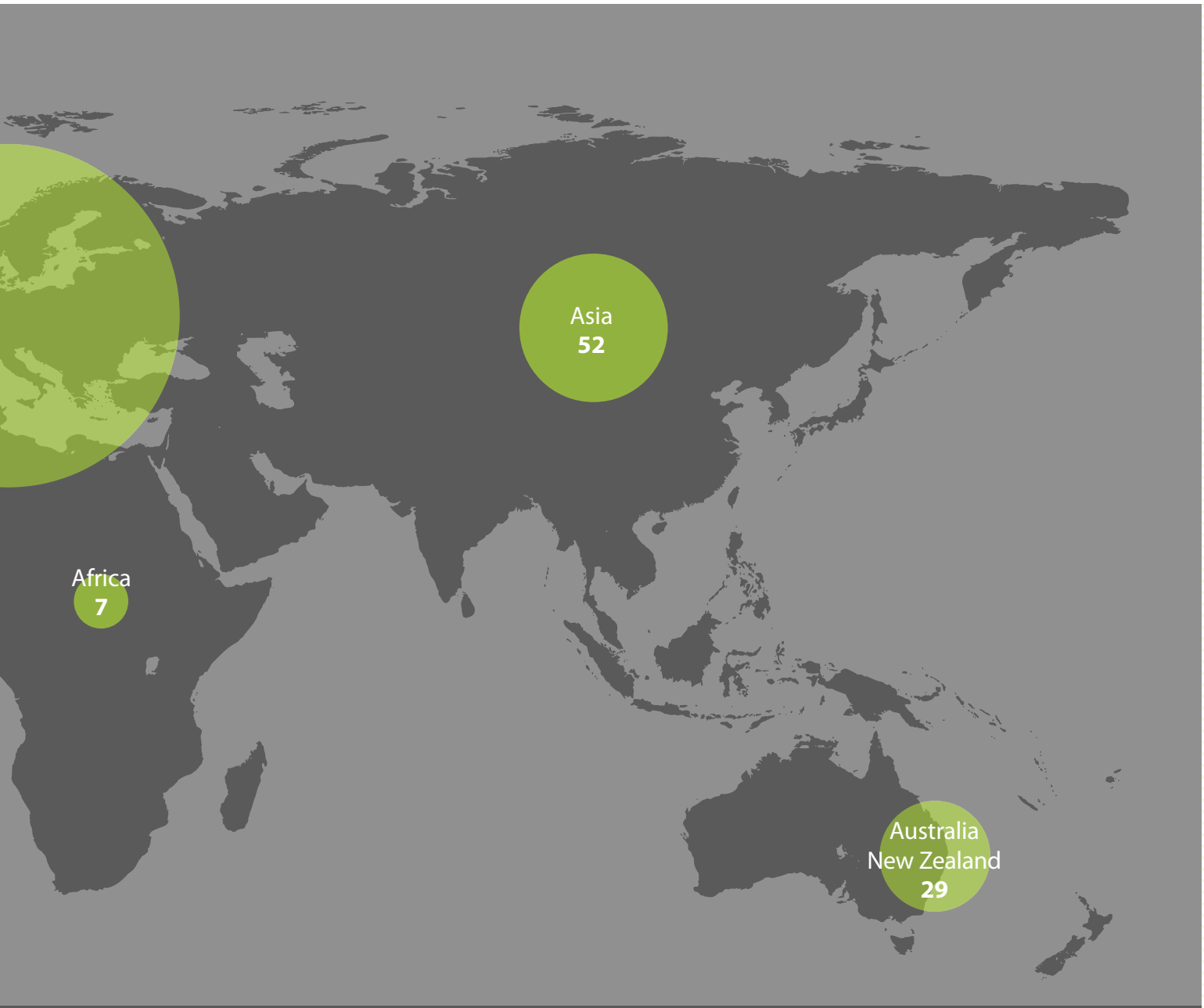
Academic	984
Other	240
DOE National Laboratory	134
Company	33

JGI Expenses **FY2016**

- **30.2%** Science & Analysis
- **30.1%** Sequence/Data Generation
- **14.9%** R&D
- **6.5%** Facilities
- **6.4%** Operations
- **5.5%** National Energy Research Scientific Computing Center (NERSC)/IT
- **3.0%** User Support/Project Management
- **1.8%** Management
- 1.7% Emerging Technologies Opportunity Program (ETOP)



North America	1017	South America	8	Europe	278		
United States	941	Brazil	5	Germany	63	Belgium	6
Canada	71	Columbia	2	France	36	Czech Republic	4
Mexico	4	Peru	1	Spain	33	Denmark	4
Panama	1			United Kingdom	28	Portugal	4
				Netherlands	20	Hungary	3
				Austria	20	Russian Federation	2
				Italy	12	Turkey	2
				Finland	11	Estonia	1
				Sweden	10	Greece	1
				Norway	8	Iceland	1
				Switzerland	8	Slovenia	1



Africa	7	Asia	52	Australia & New Zealand	29
South Africa	4	Japan	14	Australia	22
Senegal	2	China	15	New Zealand	7
Egypt	1	India	8		
		Israel	5		
		Taiwan	6		
		Republic of Korea	2		
		Malaysia	1		
		Singapore	1		

Users of **JGI Tools & Data**

DOE JGI systems also support investigators who have utilized computational and/or data resources located at the DOE JGI, but are not included in the primary user count because their projects were not conducted as part of DOE JGI's user programs.

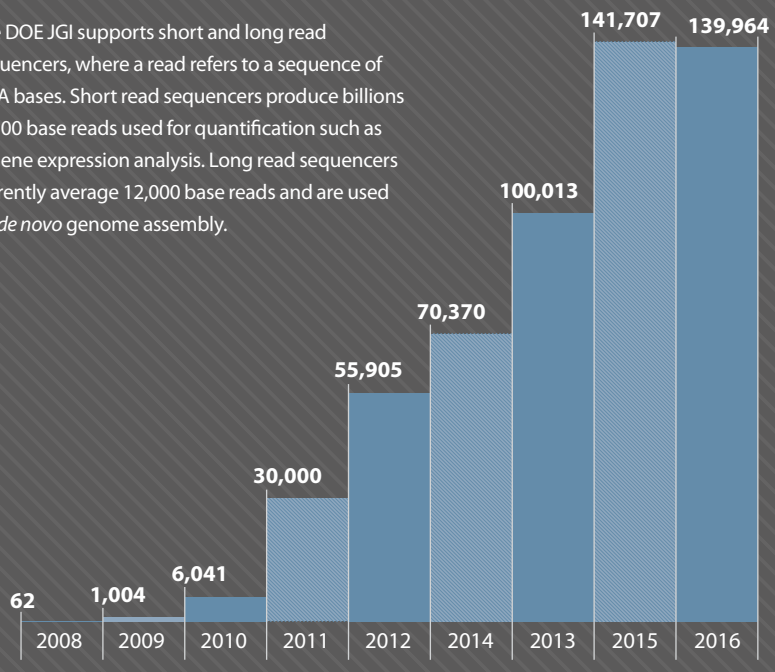
Workshops and Meetings	
Genomics of Energy & Environment	
11th Annual User Meeting Participants	**497**
Other Workshop Participants	**574**
Web Portal Visitors *(unique visits)*	
JGI Portal	**228,253**
IMG Systems	**287,099**
MycoCosm	**120,131**
GOLD	**112,359**
Phytozome	**97,769**

Sequence **Output**

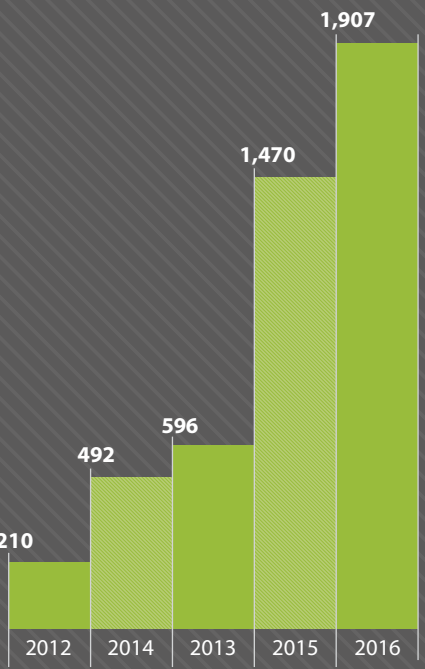
(in billions of bases or GB)

Massively Parallel Short Read Sequencing

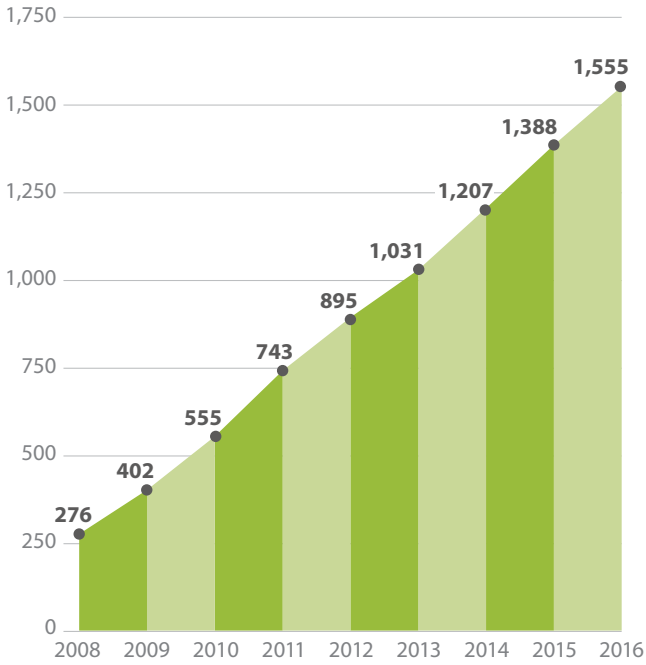
The DOE JGI supports short and long read sequencers, where a read refers to a sequence of DNA bases. Short read sequencers produce billions of 300 base reads used for quantification such as in gene expression analysis. Long read sequencers currently average 12,000 base reads and are used for *de novo* genome assembly.



Single Molecular Long Read Sequencing

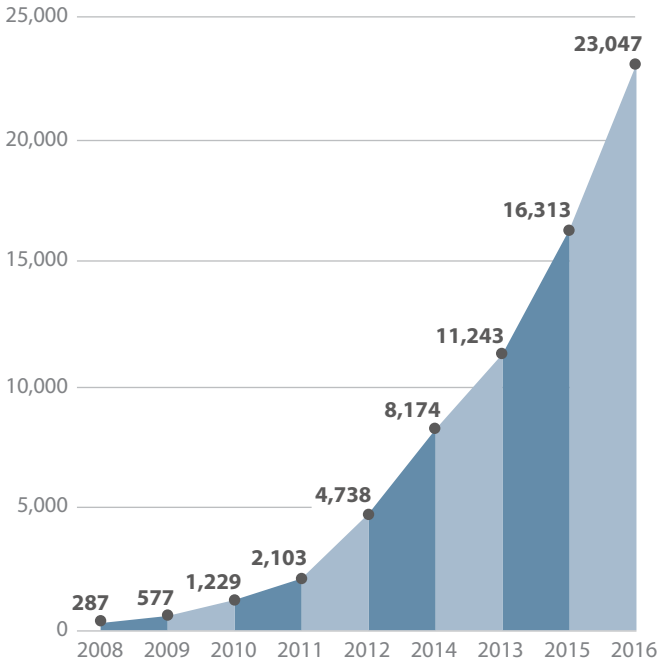


Cumulative Number of **Scientific Publications**

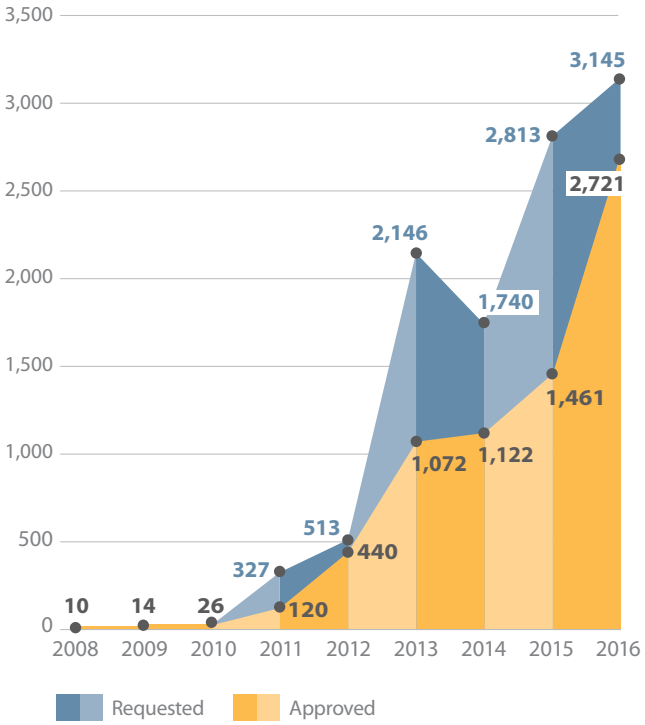


Papers published in 2011–2013 were cited by researchers nearly 5,000 times in 2014.

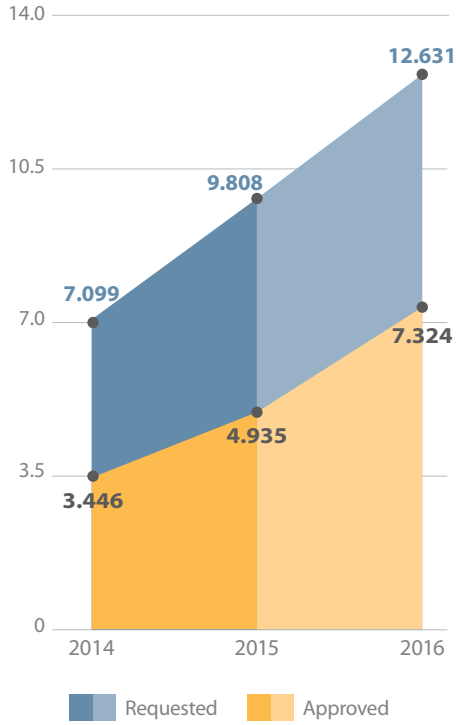
Cumulative Number of **Projects Completed**



User Demand: **CSP Single-Cell Genomes Requested & Approved**



User Demand: **Synthetic Biology Requested & Approved (Mb)**



Case Study: **A Decade of Poplar Genomics**

On September 15, 2006, the cover of the journal *Science* featured the poplar, or black cottonwood, tree, *Populus trichocarpa*. The species was the first tree to have its DNA sequenced, based on a proposal submitted by an international consortium. The genome of less than a billion bases was generated by the DOE JGI not long after it became a national user facility.

A commemorative sapling planted on the DOE JGI's front lawn in 2006 was clonally propagated from the poplar whose DNA was used to generate the sequence for the study. In a way, that tree's rapid growth reflects the rise of various genomics technologies that have been harnessed by the research community availing of the poplar genome resources.

"The impact of the genome sequence goes way beyond these quantifiable measures because it has fundamentally changed the way that we do science, and opened areas of inquiry that we didn't even anticipate in the early days," noted DOE JGI collaborator Steve DiFazio of West Virginia University, an author of the 2006 poplar study.

"The genome sequence has allowed researchers to delve into these previously-intractable areas by providing a simplified framework and infrastructure to build inquiries around," he added. "This has allowed people to invest their intellectual and monetary resources into creative and potentially risky research. Prior to the sequencing of the genome, molecular geneticists would typically focus their efforts on a handful of genes that they thought were important for a particular phenotype, usually based on work in *Arabidopsis* or another model annual plant. The genome sequence jump-started that process, so people were able to take a discovery-based approach to their research, resulting in the emergence of poplar as a true model organism that is responsible for many novel findings."

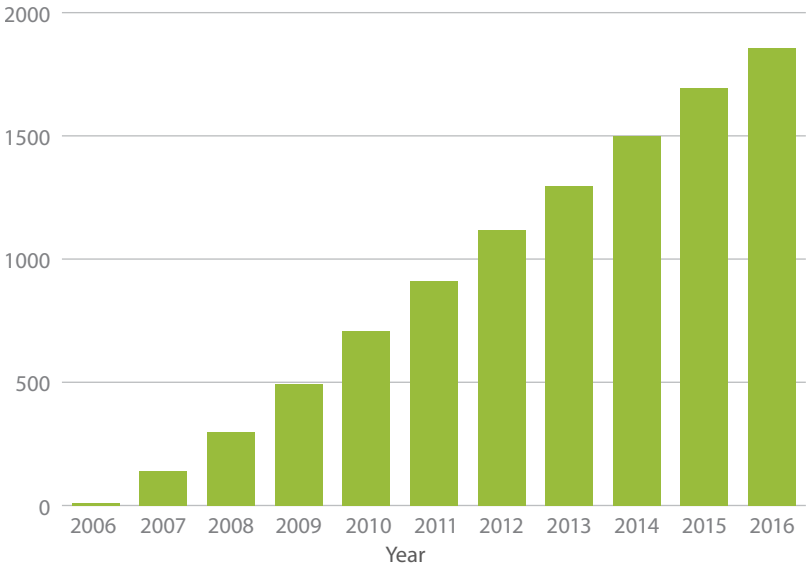
According to Web of Science, the poplar genome paper has been cited nearly 1,900 times since its publication, but that's not the only way to measure its impact.

"There are now over 1,000 *Populus* genomes that are resequenced, and have had multiple tissues characterized using RNA-seq, as well as the rich expression data now available through the Gene Atlas project at JGI," said Oak Ridge National Laboratory (ORNL) distinguished scientist Jerry Tuskan, one of the senior authors of the 2006 poplar genome paper. Tuskan heads a team studying environmental impacts on poplar plantations along the Pacific Northwest, utilizing the poplar genomes already available, and also leads a team at the BioEnergy Science Center (BESC) studying poplar growth and development. He added that numerous research projects have received funding from DOE, National Science Foundation and U.S. Department of Agriculture over the past 10 years, "due to the availability and quality of the genomics resources created by JGI." Additionally, he said, the *Populus* genomic resource "has facilitated research programs in numerous other countries, including China, Sweden, Canada, Spain, Chile, and France."

In 10 years, the commemorative sapling planted at the DOE JGI has grown from less than four feet to more than 40 feet. *(Left: 2006 planting; Right: Poplar in 2016, DOE JGI)*

Cumulative Number of **Poplar Citations**

2006–2016



Source: Web of Science

ORNL's Jerry Tuskan (left) and Wellington Muchero (right) studying poplar roots. *(DOE JGI)*

Science:
A Year in Review

Eucalyptus is one of the world's most widely planted hardwood trees, and researchers seek to harness and improve upon its potential for enhancing sustainable biofuels and biomaterials production. The *Eucalyptus grandis* genome was sequenced and analyzed by the DOE JGI as part of the Community Science Program. *(John Andrew Rice, Flickr CC BY-2.0)*

Continuing toward the goal of offering the user community access to cutting-edge capabilities for genomics research, the DOE JGI has teamed with researchers around the world to develop both techniques and software tools.

Through the DOE JGI Emerging Technologies Opportunity Program (ETOP), University of Vienna and DOE JGI researchers describe a pipeline to efficiently prepare cell extracts from soils for subsequent single-cell methods, such as high-resolution secondary ion mass spectrometry (NanoSIMS) or Raman microspectroscopy. In a study published in the October 2015 issue of *FEMS Microbiology Ecology*, the team applied the techniques to study uncultivated microbes in soil. NanoSIMS has been used to study the *in situ* function of uncultured microbes in their native environments, while Raman microspectroscopy, which allows researchers to learn about the molecular composition of a sample, has been used to characterize soil structure or mineral content. The work furthers previously reported studies from the same team where they carried out proof-of-principle experiments.

In a study published online June 28, 2016, in the *Proceedings of the National Academy of Sciences*, DOE JGI researchers helped modify a recently refined technique called BONCAT to help identify both individual active cells, and single clusters of active bacteria and archaea within microbial communities. (More on page 42).

The DOE JGI continues to develop computational tools to assemble genome sequences to further study unexplored and undiscovered life. In a paper that was part of the *PeerJ* Top Genomics Papers – October 2015 list, DOE JGI researchers offer an automated metagenome binning software tool called MetaBAT. Researchers are still working to develop methods to efficiently and accurately assemble individual microbial genomes from high-throughput metagenome shotgun sequencing datasets to learn more about each one's specific contributions to maintaining the global cycles. Tools such as MetaBAT enable researchers to better appreciate the data generated, which allows them to study microbial communities without needing to cultivate them. The team evaluated MetaBAT using both synthetic and real-world metagenome datasets and found that MetaBAT recovered "many [genomes] missed by alternative tools" and was computationally efficient.

Massive amounts of sequence and analysis data require infrastructure to manage and store the information in a manner than can be easily accessed for use. In a paper published May 16, 2016, in *Trends in Microbiology*, DOE JGI researchers called for the formation of a National Microbiome Data Center to efficiently manage the datasets accumulated globally. While technologies have scaled to allow researchers to sequence and annotate communities of microorganisms within an environment (its "microbiome") on an ever-increasing scale, the data management aspect has not been developed in parallel. By integrating and harnessing all available microbiome data and metadata, researchers could conduct larger-scale comparative analyses to address global challenges related to energy, environment, health, and agriculture.

Mycorrhizal fungi live in the roots of host plants, where they exchange sugars that plants produce by photosynthesis for mineral nutrients that fungi absorb from the soil. They include some of the most conspicuous forest mushrooms, like *Amanita muscaria*, commonly known as the fly agaric, and are of interest to bioenergy researchers because they play roles in maintaining the health of candidate feedstock crop trees. *(Michael Hartwich, Wikimedia Commons, CC BY-SA 4.0)*

Bioenergy

Fungi for Biofuel Production

Scientists have shown that fungi found in the guts of goats, horses, and sheep could help fill up your gas tank. Reported February 18, 2016, in *Science*, a team led by University of California (UC), Santa Barbara researchers reports that these anaerobic gut fungi perform as well as the best fungi engineered by industry in their ability to convert plant material into sugars that are easily transformed into fuel and other products. These enzymes — tools made of protein — work together to break down stubborn plant material. Additionally, the researchers found that the fungi adapt their enzymes to whatever scientists feed them. The findings suggest that industry could modify the gut fungi so that they produce improved enzymes that outperform the best available ones, potentially leading to cheaper biofuels and bio-based products.

"Because gut fungi have more tools to convert biomass to fuel, they could work faster and on a larger variety of plant material. That would open up many opportunities for the biofuel industry," said lead author Michelle O'Malley of UC Santa Barbara. O'Malley and her team collaborated with the DOE JGI and the Environmental Molecular Sciences Laboratory (EMSL) at Pacific Northwest National Laboratory, under a collaborative science initiative called Facilities Integrating Collaborations for User Science (FICUS).

The DOE JGI sequenced the mRNA of several gut fungi to come up with their transcriptome, which represents all the possible proteins they could make. O'Malley compared this effort to re-assembling a map from its pieces, only without seeing the complete picture. Since not all proteins are enzymes, the researchers needed to cross check their map with another one. Enter the EMSL, where researchers created that second map that identified enzymes the fungi actually produced.

Together, the maps from the DOE JGI and EMSL pointed to the treasure trove of enzymes gut fungi can produce. Compared with the industrial varieties, which top out around 100 enzymes, gut fungi can produce hundreds more. Of note, researchers found that the fungi can update their enzyme arsenal on the fly.

Healthy cassava plant. *(Simon Prochnik, DOE JGI)*

Improving Cassava Production

Cassava is a staple crop and a primary source of calories for nearly a billion people around the world. The plant is easy to cultivate — cuttings grow well on marginal land — and it is very tolerant of drought. For the U.S. Department of Energy (DOE), these traits and its starchy qualities make cassava of interest as a potential feedstock for biofuel production.

Though cassava is easy to cultivate, it is particularly vulnerable to plant pathogens, which can significantly reduce crop yields. To help improve breeding strategies for this root crop, a team led by researchers from the University of California, Berkeley and including DOE JGI researchers, has described cassava's genetic diversity in the April 18, 2016, advance online publication of the journal *Nature Biotechnology*. As cassava roots contain 20–40 percent starch that costs 15–30 percent less to produce per hectare than starch from corn, in many parts of the world, particularly Africa and Southeast Asia, it represents a strategic source of renewable energy — biomass from which ethanol is being produced for transportation fuels. With the help of genomics, researchers hope to apply advanced breeding strategies that can improve cassava's resistance to diseases and improve crop yields.

The cassava genome was initially sequenced under the aegis of the DOE JGI Community Science Program and Roche 454 Life Sciences. Since the draft sequence was released in 2009, researchers have improved it with additional data in order to develop a chromosome-scale sequence, in part to apply the information toward improved breeding strategies.

The team compared the cassava reference genome to the genomes of castor bean, rubber tree, Ceara rubber, and 53 cultivated and wild cassava varieties from around the world. The researchers found that the genetic diversity of cassava used in current breeding efforts has been greatly reduced in Africa, where viruses such as the cassava mosaic disease and the cassava brown streak disease have affected crop yields in many nations. They were able to detect the genetic signature of past cassava improvement programs going back to the 1930s, which interbred cassava and Ceara rubber, and the persistence of these Ceara rubber regions in elite cassava varieties suggests they confer desirable traits.

"The variants and population structure described here are essential inputs for marker-assisted and genome selection-based approaches to improving disease resistance and yield for this staple crop," the team noted.

Environmental Responses in Fungi

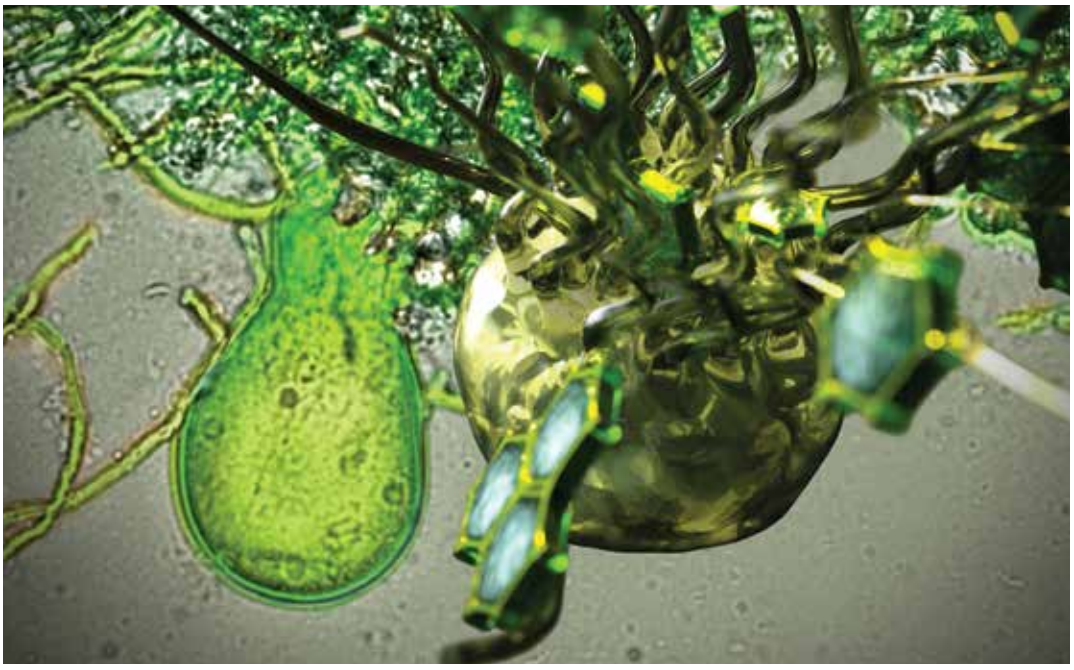
Sensory perception lies at the heart of adaptation to changing conditions, and helps fungi to improve growth and recycle organic waste, and to know when and how to infect a plant or animal host. New results based on characterizing and then conducting a comparative analysis of two genome sequences published online May 26, 2016, in the journal *Current Biology* shed light on the evolution of sensory perception in fungi.

For the study, researchers sequenced and annotated Mucoromycotina genomes, specifically those of *Phycomyces blakesleeanus* and its relative *Mucor circinelloides*. Capturing the genomic variation of fungi allows researchers to build a foundation for translating their genomic potential into practical applications. For example, understanding the mechanisms by which these environmental cues are sensed could provide insights on how some fungi can change their growth patterns to act as pathogens rather than benign organisms.

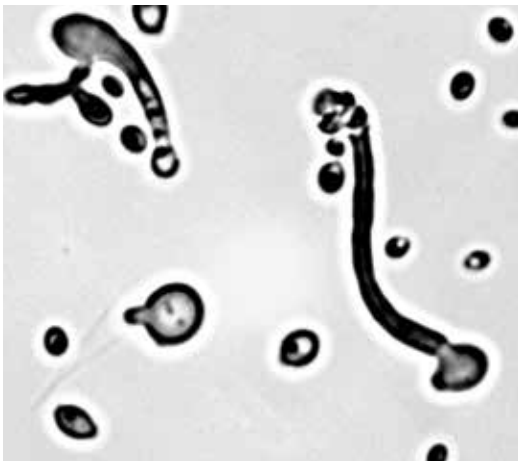
"Very little is known about basal fungi such as Mucoromycotina, and genomics may be the most efficient way to understand their metabolism," said DOE JGI Fungal Genomes program head Igor Grigoriev of this project.

The team found that both species had undergone whole genome duplication, a rare occurrence in fungi. The genome duplication led to an expansion of gene families and the development of specialized genes, providing new proteins that have enabled these fungi to refine the way they perceive signals from the environment to regulate their growth and development.

Mucor circinelloides, a related fungus, has similar environmental responses, and sometimes also acts as a human pathogen. The genome sequences of *Phycomyces* and *Mucor* provide evidence that an ancient genome duplication yielded new genes to expand signal transduction pathways and improve the mechanisms for sensing light and other signals in this group of fungi. For example, the team found that exposure to light produced massive changes in gene regulation in different fungal tissues, and hypothesized that the refined response to light is a consequence of gene duplication and specialization after whole genome duplication. These results advance understanding the role of genome dynamics in the evolution of sensory perception, which, in turn, could provide leads to genes and pathways useful for understanding fungal adaptation and for accelerating the breakdown of biomass for bioenergy.



As reported on page 23, gut fungi in herbivores break down plant mass as well as fungi engineered by industry. Here, anaerobic gut fungi colonize biomass, and secrete enzymes that release free sugars into their environment. *(Artistic rendering of the fungi by UC Santa Barbara engineering graphic designer Peter Allen)*



One of the yeast genomes sequenced was of *Scheffersomyces stipitus. (Thomas Jeffries, University of Wisconsin, Madison).* Right: Researchers found a non-stop codon reassignment in the genome of the yeast *Pachysolen tannophilus. (Courtesy of Cletus Kurtzman, USDA-ARS)*

New Findings in Yeast Genetic Diversity

Yeasts (which belong to the "kingdom" of fungi) can use a wide range of carbon and energy sources, ranging from cellulosic (6-carbon) and hemicellulosic (5-carbon) sugars to methanol, glycerol, and acetic acid. Products include ethanol and other alcohols, esters, organic acids, carotenoids, lipids, and vitamins.

"We sequenced these diverse genomes to expand the catalog of genes, enzymes, and pathways encoded in these genomes for producing biofuels and bio-based products we use in daily life," said Igor Grigoriev, DOE JGI Fungal Program Head and co-senior author of the manuscript, which appeared in the *Proceedings of the National Academy of Sciences* on August 15, 2016.

Sequencing these less-known yeasts and characterizing their metabolic pathways, added study first author Robert Riley of the DOE JGI, helps fill in knowledge gaps regarding the fungal enzymes that can help convert a wide range of sugars into biofuel. The well-known yeast *Saccharomyces cerevisiae*, for example, ferments glucose, but not the full range of sugars found in plant biopolymers.

One of the newly sequenced yeasts is *Pachysolus tannophilus*, which can ferment xylose, otherwise known as wood sugar as it is derived from hemicellulose. Along with cellulose, xylose is one of the main constituents of woody biomass. It is only distantly related to well-studied xylose fermenters such as *Scheffersomyces stipitis* — another yeast sequenced by the DOE JGI.

These distances are huge. "We might think of yeasts as simple unicellular, creatures similar to each other, but in fact their genetic diversity is like the difference between human and invertebrate sea squirt," said Riley. "We sequenced these diverse genomes to discover and facilitate the next generation of biotechnological workhorse yeasts for producing the fuels and products we use in daily life. We also discovered a genetic code change that, if not understood, will impede the yeasts' biotechnological use."

"Right with the advent of new genetic tools that can rapidly manipulate an organism's DNA, publication of these new genomic yeast sequences will open up many new platforms for bioengineering cellulose degrading, lipid producing, acid tolerant yeasts that use a wide range of substrates and produce many different primary and secondary metabolites," said senior author Tom Jeffries, Professor Emeritus at the University of Wisconsin, Madison. "The yeast scientific community owes tremendous thanks to the sequencing and annotation team at JGI."

Endophyte Genome Offers Window into Fungal Lifestyles

Endophytes reside within living plant cells and can play roles not just in plant health but also within carbon and nitrogen cycles. The evolution of endophytism and mechanisms by which they interact with their hosts are still poorly understood. To gain insights into these questions, a team including DOE JGI researchers and longtime collaborators at Clark University sequenced the genome of *Xylona heveae*, an endophyte from a Peruvian rubber tree, and then compared its genome to the genomes of 36 related fungi within the phylum Ascomycota. The report was published in the January 2016 issue of *Fungal Biology*.

Based on this dataset of 37 fungal genomes, a third of which were sequenced and annotated at the DOE JGI, the team found that endophytes such as *X. heveae*, to avoid triggering their plant hosts' immune systems, adapted by reducing their CAZymes, cellulose-degrading genes that encode carbohydrate-active enzymes. At the same time, the fungal endophyte also appears to have expanded the number of enzymes that could boost its chances of survival within the host plant's intercellular spaces. The authors theorized that in adapting to its host, *X. heveae* has reduced its wood-degrading capabilities to the point that it likely cannot switch to a different lifestyle or become pathogenic. They also suggested that this fungal endophyte could be transmitted through insects.

For this study, researchers sequenced the genome of *Xylona heveae*, a fungal endophyte from a Peruvian rubber tree, and compared it to 36 other fungal genomes. *(Marco Simola for the Center for International Forestry Research (CIFOR), Flickr, CC BY-NC 2.0)*

Fungi Can Help Host Trees Be Drought Resistant

The mutualistic relationship between tree roots and ectomycorrhizal (ECM) fungi has been shaping forest ecosystems since their inception. ECM fungi are key players supporting the growth, health and stress tolerance of forest trees globally, such as oak, pine, spruce, birch and beech, and help boost the productivity of bioenergy feedstock trees, including poplar and willow. The most common ECM fungus is *Cenococcum geophilum*, found in subtropical through arctic zones and especially in extreme environments.

To learn more about what ectomycorrhizal characteristics are dominant in *Cenococcum geophilum*, a team led by researchers at the French National Institute for Agricultural Research (INRA) and the Swiss Federal Institute for Forest, Snow and Landscape Research WSL, and including researchers at the DOE JGI, compared its genome with the genomes of close relatives, *Lepidopterella palustris* and *Glonium stellatum*, neither of which are ECM fungi. The study was published online September 7, 2016, in *Nature Communications*. The team found specific adaptations in the *C. geophilum* transcriptome — the set of its messenger RNA molecules that reflects actual biochemical activity by the fungus — that could help their hosts be more resistant to drought stress, a finding that could be useful in developing more plant feedstocks for bioenergy.

Noting that the tree root tips colonized with *C. geophilum* are highly resistant to desiccation, one of the team's key findings is that two of the three most highly induced *C. geophilum* genes in symbiosis code for water channels. "The regulation of these water channel genes is fine-tuned under drought conditions, and they might therefore play a key role in drought adaptation of host plants," said first author Martina Peter of the Swiss Federal Research Institute WSL.

"The intersection of genomics and evolutionary biology, as carried out in the MGI and the 1000 Fungal Genomes (KFG) Project, can inform our understanding of the biological principles intrinsic to mycorrhizal symbiosis," said study senior author Francis Martin of INRA. "By combining genome sequences with rigorous physiological and ecological studies, we are entering a time where linking the presence, composition, and abundance of soil mycorrhizal communities with important soil processes and forest productivity at an ecosystem scale is possible."

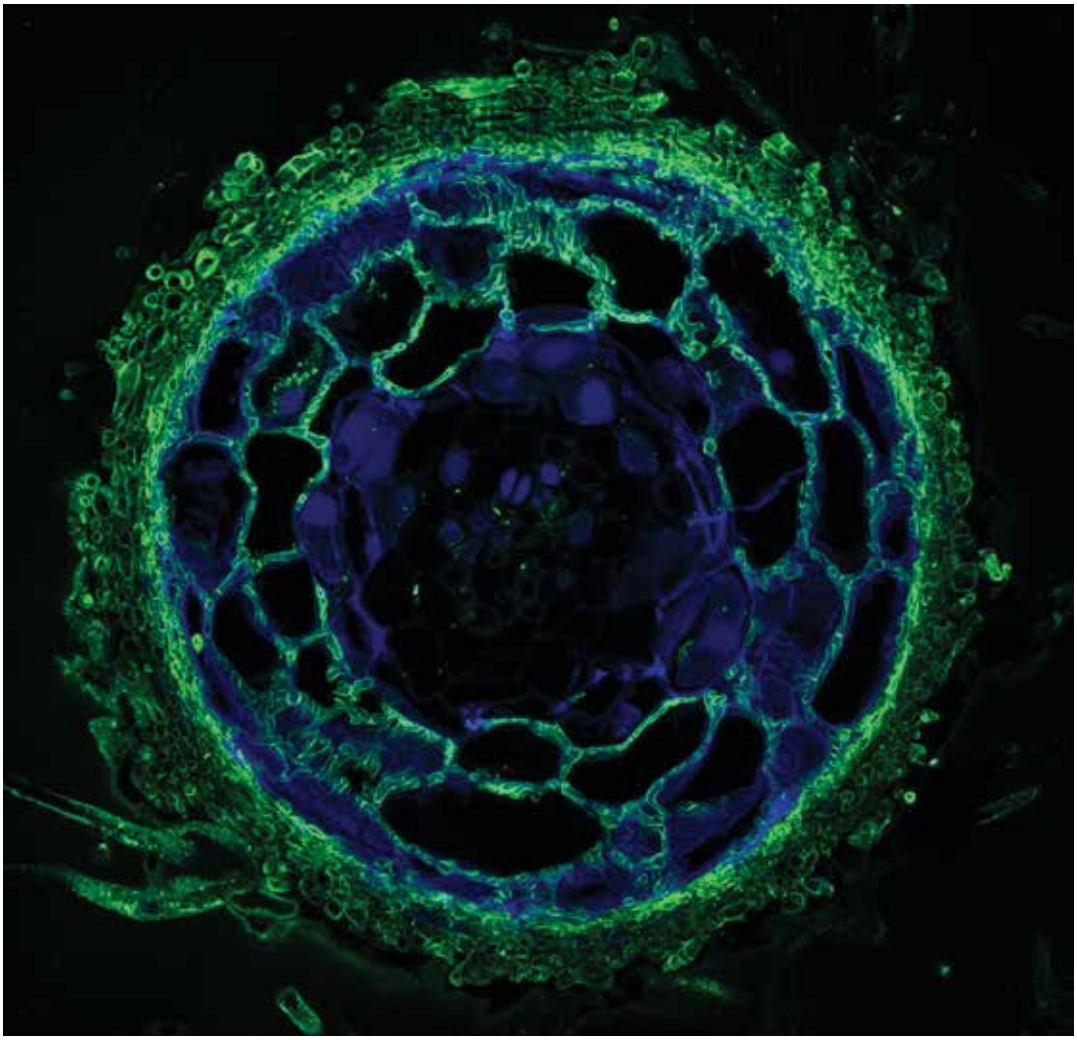
DNA Methylation in Prokaryotes

The epigenome of a cell is the collected set of changes made to specific bases in its genomic DNA that affects how the genome is actually used and results from chemical modification (usually methylation). DNA methylation is a process eukaryotes use to regulate gene expression (for example, keeping certain genes from turning on). Though prokaryotes (bacteria and archaea) are also known to have methylated DNA, the roles this process might play in these single cell organisms is less well understood. To learn more, a team including DOE JGI researchers relied on single-molecule, real-time (SMRT) sequencing at the DOE JGI and Pacific Biosciences to reveal DNA methylation patterns in 230 bacterial and archaeal genomes. The report was published February 12, 2016, in *Plos Genetics* and was highlighted February 29, 2016, in *Nature Reviews Genetics*.

The team found evidence of DNA methylation in 215 microbes (93 percent of those sequenced). These data enabled the annotation of 600 enzymes that methylate DNA (MTases), a massive increase over known annotations. While many DNA methylating enzymes are part of restriction modification systems, consistent with their known role in defense against phages and viruses, the findings suggest that a substantial number of others may be involved in genome regulation, and have a more crucial role in prokaryotic physiology and biology than had been previously suspected. By mapping and characterizing the epigenetic changes, scientists can associate those targeted genes with environmental adaptations and metabolic activities. Better understanding of such controls on gene expression and under what circumstances they are observed will improve the ability to predict when and where such microbes are detected. In addition, this will inform how microbes interact with plants and other microbes involved in DOE mission interests, such as plant bioenergy feedstock growth, advanced biofuel generation, and soil carbon processing.

Reconstructing the Origins of Rot Fungi

Fungi are natural degraders of plant material and contain enzymes that biofuels researchers hope to harness for commercial applications. Researchers are particularly interested in wood-decaying fungi because their enzymes allow them to break down plant materials such as cellulose, hemicellulose, and lignin. Brown rot fungi are typically described as those that can only break down cellulose and hemicellulose, while white rot fungi can break down all



The *C. geophilum* crosscut shows the fungal tissues — the fungal mantle around the root tip and the fungal network of tendrils that penetrates the root of plants, or Hartig Net, between *Pinus sylvestris* plant root cells — in green. *(Maira de Freitas Pereira, INRA Nancy.)*

the plant cell wall components — cellulose, hemicellulose, and lignin. The wide range of fungal genome sequences now available has also demonstrated the breadth of wood-decaying strategies fungi employ; for example, while some fungi may break down plant matter in the same manner as white rots, they lack some of the gene families typically associated with white rot fungi.

In a study published in the April 2016 issue of *Molecular Biology and Evolution*, DOE JGI researchers and collaborators at the INRA and Clark University in Massachusetts conducted a comparative genomics analysis with the help of 10 fungi sequenced, assembled, and annotated by the DOE JGI in order to better reconstruct the origins of lignin-degrading fungal enzymes. The work builds off of previous work in which DOE JGI and Clark University researchers suggested that the evolution of white rot fungi some 300 million years ago coincided with the end of the coal-forming Carboniferous period, impacting the global carbon cycle.

That analysis did not include genome sequences for several groups of early diverging fungi, some of which include species that suggest white rot evolution could have happened even earlier. With draft genomes representing early-diverging fungi in hand now, a more refined analysis shows that "the majority of oxidative enzyme diversity used by white rot fungi has emerged after the origins of the first white rot species," a finding that adds credence to the earlier report. In addition to clarifying fungal biomass degrading enzyme evolution, these additional fungal genomes identify additional candidate degrading enzymes of possible interest to bioenergy researchers.

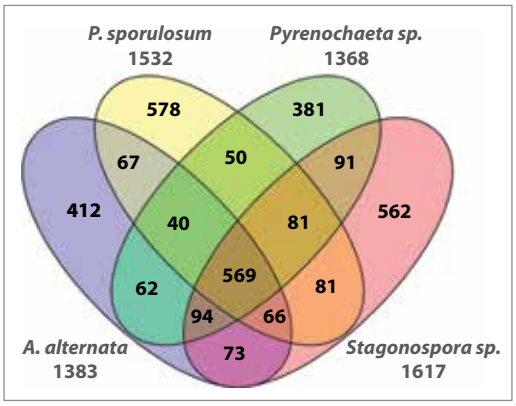
One of the 10 early-diverging fungi sequenced, assembled, and annotated by the DOE JGI to learn more about lignin-degrading fungal enzymes is the white rot *Calocera viscosa* or yellow stagshorn. *(abejorro34, Flickr, CC BY-NC 2.0)*

Uncovering Carbon Compound Degradation Pathways

Their unassuming appearances may cause them to be overshadowed by the plants or animals in their natural habitats, but fungi play key roles in maintaining their ecosystems. Fungal secretomes, those collections of all molecules secreted by a cell, contain enzymes that can break down plant cell wall components such as cellulose, hemicellulose, and lignin. These capabilities make them of interest to bioenergy researchers looking for cost-effective ways to convert plant mass into sustainable, alternative transportation fuels. In a study published online July 19, 2016, in *Plos ONE*, a team led by researchers at Harvard University and Woods Hole Oceanographic Institution (WHOI) conducted a comparative analysis of the secretomes of four recently-isolated and sequenced filamentous Ascomycete fungi to learn more about the variety of pathways they deploy to break down carbon compounds.

The team studied four Mn(II)-oxidizing fungi known to play roles in cleaning up metal contaminated waters. *Alternaria alternata* SRC1lrK2f, *Stagonospora* sp. SRC1lsM3a, and *Pyrenochaeta* sp. DS3sAY3a were all isolated from coal mine drainage treatment systems in central Pennsylvania. *Paraconiothyrium sporulosum* AP3s5-JAC2a was isolated from a remediated freshwater lake in Massachusetts. All four fungi had had their genomes sequenced and annotated as part of the collaborative science project involving DOE JGI and EMSL resources under the FICUS program to enable genomic and proteomic characterization of enzymes involved in manganese and carbon oxidation by these isolates.

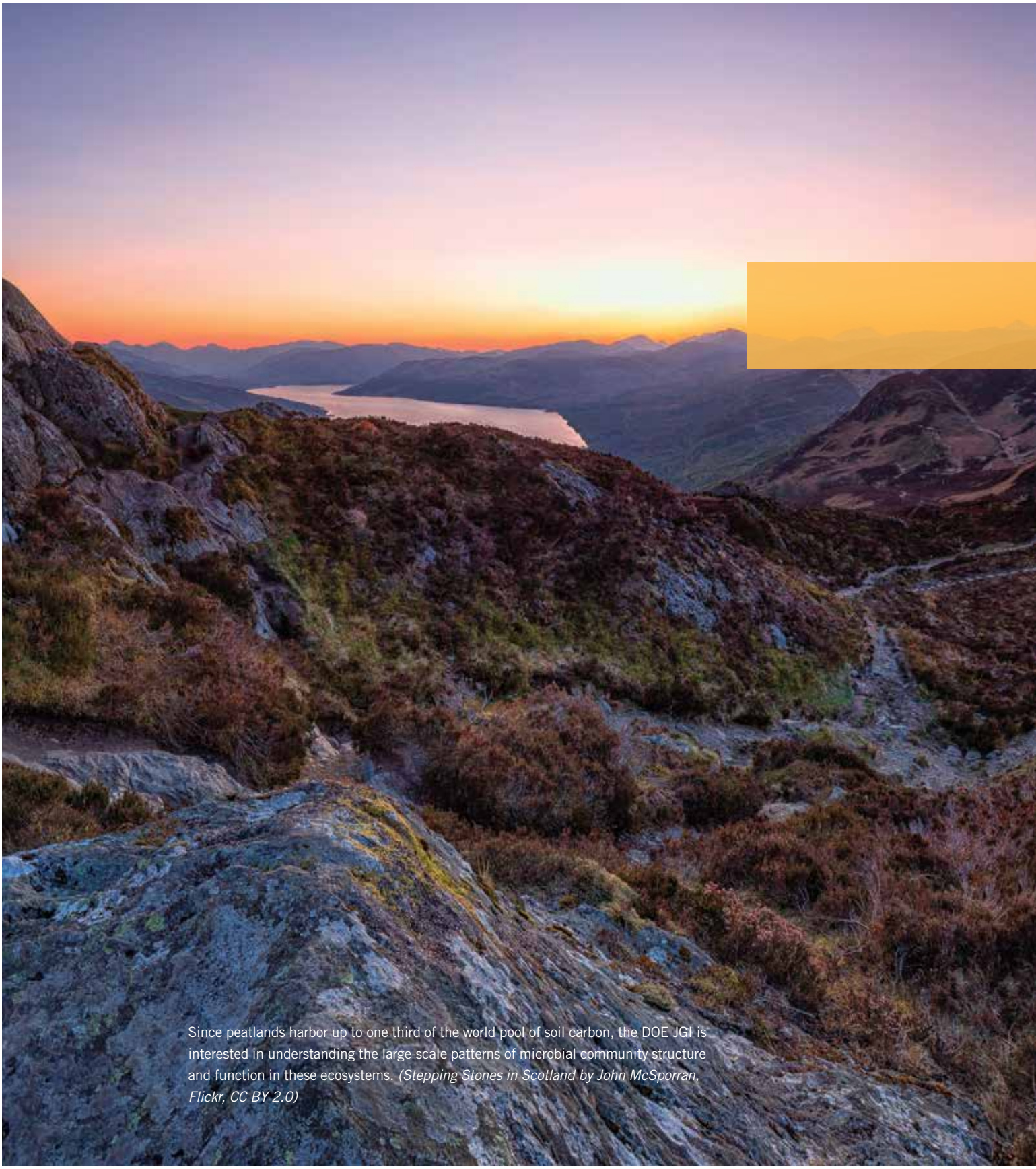
Sam Purvine (left), of Pacific Northwest National Laboratory, and first author Carolyn Zeiner (right), now at Boston University, are among the authors of the Ascomycete fungi comparative analysis. *(DOE JGI)*



Venn diagram showing number of unique and shared proteins experimentally identified in Ascomycete fungi secretomes. Proteins identified via LC-MS/MS over a 21-day study. Total number of proteins identified for each fungus is indicated outside of diagram. Diagram generated with Venny 2.0. *(Figure from Zeiner CA et al. 2016. PLoS ONE. CCO 1.0)*

"Through collaboration with EMSL, we utilized state of the art mass spectrometry facilities and leveraged expertise in quantitative, comparative proteomic methods — specifically, isobaric tags for relative and absolute quantitation (iTRAQ) — to identify over 1,300 secreted proteins per species. Through collaboration with JGI, we sequenced the genomes of our four Ascomycetes, resulting in a 400 percent increase in the number of protein matches that we were able to obtain compared to searching our datasets against the genomes of closely related species," said study first author Carolyn Zeiner, now at Boston University.

"This vastly improved our ability to harness our large datasets in drawing biological conclusions about the role of these species in carbon cycling. Finally, we collaborated with JGI bioinformaticists to generate genome-based predicted secretomes of the four species, which enhanced the functional annotation of our experimental data, aided in identifying intracellular proteins that were released into the secretome via lysis, and provided more robust assignments of carbohydrate-active enzymes [CAZymes]."



Since peatlands harbor up to one third of the world pool of soil carbon, the DOE JGI is interested in understanding the large-scale patterns of microbial community structure and function in these ecosystems. *(Stepping Stones in Scotland by John McSporran, Flickr, CC BY 2.0)*

Biogeochemistry

The First Global Viral Distribution Map

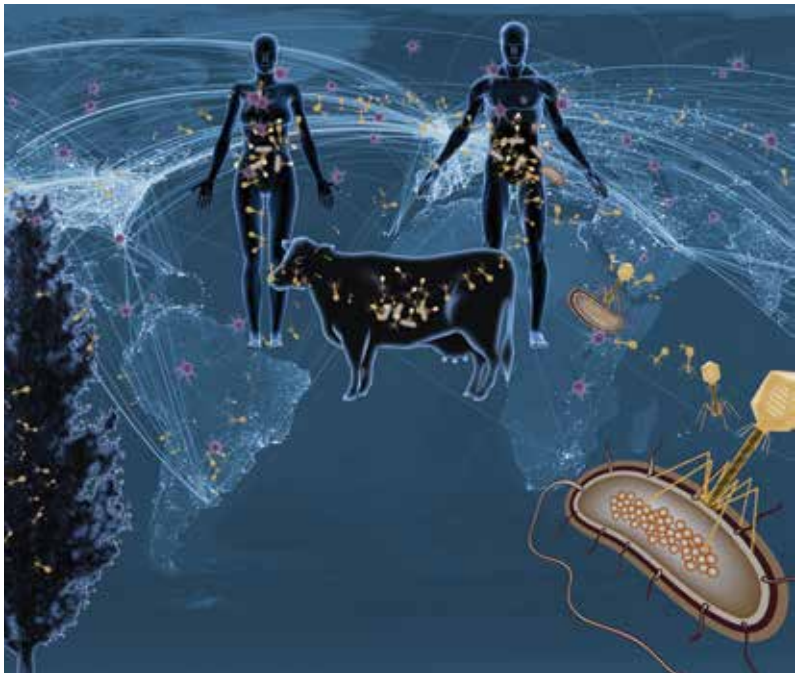
The number of microbes in, on, and around the planet — on the order of a nonillion, or 10^{30} — is estimated to outnumber the stars in the Milky Way. Microbes are known to play crucial roles in regulating carbon fixation, as well as maintaining global cycles involving nitrogen, sulfur, and phosphorus and other nutrients, but the majority of them remain uncultured and unknown. To understand Earth's microbial diversity requires learning more about the poorly studied relationships between microbes and the viruses that infect them, viruses that impact the microbes' abilities to regulate global cycles. In a study published online August 17, 2016, in *Nature*, researchers at the DOE JGI utilized the largest collection of assembled metagenomic datasets from around the world to uncover over 125,000 partial and complete viral genomes, the majority of them infecting microbes.

The team analyzed over five trillion bases of sequence available in the DOE JGI's Integrated Microbial Genomes with Microbiome Samples (IMG/M) system collected from 3,042 samples around the world from 10 different habitat types. Their efforts to sift through the veritable haystack of datasets yielded over 125,000 viral sequences containing 2.79 million proteins. The team matched viral sequences against multiple samples in multiple habitats.

"One of the most important aspects of this study is that we did not focus on a single habitat type. Instead, we explored the global virome and examined the flow of viruses across all ecosystems," said Kyrpides. "We have increased the number of viral sequences by 50 times, and 99 percent of the virus families identified are not closely related to any previously sequenced virus. This provides an enormous amount of new data that will be studied in more detail in the years to come. We have more than doubled the number of microbial phyla that serve as hosts to viruses, and have created the first global viral distribution map. The amount of analysis and discoveries that we anticipate will follow this dataset cannot be overstated."

By analyzing a CRISPR-Cas system — an immune mechanism in bacteria that confers resistance to foreign genetic elements by incorporating short sequences from infecting viruses and phages — the team generated a database of 3.5 million spacer sequences in IMG. These spacers, fragments of phage genetic sequences retained by the host, can then be used to explore viral and phage metagenomes for where the fragments may have originally come from. Also, using mainly this approach, the team computationally identified the host for nearly 10,000 viruses.

Jan-Fang Cheng, head of the DOE JGI's Functional Genomics group, said the work being done by Kyrpides' group in identifying new viral sequences will help the DOE JGI Synthetic Biology group develop novel promoters that can work in many bacterial hosts. Cheng also anticipated that the expanded viral sequence space generated by Kyrpides' team will allow researchers to look for other genetic sequences known as proto-spacer adjacent motifs



DOE JGI researchers utilized the largest collection of assembled metagenomic datasets from around the world to uncover over 125,000 partial and complete viral genomes, the majority of them infecting microbes. *(Graphic by Zosia Rostomian, Berkeley Lab Creative Services)*

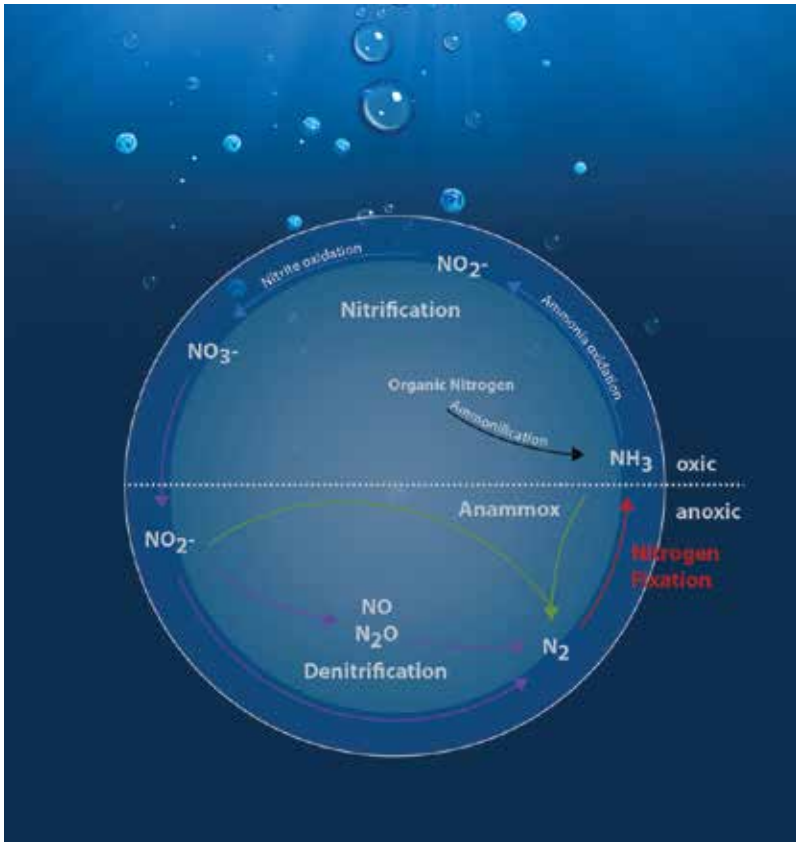
(PAMs). These sequences lie next to spacer sequencers in phages and are used as beacons by CRISPR-Cas proteins, triggering actions such as editing or regulating a gene. "People are looking for new PAM sequences and new Cas9s, and with this new information, if you can map the spacer sequence back to the same phage and align them and see what's in common in neighboring sequences, then you could ID new PAM sequences."

SAR11 Lineages Reveal Adaptions to Oxygen-Poor Ocean Zones

Oxygen minimum zones (OMZs) extend over about eight percent of the oceanic surface area, but account for up to 50 percent of the total loss of bioavailable nitrogen and thus play an important role in regulating the ocean's productivity by substantially affecting the nitrogen cycle. Though SAR11 bacteria are estimated to make up as much as half of the total microbial community in the ocean's oxygen-rich surface layers, they have also been found to be abundant in OMZs.

To understand how SAR11 bacteria can thrive in what was thought to be an inhospitable environment for them, a team led by Georgia Tech scientists and supported by a DOE JGI Community Science Program proposal analyzed single amplified genomes (SAGs) generated from water samples collected in the nitrogen reduction zone of the Eastern Tropical North Pacific OMZ. Located off Mexico, this is the world's largest OMZ and comprises about 40 percent of the total OMZ surface area.

DOE JGI researchers conducted a series of single-cell sorting and synthetic biology experiments that helped to demonstrate how SAR11 cells adapted to anoxic environments by acquiring the biochemical machinery to respire nitrate, in effect initiating oceanic nitrogen loss. The study appeared online August 3, 2016, in *Nature*. "Together," the team concluded, "these findings redefine the ecological niche of one of the planet's most dominant groups of organisms."



In the *Nature* paper, the team found functional nitrate reductase pathways, shown on the left-hand side of the nitrogen cycle, in SAR11 microbes. *(Nitrogen cycle graphic by Zosia Rostomian, Berkeley Lab Creative Services)*

A Predictive Marine Microbiome Math Model

There are areas in the ocean where there is not enough oxygen for life to exist due to factors such as warming temperatures and waste runoff, impacting the marine multicellular life that all require oxygen to thrive. There are currently more than 500 oxygen minimum zones (OMZs) worldwide, and the metabolism of the microbial communities in these oxygen-starved waters affects both nutrient and energy conversion processes, including the production and consumption of the greenhouse gases carbon dioxide, methane, and nitrous oxide.

"It takes a network to study networks. A drop of seawater can contain millions of single-celled microbes that collectively form the basis for nutrient and energy cycles in the ocean," said DOE JGI collaborator and University of British Columbia microbial ecologist Steven Hallam. "Understanding how microbial processes contribute to these cycles is vital in a time of climate change. Our model provides a step change in more accurate projections of microbial processes with potential feedback on climate change and ecosystem health." The omics component of the model comes from Hallam's work with the DOE JGI through the latter's Community Science Program.

This work was accomplished in partnership with two user facilities stewarded by the DOE Office of Biological and Environmental Research: the DOE JGI at Lawrence Berkeley National Laboratory (Berkeley Lab), and the Environmental Molecular Science Laboratory (EMSL) at Pacific Northwest National Laboratory. The study appeared online September 21, 2016, in the *Proceedings of the National Academy of Sciences*.

An artistic representation of the tree of life, with the many groups of bacteria on the left, the uncultivable bacteria at upper right (purple), and the Archaea and eukaryotes (green) — which includes humans — at the lower right. *(Graphic by Zosia Rostomian, Berkeley Lab Creative Services)*

Reorganizing the Tree of Life

In 1977, microbiologist Carl Woese proposed a third branch, Archaea, on the tree of life. In a study published April 11, 2016, in *Nature Microbiology*, and with help from the DOE JGI, a team led by longtime collaborator Jill Banfield of UC Berkeley featured a new depiction of the tree of life that better reflects the microbial diversity revealed through cultivation-independent techniques and bioinformatics methods. The team utilized the 30,437 eukaryote, bacterial, and archaeal genomes publicly accessible through the DOE JGI's Integrated Microbial Genomes (IMG) database, and 1,011 newly-reconstructed genomes from previously uncharacterized lineages.

The work builds off previous studies by the Banfield lab in which over 35 new groups or phyla of bacteria and nine new groups of archaea were identified. Many of the bacteria identified in the earlier study were sequenced as part of a 2015 DOE JGI Community Science Program (CSP) project that involved groundwater samples from a bioremediation site at Rifle, Colorado. This CSP project is also part of the Berkeley Lab Genomes-to-Watershed Scientific Focus Area (SFA).

With more accurate perspectives on microbial diversity and their relationships, scientists can make better inferences about adaptations to different environments, which can lead to discoveries of new enzymes and pathways to help address DOE missions in bioenergy and environmental processes.

Surface Microbes and Deep Subsurface Colonies

Microbes play key roles in maintaining the planet's biogeochemical cycles. Hydraulic fracturing, or "fracking," a technique to extract natural gas, can impact deep subsurface and salt tolerant microbial communities. While the economic benefits and environmental impacts of fracking are being studied, less attention has been focused on the impacts to the deep subsurface and the microbial communities therein. To answer that question, a team led by The Ohio State University's Kelly Wrighton compared microbial communities from the Marcellus shale in the Appalachian Basin, which is expected to produce triple the amount of natural gas sourced from any other shale formation, and from the Utica, a shale formation in Ohio.

For the first time, the team reconstructed 31 microbial genomes — including one from a novel bacterial genus they dubbed "Frackibacter" — from the fractured shale and studied their metabolic interactions for nearly a full

What a site looks like during the drilling process. *(Courtesy of the MSEEL [Marcellus Shale Energy and Environment Laboratory www.mseel.org], where the Wrighton Lab is also conducting research.)*

year. This work was accomplished in partnership with two user facilities stewarded by the Office of Biological and Environmental Research within DOE's Office of Science: the DOE JGI and EMSL through a collaborative science initiative called Facilities Integrating Collaborations for User Science (FICUS).

As fracking involves injecting freshwater along with chemicals and other fluids deep into the earth, the team found that microbes from the surface were also being injected and colonizing the deep subsurface, 2.5 kilometers underground, and thus affecting the resident microbial community in unknown ways. Watch Wrighton discuss her FICUS project at the DOE JGI 2016 Annual Genomics of Energy and Environment Meeting: **http://bit.ly/JGI2016Wrighton**. The study appeared online September 6, 2016, in *Nature Microbiology* and was recognized as one of *Discover* magazine's Top 100 Science Stories of 2016.

Microbial Evolution in Freshwater Lakes

A team led by University of Wisconsin scientists was able to reassemble the genomes of dozens of freshwater microbes and follow how their genetic makeup changed over time. The team's results show for the first time that different populations coexisting in the same environment likely have dramatically different rates of genetic exchange, and that these differences allow populations to evolve according to different non-exclusive theoretical models. This was the first conclusion to come out of a project approved under the DOE JGI's 2011 Community Science Program portfolio and was reported ahead online January 8, 2016, in *The ISME Journal*.

The team found different evolutionary processes at play within different microbial populations living in the same environment. For example, the team found one bacterial population undergoing a genome-wide selective sweep, a process predicted by the most prominent microbial evolutionary model (i.e., the ecotype model), but never before observed in the wild.

During a genome-wide selective sweep, a single member of a population out-competes all others once it acquires an advantageous trait, thus purging genetic diversity within the population. In contrast, the team also found other bacterial populations that appear to experience gene-specific sweeps, indicating that high rates of genetic interchange within these populations preserved their diversity. During a gene-specific sweep, only the small piece of DNA encoding the advantageous trait is shared with other members of the population through genetic recombination, thereby preventing a takeover by a single dominant strain.

Researchers studied peat soils extracted from peatland in Bavaria similar to this fen, also located in Germany. *(Paul Schulze, Flickr, CC-BY 4.0)*

Rare Sulfate Reducers Affect Peatlands Methanogenesis

Peatlands store almost a third of the world's terrestrial carbon, and there is a risk that some of this carbon could be released as global temperatures rise. Although the sulfur cycle of these peatlands has not been well studied, it plays a role in mitigating methane production from this ecosystem.

A research team led by Bela Hausmann of the University of Vienna and including DOE JGI researchers took samples of acidic peat soil from Germany, incubated them with various amounts of additional sulfate, and looked at how the sulfate levels impacted methane production in these artificial peat microcosms. As reported in the study published online March 25, 2016, in *The ISME Journal*, through sequencing metagenomic DNA and metatranscriptomic RNA, the researchers identified low-abundance microbes that respond to the presence or absence of sulfate and play roles in regulating methane production in these ecosystems.

For example, the researchers found that methane production in sulfate-stimulated microcosms was reduced by 83 to 100 percent compared with the control peat microcosms with no added sulfate. These results emphasize the importance of the rare microbial biosphere not only as a reservoir of dormant microorganisms, but also as active participants in biogeochemical processes that buffer against climate change. Understanding these rare microbes and their roles as mediators of biogeochemical processes offers researchers insights into both the characterization of microbial ecological functioning and into mitigating global climate change.

Improving Microbial Community Characterization

Microbes play key roles in maintaining the planet's biogeochemical cycles, but only a fraction of them have been characterized. Characterizing and classifying these species make use of 16S rRNA gene sequences, which represent a phylogenetic marker commonly used to assign all life into a particular classification. Characterizing the planet's microbial diversity provides researchers with more information about the roles these microscopic organisms play in maintaining the planet's biogeochemical cycles, and provides the DOE with insights into strategies for remediating environmental contamination. In the last decade, the ascension of high throughput,

short-read sequencing platforms over the more costly but tried-and-true Sanger sequencing platform has allowed researchers to characterize tens of thousands of microbes, but at the cost of generating microbial community profiles at lower taxonomic resolution.

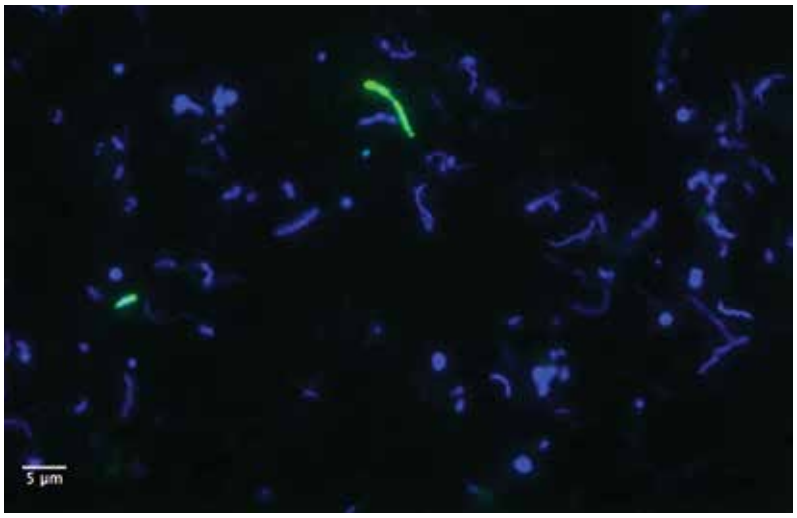
In a study that appeared online February 9, 2016, in *The ISME Journal*, DOE JGI researchers describe an approach called "PhyloTags" that harnesses the long reads generated by the Pacific Biosciences sequencing platform. PhyloTags are then evaluated against the iTags generated from Illumina's high-throughput, short-read sequencing platform. The researchers found PhyloTags provide a reliable adjunct or alternative to cost-effective iTags, enabling more accurate phylogenetic resolution of microbial communities and predictions on their metabolic potential. Specifically, some microbial genera involved in nitrogen and methane cycling across the lake's water column could only be resolved with PhyloTags, being missed by the short-fragment iTags. The team describes its work as the first benchmark study to offer a comprehensive comparison between these two sequencing approaches, allowing other researchers to see how using one platform over another impacts the characterization of a microbial community. Each approach has advantages, but each comes with costs as well. This work clarifies the benefits of each, enabling scientists to select the approach best suited to their experimental needs.

New Microbial Lineages Found in Hot Springs

Although global microbial populations are orders of magnitude larger than nearly any other population in, on, or around the planet, only a fraction have been identified thus far. The DOE seeks to uncover the true extent of the planet's microbial diversity in order to learn more about the genes, enzymes, and metabolic pathways that play key roles in regulating critical biogeochemical cycles.

In a study published January 27, 2016, in *Nature Communications*, a team led by DOE JGI utilized the largest collection of metagenomic datasets to uncover a completely novel bacterial phylum that the researchers have dubbed "Kryptonia." The team started with 5.2 trillion bases (Terabases or Tb) of sequence in the Integrated Microbial Genomes with Microbiome Samples (IMG/M) system. The team then identified long sequences that contained a phylogenetic marker (DNA corresponding to ribosomal RNA, rRNA) commonly used to assign all life (bacteria, archaea, and eukaryotes) into a particular classification system.

The team identified sequences from four different geothermal springs that could not be placed into any recognizable phylum. Reconstructing the genomes from metagenomic datasets and single cell genomes yielded four lineages belonging to the novel candidate phylum, named Kryptonia (*Candidatus Kryptonia*) from the Greek word for "hidden."



The novel candidate bacterial phylum Kryptonia. *(Composite image by Emiley Eloe-Fadrosh, DOE JGI)*

Given that there are currently 35 cultured bacterial and archaeal phyla, and roughly the same number of recognized uncultured phyla, study first author Emiley Eloe-Fadrosh said the identification of a novel candidate phylum was a surprise. "It's not every day that you find a completely new phylum. With all the studies that have been conducted in hot springs, there's an assumption that all novelty has been found. But we found these unknown lineages in high abundance."

Analyses of Kryptonia reveal that the bacteria need to rely on other microbes for several nutritional requirements, suggesting a reason this candidate phylum had not been found previously despite its abundance in geothermal springs.

In a complementary article published February 1, 2016, in *Nature Microbiology*, DOE JGI researchers wondered how many other microbial lineages might have been missed, as Kryptonia had been, due to being overlooked by the polymerase chain reaction (PCR) primers most commonly used in these kinds of analyses.

With help from the National Energy Research Scientific Computing Center (NERSC), the team analyzed more than 50,000 16S genes — gene sequences that are found in every microbe and can serve as an identifying marker — from metagenomic datasets and were able to find that as much as 10 percent of those sequences would be missed by the currently used PCR primers. The team then asked how many of those sequences might be unaffiliated with currently known microbial lineages and found what DOE JGI Prokaryote Super Program Head Nikos Kyrpides described as "a significant number of genes that could represent novel microbial phyla."

The team was also able to pinpoint specific habitats that could harbor most of the novel sequences. In addition to revealing more of the "microbial dark matter," the vast realm of as-yet-undiscovered microbes on Earth, this work will help better characterize the responses of microbial life to environmental changes.

Contributions of Microbial Communities Towards Mat Formation

Microbial mats serve as model systems for studying microbial interactions and their influence over biogeochemical processes. Understanding how communities of microbes establish mats over time provides insights on how their environments determine the mechanisms they employ. Microbes that establish mats thrive in environments like Yellowstone Hot Springs, and have been part of studies conducted at the DOE JGI for their bioenergy and environmental applications.

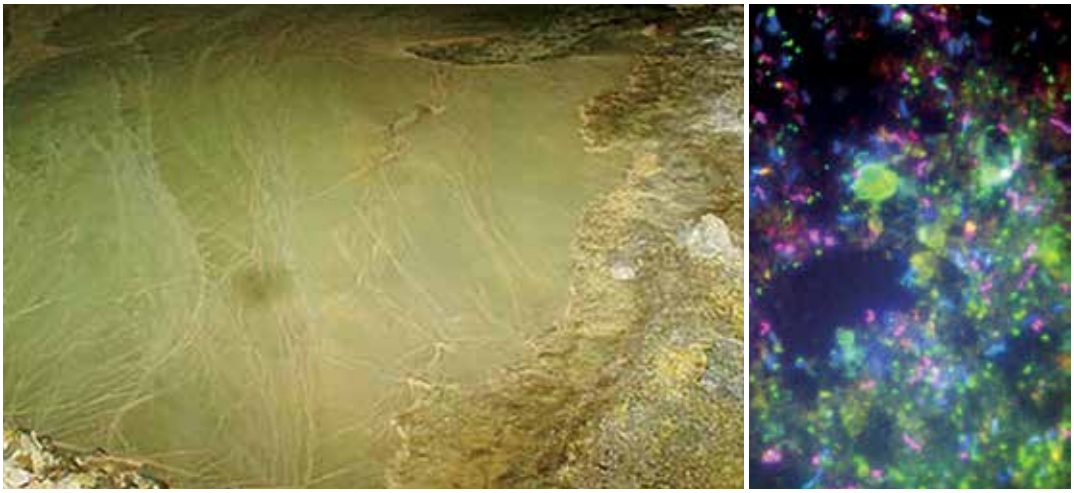
Originally published in *Frontiers in Microbiology* on February 16, 2016, the research, which was conducted by a team of scientists from the Pacific Northwest National Laboratory and DOE JGI, developed a conceptual model that details how microbial mats are formed in hot, acidic springs in the Yellowstone caldera. The team sequenced DNA samples extracted from two acidic geothermal springs at various time points over two months in Norris Geyser Basin at Yellowstone National Park. The data allowed the team to track the formation of microbial mats, beginning with primary colonization by *Hydrogenobaculum* species and *Metallosphaera yellowstonensis*, and how these populations as well as those of other microbes that colonized later changed over time in response to availability of nutrients such as oxygen and carbon.

These studies continue to build on the decades of microbial field studies senior author Bill Inskeep of Montana State University and his team have done at Yellowstone National Park. The insights gained from this model, the team noted, could provide insights into microbial life at other hot springs ecosystems and, potentially, on other planets.

Linking Phages to Hosts in Ecological Studies

Interactions between uncultivated microbes and their phage influence the local ecosystem as well as each other. As microbes play crucial roles in geochemical cycles and environmental processes, DOE researchers want to learn more about their diversity and understand their activities.

One method employed by longtime DOE JGI collaborator Jill Banfield and her team at UC Berkeley involves looking at Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) loci, a microbial immune system that is known to protect bacteria from being infected by viruses. Banfield's team can look for CRISPR spacer sequences in host microbes and link them back to the infective phages, as well as distinguish ancient from more recent encounters based on positioning within the locus.



Left: In this 2004 image, a near-continuous sheet of biofilm is seen over the surface of the acid mine drainage pool in Richmond Mine at Iron Mountain, near California's Mount Shasta. Right: FISH micrograph of mine communities with archaea and bacteria in blue. *(Courtesy of Banfield Lab, UC Berkeley)*

In a study published September 22, 2015, in *The ISME Journal*, Banfield and her team evaluated their ability to look at the genetic diversity and population histories of microbial communities using sequence datasets developed in collaboration with the DOE JGI. The team developed tools to analyze the CRISPR sequences in the datasets, including a custom script that extracts the CRISPR sequences from the metagenome sequencing reads. Specifically, they were able to identify spacers from *Leptospirillum* bacteria in the datasets and examine the spectrum of phage the bacteria had encountered over time. They also found two different CRISPR systems in the metagenomic datasets.

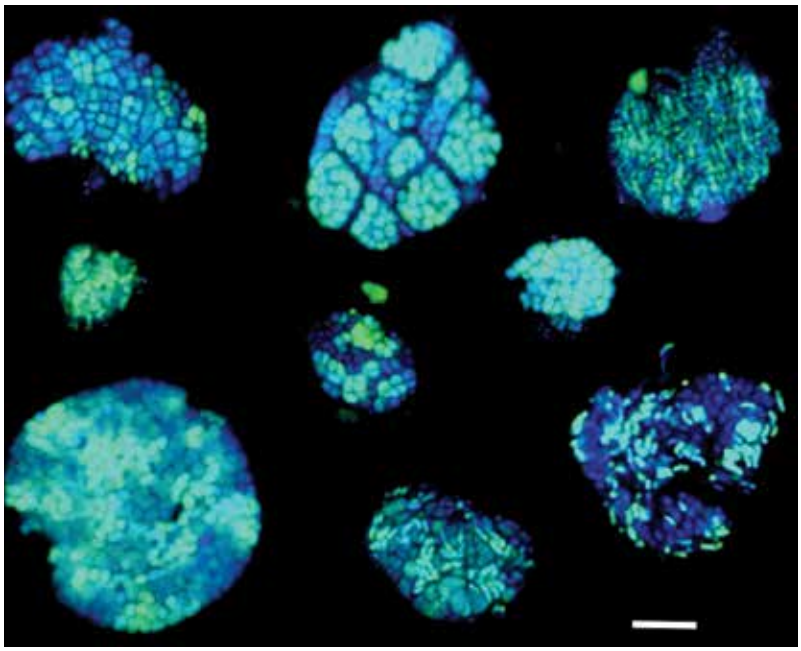
The Possibility of Other Genetic Codes

Yale University and DOE JGI researchers have discovered that microorganisms recognize more than one codon for selenocysteine. Their finding, published March 16, 2016, in the journal *Angewandte Chemie International Edition,* lends credence to recent studies indicating that an organism's genetic vocabulary is not as constrained as had been long held.

The combined team scanned trillions of base pairs of public microbial genomes and unassembled metagenome data in the National Center for Biotechnology Information and the DOE JGI's Integrated Microbial Genomes (IMG) data management system to find stop codon reassignments in bacteria and bacteriophages. Delving into genomic data from uncultured microbes afforded researchers the opportunity to learn more about how microbes behave in their natural environments, which in turn provides information on their management of the various biogeochemical cycles that help maintain Earth.

From approximately 6.4 trillion bases of metagenomic sequence and 25,000 microbial genomes, the team identified several species that recognize the stop codons UAG and UAA, in addition to 10 sense codons, as acceptable variants for the selenocysteine codon UGA.

The results, the team reported, "open our minds to the possible existence of other coding schemes… Overall our approach provides new evidence of a limited but unequivocal plasticity of the genetic code whose secrets still lie hidden in the majority of unsequenced organisms."



The depicted cell consortia are composed of archaea and bacteria that, by combining their individual metabolic powers, together achieve oxidation of the greenhouse gas methane in the absence of oxygen. For the first time, these consortia have now been analyzed for their protein synthesis activity under a variety of conditions. In this image, nucleic acids within individual cells are stained in blue. Green color indicates that new proteins have been made. The scale bar equals 10 micrometer, or 1/8th the width of a human hair. *(Roland Hatzenpichler)*

New Techniques to Visualize Uncultured Microbial Cell Activity

In a study published online June 28, 2016, in the *Proceedings of the National Academy of Sciences*, researchers at the California Institute of Technology (Caltech) and the DOE JGI utilized a recently refined technique called Bio-Orthogonal Non-Canonical Amino Acid Tagging (BONCAT) to identify both individual active cells, and single clusters of active bacteria and archaea within microbial communities. Many uncultured microbes play unknown roles in regulating Earth's biogeochemical processes — from regulating plant health to driving nutrient cycles in both terrestrial and marine environments — which can impact global climate. Understanding the true scope of the planet's microbial diversity is of interest to the DOE in order to learn how they can be harnessed for a wide range of energy and environmental challenges.

While researchers are harnessing multiple approaches to identify these microbes, referred to as "microbial dark matter," and determine what these microbes do, most techniques don't allow them to do both at once. Short for BioOrthogonal Non-Canonical Amino Acid Tagging, BONCAT is a technique developed at Caltech for bioengineering studies, but was adapted by Roland Hatzenpichler and Victoria Orphan for use in microbial ecology investigations. BONCAT uses synthetic amino acids to label protein-making cells. For this study, Hatzenpichler and his colleagues used sediment samples collected from deep-sea methane seep sediments off the coasts of Oregon and California for a series of incubation experiments, tracking the slow growth of methane-metabolizing bacterial and archaeal populations. The microbial communities in these sediments include aggregates of methane-oxidizing archaea called ANME (for ANaerobic MEthanotrophs) and sulfate-reducing bacteria that live together symbiotically and help to remove some 80 percent of the methane released from ocean sediments.

The Caltech researchers combined BONCAT with fluorescent *in situ* hybridization to analyze active microbes within these simulated environments and identify under which conditions they were active. DOE JGI researchers then helped develop a process that allowed the team to apply BONCAT to flow cytometry. Using BONCAT-Fluorescence Activated Cell Sorting, the team found that the methane-oxidizing archaea weren't interacting with just the known sulfate-reducing bacteria, as had been expected, but were also sometimes associating with members of the less well-studied Verrucomicrobia phylum. These interactions had not been seen previously, and the finding suggests to researchers that methane-oxidizing archaea have a broader range of symbiotic relationships than had been thought.

Zostera marina meadow in the Archipelago Sea off Finland, from which the sequenced eelgrass sample was taken. *(Christoffer Bostrom)*

Eelgrass Genome Sequence Offers Salt Tolerance Insights

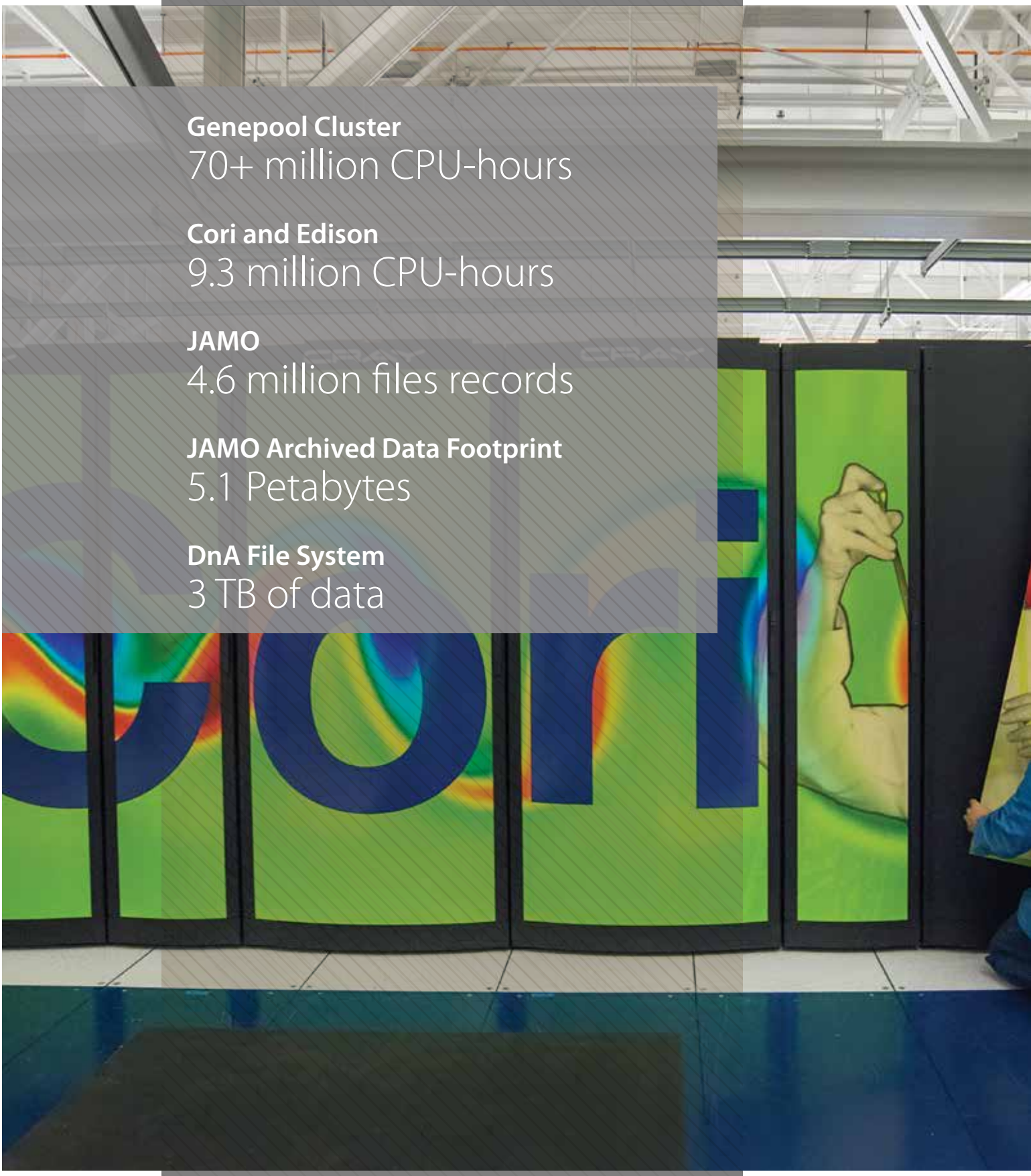
Coastal seagrass ecosystems cover some 200,000 square kilometers, account for an estimated 15 percent of carbon fixed in global ocean, and and affect sulfur and nitrogen cycles.

Appearing on the February 18, 2016 cover of the journal *Nature*, a study by a European team including researchers from the DOE JGI sequenced a seagrass genome: that of the eelgrass *Zostera marina*, taken from the Archipelago Sea off Finland. *Z. marina* is the first marine flowering plant to be fully sequenced, work done through the DOE JGI's Community Science Program. To better understand the adaptations the plant made in returning to a saltwater environment, the team compared the eelgrass genome to its freshwater relative, Greater duckweed (*Spirodela polyrhiza*). The duckweed genome was also sequenced and analyzed by the DOE JGI.

The team compared the eelgrass genome with duckweed, one of the simplest flowering plants and *Z. marina*'s closest sequenced relative. The team noted differences in genes related to cell wall structure due to adaptations to freshwater or terrestrial conditions. For example, plants such as duckweed have seemingly lost genes that help plants retain water in the cell wall, while eelgrass has regained these genes to better deal with osmotic stress at low tide.

"They have re-engineered themselves," said study lead author Jeanine Olsen of the University of Groningen of the changes affecting the eelgrass cell walls. "Crop breeders may benefit from lessons on how salt tolerance has evolved in these plants."

Learning more about eco-evolutionary interactions is also relevant to the development of genomics-based, early-warning indicators that may foreshadow seagrass ecosystem collapse. Jeremy Schmutz, head of the DOE JGI's Plant Program, emphasized that while eelgrasses are key players in coastal marine ecosystem functions and are considered the "lungs of the sea," they are also endangered. "There are estimates that nearly a third of the eelgrass meadows worldwide have been destroyed by runoff into the ocean," he said, "reducing their potential capabilities as carbon sinks. Thus, studying the adaptive capacity of eelgrass is urgent to assist conservation efforts."



Genepool Cluster
70+ million CPU-hours

Cori and Edison
9.3 million CPU-hours

JAMO
4.6 million files records

JAMO Archived Data Footprint
5.1 Petabytes

DnA File System
3 TB of data

Computational Infrastructure

The Department of Energy Joint Genome Institute's (DOE JGI) capacity as a next-generation genomics user facility has generated petabytes of high-quality sequence data and analysis. To support this workload, the DOE JGI has invested significant resources in its high-performance compute cluster, Genepool, as well as its storage and web infrastructure.

The success of these computing projects is in part due to the DOE JGI's ongoing partnership with the National Energy Research Scientific Computing Center (NERSC), one of the nation's foremost centers for high-performance computing. In 2010, all of the DOE JGI's computational resources were moved to NERSC, and both sides have learned a great deal through this partnership. The infrastructure advancements to Genepool and other DOE JGI portals mean rapid and smooth access for users across the globe. Partnering with NERSC enables DOE JGI researchers and users to devote more of their time to cutting-edge genomics research.

The DOE JGI used several million central processing unit hours on NERSC's petascale supercomputers, Edison and Cori. Many of these calculations could not have been completed on the Genepool cluster because of the computing scale required. In 2016 NERSC accepted Cori, a $70 million investment in data-intensive and high-performance computing infrastructure. One of the core new features of Cori is new hardware called the Burst Buffer, an array of non-volatile random-access memory (NVRAM) nodes that will enable input/output-intensive workloads to run at scale on the new system. NERSC also deployed new software, called Shifter, to allow Docker containers to run on the Cray supercomputers. This work grew out of collaboration with the DOE JGI staff and has reduced the hurdles required to move bioinformatics, light source, and astronomy workloads to Cori.

The DOE JGI and NERSC co-organized a workshop at the DOE JGI Genomics of Energy and Environment Meeting in 2016. In this workshop, users learned about the Docker technology and got hands-on experience running containers in the cloud. In 2017, the DOE JGI and NERSC user communities will be able to apply for the first Facilities Integrating Collaborations for User Science (FICUS) program. This joint call issued by the two DOE user facilities will give scientists access to state of the art computing and large quantities of high quality genomic data. The workshop was successful and will be held again at the 2017 DOE JGI User Meeting.

In 2015, NERSC moved to a new, state of the art facility on the main Berkeley Lab site. The minimally disruptive move of the DOE JGI's computational infrastructure was completed in early 2016. The high-speed link between the two facilities enabled the transfer of the 3+ petabytes of DOE JGI data from Oakland to Berkeley without any disruption to the DOE JGI staff. NERSC's new home, the Computational Research and Theory (CFT) Facility will be on the forefront of high-performance supercomputing research and will be DOE's most efficient facility of its kind. Designed to take advantage of the cool Berkeley climate, the CRT facility is anticipated to set a new standard in energy efficiency for high-performance computing.

Appendices

Appendix A
Acronyms at a Glance

AF	Alignment fraction
ANI	Average Nucleotide Identity
ANME	ANaerobic MEthanotrophs
BER	DOE Office of Biological and Environmental Research
BERAC	Biological and Environmental Research Advisory Committee
BESC	BioEnergy Science Center
BLAST	Basic Local Alignment Search Tool
BONCAT	BioOrthogonal Non-Canonical Amino Acid Tagging
BRC	Bioenergy Research Center (i.e., BESC, GLBRC, JBEI)
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRT Facility	Computational Research and Theory Facility
CSP	Community Science Program
DOE	Department of Energy
EMSL	Environmental Molecular Sciences Laboratory (at PNNL)
ETOP	Emerging Technologies Opportunity Program
FACS	Fluorescence Activated Cell Sorting
FICUS	Facilities Integrating Collaborations for User Science
FISH	Fluorescent *In Situ* Hybridization
GEBA	Genomic Encyclopedia of Bacteria and Archaea
GLBRC	Great Lakes Bioenergy Research Center
GOLD	Genomes OnLine Database
HPC	High Performance Computing
HPPS	High Performance Storage System
IMG	Integrated Microbial Genomes data management system
ISM	Integrated Safety Management
ITS	Integrated Tracking System
JAMO	JGI Archive and Metadata Organizer
JBEI	Joint BioEnergy Institute
KBase	DOE Systems Biology Knowledgebase
LANL	Los Alamos National Laboratory
LBNL	Lawrence Berkeley National Laboratory
LLNL	Lawrence Livermore National Laboratory
MGM	Microbial Genomics & Metagenomics
NERSC	National Energy Research Scientific Computing Center
NGEE	Next-Generation Ecosystem Experiments
NREL	National Renewable Energy Laboratory
ORNL	Oak Ridge National Laboratory
PMO	Project Management Office
PNNL	Pacific Northwestern National Laboratory
SAC	Scientific Advisory Committee
SAG	Single Amplified Genome
SFA	Scientific Focus Area
WIP	Work Initiation Process

Appendix B
Glossary

Annotation: The process of identifying the locations of genes in a genome and determining what those genes do to improve accuracy of genetic information collected.

Archaea: One of the three domains of life (Eukarya and Bacteria being the others) that include primitive microorganisms that can tolerate extreme environmental conditions (temperature, acid, etc.).

Assembly: Aligning and merging fragments of a much longer DNA sequence in order to reconstruct the original sequence. This is required, as DNA sequencing technology cannot read whole genomes at once, but rather reads small pieces of between 20 and 1,000 bases, depending on the technology used.

Barcoding: The practice of appending known unique synthetic DNA sequences to sequencing libraries to allow pooling of libraries for next-generation sequencing, after which sequence data can be assigned to particular libraries or samples based on the barcode sequence.

Base: A unit of DNA. There are four bases: adenine (A), guanine (G), thymine (T), and cytosine (C). The sequence of bases constitutes the blueprint of life.

Base pair: Two DNA bases complementary to one another (A and T or G and C) located on the complementary strands of the DNA double helix.

Biogeochemistry: The field of study of the biosphere's interactions with the Earth's chemical environment.

Bioinformatics: The use of computers to collect, store, and analyze biological information.

Biomass: Material derived from living or recently living organisms, usually referring to plants or plant-derived material (lignocellulosic biomass). Biomass can serve as an energy source directly by burning or indirectly, after conversion into biofuels.

Bioprospecting: Searching nature for genes and proteins that can be applied to help scientists solve energy and environment challenges.

Bioremediation: The use of microorganisms to break down contaminants and other unwanted substances in waste and other substances.

Bioscriber: A scientist who uses synthetic biology to coax microbes into producing helpful compounds such as biofuels or antibiotics.

Bridge amplification: A proprietary technique used by Illumina sequencing platforms to generate single-stranded clusters of template DNA.

Carbon cycle: The biogeochemical process by which carbon is exchanged among the planet's atmosphere, land, and oceans.

CAZymes: Carbohydrate-active enzymes that can break down plant polysaccharides such as cellulose into small sugars.

Cellulose: An organic compound made of a long chain of several hundred to over 10,000 glucose units. It is a critical part of the cell wall of plants and many algae.

Cellulosic biofuel: A type of liquid transportation fuel produced from lignocellulose,a structural material that makes up much of the mass of plants. Lignocellulose is composed mainly of cellulose, hemicellulose, and lignin.

ChIP-Seq: A method of analyzing protein interactions with DNA.

Contig: A contiguous sequence resulting from the assembly of smaller sequence fragments.

Coverage: The number of times a region of the genome has been sequenced during whole-genome shotgun sequencing.

Curation: Analysis of genome annotations to improve and maintain data presentation.

Cyanobacteria: A phylum of bacteria that obtain their energy through photosynthesis, and named for the color of the bacteria. Although often called blue-green algae, that name is a misnomer as cyanobacteria are prokaryotic and algae are eukaryotic.

Draft genome (also called a draft assembly): The term for an incomplete genome sequence. It can be applied to a wide range of sequences, from those that have the minimum amount of information needed for submission to a public database, to assembled genomes that have undergone manual and automatic review but still have sequence errors that need to be corrected.

Enzyme: A protein used to induce or speed up a chemical reaction.

Eukaryotes: The domain of life containing organisms that consist of one or more cells, each with a nucleus and other well-developed intracellular compartments. Eukaryotes include all animals, plants, and fungi.

Finished genome: In accordance with the 1996 Bermuda standard, a gapless sequence with a nucleotide error rate of one or less in 10,000 bases.

Flow cell: Resembles a microscopic slide, only with eight channels, on which DNA samples are loaded for analysis on Illumina sequencing platforms.

Fluorescence-activated cell sorting: A specialized type of flow cytometry used to study and purify cells. A heterogeneous mixture of cells passes through laser beams and is sorted into two or more containers, one cell at a time, based upon the specific light-scattering and fluorescent characteristics of each cell.

Fosmid: A vector suitable for cloning genomic inserts approximately 40 kilobases in size.

GenBank: Open-access, publicly available collection of annotated sequences submitted by individual laboratories and large-scale sequencing centers that is overseen by the National Center for Biotechnology Information.

Halophile: A microbe that thrives in environments with high salinity.

Hemicellulose: An organic compound that is part of most plant cell walls and is made of 5-carbon sugars. Unlike cellulose, which is crystalline, strong, and resistant to being broken down, hemicellulose is much more fragile, and has a random structure.

Informatics: The science of information and computer information systems. At the DOE JGI, it is the science of managing and interpreting genomic information with computational tools.

Library: A collection of DNA fragments.

Lignin: A complex polymer of aromatic alcohols known as monolignols, usually derived from wood. It is a critical part of the cell wall of plants and many algae.

Lignocellulosic biomass: Biomass derived from plants, the most abundant raw material for the production of biofuels.

Locus (plural loci): The specific location of a gene or DNA sequence or position on a chromosome.

Mapping: Charting the location of genes on chromosomes.

Mass spectrometry: An analytical technique that can identify unknown compounds through their molecular weight. It can also be used to determine a molecule's structure and chemical properties.

Metabolomics: A comparison of biological samples based on their metabolite profiles.

Metagenomics (also environmental genomics or community genomics): The study of genomes recovered from environmental samples through direct methods that do not require isolating or growing the organisms present. This field of research allows the genomic study of organisms that are not easily cultured in a laboratory.

Metatranscriptomics: The study of the region of the complete genetic code that is transcribed into RNA molecules and provides information on gene expression and gene function.

Microbe: Another name for a microorganism.

Microbiome: A defined environment within which a community of microbes exists and interacts.

Molecular cloning: The use of specialized DNA technology to produce multiple exact copies of a single gene or other segment of DNA to obtain enough material for further study.

Multiple displacement amplification (MDA): Method of amplifying tiny amounts of DNA in a cell so that it can be used for sequencing through single-cell genomics.

Nitrogen cycle: The biogeochemical process by which nitrogen is exchanged among the planet's atmosphere, land, and oceans.

Paired-end reads: DNA library preparation technique that lets researchers look at both the forward and reverse template strands of a large DNA fragment and that provides positional information.

Peptide: Short chain of amino acids, the same compounds that make up proteins. Peptide chains are much shorter than the chains of amino acids that make up proteins.

Phylogeny: The evolutionary history of a molecule such as a gene or protein, or a species.

Polymerase chain reaction (PCR): A method of DNA amplification.

Prokaryotes: Unlike eukaryotes, these organisms, (e.g., bacteria) are characterized by the absence of a nuclear membrane and by DNA that is not organized into chromosomes.

Promoter: A region of DNA that sends signals to a cell to tell it where a gene begins and when the gene is read. An inducible promoter only signals the cell under certain conditions while a constitutive promoter is always signaling the cell.

Proteomics: The large-scale study of proteins, as well as their structures and functions.

Psychrophile: A cold-loving microbe that optimally grows in environments with temperatures of 15°C (60°F) or less.

Read length: The number of nucleotides tabulated by the DNA analyzer per DNA reaction well.

Rhizosphere: Microecosystem defined by a thin layer of soil where plant roots interact with microorganisms in the soil.

Selfing: Self-pollination or self-fertilization.

Sequence: Order of nucleotides (base sequence) in a nucleic acid molecule. In the case of DNA sequence, it is the precise ordering of the bases (A, T, G, C) from which the DNA is composed. Also used as a verb to describe the process of determining the nucleotide order.

Sequencing by synthesis: Proprietary sequencing technique used by Illumina systems in which four fluorescently labeled nucleotides determine the sequence of a DNA fragment, one base at a time.

Single-cell genomics: Method for sequencing a genome using DNA derived from a single cell that is used to study uncultured or nonculturable organisms.

Single-molecule real-time (SMRT) sequencing: Single-molecule DNA sequencing performed in zero-mode waveguide (ZMW) chambers on a chip.

Subcloning: The process of transferring a cloned DNA fragment from one vector to another.

Sulfur cycle: The biogeochemical process by which sulfur is exchanged between the planet's atmosphere, land, and oceans.

Synthetic biology: A field of research concerned with purposeful editing of biological systems. For the DOE JGI's objectives, this process refers to assembling DNA sequence fragments with the goal of synthesizing sequences to experimentally validate their functions and applications.

Transcriptome: A collection of all the RNA transcripts in a given cell that serves as a snapshot of global gene expression.

Appendix C
2016 User Program Supported Projects

Community Science Program Projects

INVESTIGATOR	AFFILIATION	DESCRIPTION
Aitken, Karen	CSIRO (Australia)	Understanding polyploidy through the generation of the first sugarcane genome sequence
Attwood, Graeme	AgResearch Ltd (New Zealand)	Defining gene function in rumen microbes
Bohlmann, Joerg	University of British Columbia (Canada)	Exploring the G3 "Gymnosperm Giga-Genomes" for carbon sequestration, biofuels, and bioproducts
Bonito, Gregory	Michigan State University	Functional gene network regulation in tripartite plant-fungal-bacterial mutualisms of bioenergy plants and algae
Bruns, Thomas	University of California, Berkeley	Functional genomics of pyrophilous fungi — determining the fate of pyrolyzed carbon in post-fire soils
Buchan, Alison	University of Tennessee Knoxville	Geo-metabolomics of a saltmarsh: Combining *in situ*, bulk, genomic, transcriptomic, and DOMeomic data streams
Coleman-Derr, Devin	USDA-ARS	Exploring the role of drought-induced plant associated microbes in promoting plant fitness in sorghum bicolor and *Oryza sativa*
Cox, Michael	University of Wisconsin-Madison	Defining the molecular basis of extreme resistance to ionizing radiation
Crowe, Sean	University of British Columbia (Canada)	Mapping the global methanome
Cullen, Daniel	Forest Products Laboratory	Metatranscriptome analysis of fungal decay of *Pinus contorta*
Cushman, John	University of Nevada	Ice plant gene atlas resource development for *Mesembryanthemum crystallinum*, a facultative crassulacean acid metabolism (CAM) model for improved water-use efficiency of bioenergy feedstocks
Des Marais, David	Harvard University	Perenniality, abiotic stress tolerance, and biomass allocation in *Brachypodium*, a model grass genus for bioenergy
Dijkstra, Paul	Northern Arizona University	Stress in microbial communities in response to changes in carbon and nitrogen availability
Dopson, Mark	Linnaeus University (Sweden)	Exploring deep biosphere microbial communities by single-cell DNA sequencing

Community Science Program Projects *(continued)*

INVESTIGATOR	AFFILIATION	DESCRIPTION
Doty, Sharon	University of Washington	Functional genomics of poplar endophytes for elucidation of mechanisms of improved plant growth under challenging conditions
Dudycha, Jeffry	University of South Carolina	Unlocking the photosynthetic diversity of cryptophyte algae through whole-genome sequencing
Francis, Christopher	Stanford University	Spatiotemporal characterization of microbial communities controlling estuarine nitrogen and carbon cycling in the San Francisco Bay Delta
Göker, Markus	DSMZ (Germany)	The One Thousand Microbial Genomes Phase 4 Project (KMG-4) — sequencing the most valuable type-strain genomes for metagenomic binning, comparative biology, and taxonomic classification
Harris, Steven	University of Nebraska–Lincoln	Fungal interaction networks in biological soil crusts
Heyduk, Karolina	University of Georgia	Genome sequencing of C3 and CAM *Yucca* species
Johnson, Matthew	Woods Hole Oceanographic Institute	The role of acquired phototrophy in phytoplankton blooms: Insights from the *Mesodinium rubrum* genome
Kellogg, Elizabeth	Donald Danforth Plant Science Center	Pan-genomics of big bluestem, a broadly adapted dominant grass
LeBoldus, Jared	Oregon State University	RNAseq enabled metabolic modeling of disease resistance to Septoria canker in the DOE flagship *P. trichocarpa*
Lopez Peredo, Elena	Marine Biological Laboratory	Protecting photosynthesis during desiccation: Do the genomes of desert-derived and aquatic *Scenedesmus* species hold the key to understanding extreme desiccation tolerance among green algae?
MacGregor, Barbara	University of North Carolina at Chapel Hill	Single-cell (meta-)genomics of uncultivable large sulfur bacteria and their epibionts: investigating host-microbe mediation of biogeochemical cycling
Mockler, Todd	Donald Danforth Plant Science Center	A complete-sequence population for pan-genome analysis of sorghum
Moran, Mary Ann	University of Georgia	Dynamics of bacterial carbon and sulfur cycling in a coastal environment

Community Science Program Projects *(continued)*

INVESTIGATOR	AFFILIATION	DESCRIPTION
Nguyen, Nhu	University of Hawai'i at Manoa	A genome atlas of the ectomycorrhizal genus *Suillus*: Phylogenetic diversity and population genomics of a keystone guild of symbiotic forest fungi
Nicholson, Wayne	University of Florida	Transcriptomic and methylomic responses of *Carnobacterium* species to extreme low pressure
Pett-Ridge, Jennifer	Lawrence Livermore National Laboratory	Microbial carbon transformations in wet tropical soils: effects of redox fluctuation
Raff, Jonathan	Indiana University	Combined flux chamber and genomics approach to understanding soil emissions of reactive nitrogen oxides in a forested environment
Shade, Ashley	Michigan State University	Greater than the sum of its parts? A synthetic microbial community approach to untangle member interactions and exometabolite production
Smart, Christine	Cornell University	Genetic diversity of shrub willow pathogen *Melampsora americana* aided by genome sequence
Stepanauskas, Ramunas	Bigelow Laboratory for Ocean Sciences	Expanding the dark matter reference catalog by targeting taxonomic blind spots
Whitman, William	University of Georgia	Core and pangenomes of soil and plant-associated prokaryotes
Yang, Xiaohan	Oak Ridge National Laboratory	Gene atlas for *Kalanchoe laxiflora*, a obligate crassulacean acid metabolism (CAM) model for genetic improvement of water-use efficiency in bioenergy feedstocks
Zhang, Chi	University of Nebraska – Lincoln	Genome sequencing of Zygnematales, the closest algal lineage to land plants, as a foundation for comparative genomic, transcriptomic, epigenetic, evolutionary, and biochemical studies

Small-Scale Proposals

INVESTIGATOR	AFFILIATION	DESCRIPTION
Alvarez-Cohen, Lisa	University of California, Berkeley	Unraveling functional dynamics and regulation crucial for the stability of an anaerobic ammonium oxidizing (anammox) community via community metatranscriptomics and 16S rRNA sequencing
Andras, Jason	Mount Holyoke College	The effect of ecological restoration on the structure and function of soil microbial communities in coastal wetlands
Bowen, Jennifer	Northeastern University	Building a foundation for understanding carbon and nitrogen cycling in microbially diverse salt marsh sediments
Coleman, Maureen	University of Chicago	Teasing apart coexisting picocyanobacteria and their contributions to biogeochemistry
Cotner, James	University of Minnesota-Duluth	Translating stoichiometric diversity into genomic diversity: What elements are responsible for variability in bacterial biomass stoichiometry?
D'Agostino, Paul	Technische Universität München (Germany)	Genome mining and synthetic biology of underrepresented organisms in the search for novel (bio-)chemistry
Dove, Nicholas	University of California, Merced	Immediate effects of prescribed fire on microbial communities, decomposition, and nitrification
Dynarski, Katherine	University of California, Davis	Microbial community composition and function throughout decomposition in temperate forests across a bedrock nitrogen gradient
Freedman, Zachary	West Virginia University	Assessing the recovery of microbial traits in bioenergy crop agroecosystems on reclaimed surface mines
Hatzenpichler, Roland	Montana State University	Genomic characterization of cosmopolitan sediment-dwelling archaea hypothesized to be involved in anaerobic carbon cycling
Hatzinikolaou, Dimitris	University of Athens (Greece)	Genome sequencing of lignin and cellulose degrading isolates from Greek habitats
Hunt, Dana	Duke University	Seasonal and disturbance-related alterations in the biogeochemical cycling of estuarine carbon

Small-Scale Proposals *(continued)*

INVESTIGATOR	AFFILIATION	DESCRIPTION
Jensen, Paul	University of California, San Diego	Extracellular electron shuttles: a mechanism for obligate aerobes to survive anaerobic conditions
Lebeis, Sarah	University of Tennessee	Characterizing functional chemotaxis receptors in different root zones
Lindemann, Steve	Purdue University	Identifying genome properties and environmental conditions governing the assembly of stable cyanobacterial-heterotroph consorta
Mavrodi, Dmitri	University of Southern Mississippi	Transcriptomic responses of beneficial *Pseudomonas* rhizobacteria to *Brachypodium* root exudates
McMahon, Katherine	University of Wisconsin-Madison	Reference genomes for freshwater bacteria — alpha- and beta-proteobacteria
Müller, Henry	Graz University of Technology (Austria)	Elucidating the relevance of DNA methylation of strain-specific plant-microbe interactions
Oliveira, Rafael	State University of Campinas (Brazil)	Unravelling microbial communities associated with native plant species from P-impoverished soils of a global biodiversity hotspot
Pan, Chongle	Oak Ridge National Laboratory	Multi-cycle selection of sorghum microbiomes for biological nitrogen fixation
Richardson, Ruth	Cornell University	Metatranscriptomic responses of a butyrate-to-methane enrichment culture to sulfate and iron availability: elucidating the identities and strategies of metabolically versatile sulfate and iron reducing populations
Schmer, Marty	USDA-ARS	Nitrogen effects on soil microbial communities and soil organic carbon in a resilient bioenergy cropping system
Valentine, David	University of California, Santa Barbara	Metagenomic reconstruction of novel archaeal genomes from energy-limited hypersaline sediments
Veley, Kira	Donald Danforth Plant Science Center	Epigenetic and transcriptional changes that occur in populations of bacteria colonizing plant hosts
Venturi, Vittorio	International Centre of Genetic Engineering and Biotechnology (ICGEB) (Italy)	Sequencing of a set of identified and characterized rice bacterial endophytes

Synthesis Proposals

INVESTIGATOR	AFFILIATION	DESCRIPTION
Alper, Hal	The University of Texas, Austin	Systematic testing of gene sensitivities in yeast via a sgRNA library approach
Blaby, Ian	Brookhaven National Lab	Sequence to function: integrative high-throughput approaches for functional prospecting in photosynthetic organisms
Christen, Beat	ETH Zurich (Switzerland)	A high-throughput strategy to tame bacterial genomes
Erb, Tobias	Max-Planck-Institute for Terrestrial Microbiology (Germany)	SYNCOPe – realizing SYNthetic CO_2 fixation Pathways
Ghodge, Swapni	University of North Carolina at Chapel Hill	Enzyme function discovery as a platform for exploring novel bacterial secondary metabolites exploring the perception landscape of the strigolactone receptor
Hussey, Steven	University of Pretoria (South Africa)	High-throughput screening and reconstruction of transcriptional networks underlying secondary cell wall formation for lignocellulosic biomass improvement in Eucalyptus
Lewis, Jared	University of Chicago	Biorefining using phylogenetically diverse sets of enzymes and artificial metalloenzymes
Niyogi, Kris	University of California, Berkeley	Synthetic biology tools for engineering of *Nannochloropsis oceanica*
Oikawa, Hideaki	Hokkaido University (Japan)	Deciphering programming rule for fungal iterative polyketide synthases by functional analysis of individual catalytic domain
Ruffing, Anne	Sandia National Laboratories	Optimization of CRISPR-Cas9 tools for engineering *Nannochloropsis gaditana*
Santos, Christine	Manus Biosynthesis	Generating a library of bioactive natural compounds in engineered *E. coli*
Schmidt, Eric	University of Utah	Synthetic biological control of bacterial interactions
Wheeldon, Ian	University of California, Riverside	CRISPR-enabled genome-wide screening of the oleaginous yeast *Yarrowia lipolytica*
Zhang, Wenjun	University of California, Berkeley	DNA synthesis of antimycin-type depsipeptide genes to advance combinatorial biosynthesis capabilities and understand the natural roles of this family of secondary metabolites

JGI-EMSL Collaborative Science Initiative Proposals

The DOE JGI and the Environmental Molecular Sciences Laboratory (EMSL) accepted 10 projects submitted during the 2017 call for Collaborative Science Initiative proposals for their joint Facilities Integrating Collaborations for User Science (FICUS) initiative. The call represents a unique opportunity for researchers to harness the combined power of genomics and molecular characterization in one research project to help advance the missions of the Department of Energy's Office of Biological and Environmental Research. The accepted proposals began on October 1, 2016, providing the researchers with access to the capabilities of both user facilities and datasets beyond what could be generated by either facility alone.

JGI-EMSL collaborative FICUS proposals:

INVESTIGATOR	AFFILIATION	DESCRIPTION
Baldrian, Petr	Institute of Microbiology ASCR (Czech Republic)	The impacts of nitrogen availability and seasonal dynamics on plant-microbial interactions affecting C and N cycling in coniferous forest soils
Bartley, Laura	University of Oklahoma	Systems analysis of grass secondary cell wall development and regulation for biofuel production
Bianchi, Thomas	University of Florida	The role of priming effects on the conversion of blue carbon to CO_2 in the coastal zone
Blanchard, Jeffrey	University of Massachusetts Amherst	Molecular mechanisms underlying changes in the temperature sensitive respiration response of forest soils to long-term experimental warming
Cattolico, Rose Ann	University of Washington	Global warming induced salinity shifts: metabolic responses by algal-bacterial consortia
Fendorf, Scott	Stanford University	Metabolic constraints on organic matter decomposition and metal cycling in sediment deposits
Liao, Hui-Ling	Duke University	Combined 'omics approaches for the study of ectomycorrhizal symbiosis between *Suillus* and *Pinaceae*, with emphasis on their role in nutrient cycling
Rich, Virginia	The Ohio State University	Something old, something new: systems-level insights into plant-microbial-permafrost carbon dynamics by parallel high-resolution organic matter and microbial meta-omics
Skerker, Jeffrey	University of California, Berkeley	Understanding conversion of biomass-derived carbon into lipids and terpenoids in the oleaginous yeast *Rhodosporidium toruloides*
Wrighton, Kelly	The Ohio State University	Deciphering controls on plant decomposition in Arctic ecosystems: Identifying unknown microbial condensed tannin degradation pathways

Appendix D
Advisory and Review Committee Members

The Scientific Advisory Committee (SAC)

The Scientific User Advisory Committee is a board convened by the DOE JGI Director to provide a scientific and technical overview of the DOE JGI. Responsibilities of this board include providing technical input on large-scale production sequencing, new sequencing technologies, and related informatics; an overview of the scientific programs at the DOE JGI; and an overview of the Community Science Program (CSP). A crucial job of the committee is to take the input from the CSP Proposal Study Panel on prioritization of CSP projects and, with the DOE Office of Biological and Environmental Research (BER) concurrence, set the final sequence allocation for this program.

Members

Mark Adams, *The Jackson Laboratory (Chair)*

Carol Bult, *The Jackson Laboratory*

Steve Briggs, *University of California, San Diego*

Jeff Dangl, *University of North Carolina*

Claire M. Fraser, *University of Maryland*

N. Louise Glass, *Lawrence Berkeley National Laboratory*

Glenn Kubiak, *Lawrence Berkeley National Laboratory*

Trina McMahon, *University of Wisconsin-Madison*

Deirdre Meldrum, *Arizona State University*

Juan Meza, *University of California, Merced*

Mary Ann Moran, *University of Georgia*

Julian Parkhill, *The Sanger Institute*

Informatics Advisory Committee (IAC)

Members

Adam Arkin, *Lawrence Berkeley National Laboratory; University of California, Berkeley*

Judith Blake, *The Jackson Laboratory*

David Dooling, *Monsanto*

Paul Flicek, *European Molecular Biology Laboratory (EMBL)-European Bioinformatics Institute (EBI)*

Saul Kravitz, *The Howard Hughes Medical Institute (HHMI)*

Jill Mesirov, *University of California, San Diego (Chair)*

Granger Sutton, *J. Craig Venter Institute*

Cathy Wu, *Georgetown University*

Kathy Yelick, *Lawrence Berkeley National Laboratory*

Fungal Program User Advisory Committee

Members

Scott Baker, *Pacific Northwest National Laboratory*

Randy Berka, *Novozymes*

Ronald de Vries, *CBS (Netherlands)*

Audrey Gasch, *University of Wisconsin-Madison; Great Lakes Bioenergy Research Center*

N. Louise Glass, *University of California, Berkeley; Lawrence Berkeley National Laboratory*

Stephen Goodwin, *Purdue University*

David Hibbett, *Clark University*

Francis Martin, *INRA (France)*

Michelle O'Malley, *University of California, Santa Barbara*

Joseph Spatafora, *Oregon State University*

Kathleen Treseder, *University of California, Irvine*

Adrian Tsang, *Concordia University (Canada)*

Plant Program User Advisory Committee

Members

Siobhan Brady, *University of California, Davis*

Gloria Coruzzi, *New York University*

Jeff Dangl, *University of North Carolina*

Joe Ecker, *The Salk Institute for Biological Studies*

Samuel Hazen, *UMass Amherst*

Tom Juenger, *University of Texas, Austin*

Toby Kellogg, *Donald Danforth Plant Science Center*

Sabeeha Merchant, *University of California, Los Angeles*

Stephen Moose, *University of Illinois*

Sue Rhee, *Carnegie Institution for Science, Stanford*

Bob Schmitz, *University of Georgia*

Gary Stacey, *University of Missouri*

Prokaryotic Super Program Advisory Committee

Members

Jill Banfield, *University of California, Berkeley*

Cameron Currie, *University of Wisconsin*

Ed DeLong, *University of Hawaii at Manoa*

Jonathan Eisen, *University of California, Davis*

George Garrity, *Michigan State University (Chair)*

Steve Hallam, *University of British Columbia*

Phil Hugenholtz, *University of Queensland (Australia)*

Janet Jansson, *Pacific Northwest National Laboratory*

Monica Medina, *Penn State University*

Trina McMahon, *University of Wisconsin-Madison (Vice-Chair)*

Mary Ann Moran, *University of Georgia*

Nancy Moran, *University of Texas at Austin*

Victoria Orphan, *California Institute of Technology*

Rich Roberts, *New England Biolabs*

Ramunas Stepanauskas, *Bigelow Laboratory for Ocean Sciences*

Matt Sullivan, *The Ohio State University*

Synthetic Biology Advisory Committee

Members

Richard Bailey, *Independent consultant*

Doug Cameron, *Firstgreen Partners*

Sunil Chandran, *Amyris, Inc.*

James Flatt, *Hampton Creek*

Megan Palmer, *University of California, Berkeley*

Elizabeth Sattely, *Stanford University*

Elizabeth Shank, *University of North Carolina, Chapel Hill*

David Weller, *USDA-ARS*

Appendix E
2016 Genomics of Energy and Environment Meeting

With a crowd at full capacity in attendance at the Walnut Creek Marriott, the 11th Annual DOE JGI Genomics of Energy and Environment Meeting took place March 22–24, 2016.

Keynote Speakers

Dianne Newman of Caltech opened the 11th Annual DOE JGI Genomics of Energy and Environment Meeting. One of her messages, repeated by several speakers over the course of the meeting, is that much remains unknown about the underlying mechanisms and pathways of the systems being studied. Newman spoke about the importance of secondary metabolites, using a model system reliant on *Shewanella oneidensis* communities.

Read about the meeting at http://bit.ly/JGIPrimerSpring16

Learn more about the meeting talks at http://usermeeting.jgi.doe.gov/past-meetings/2016-agenda/

A recap of the annual meeting as it happened is available at http://bit.ly/JGI2016storify

Videos of the talks are available on DOE JGI's YouTube channel at http://bit.ly/JGI2016videos

In her closing keynote, **Margaret McFall-Ngai** from the University of Hawaii focused on the relationship between the bobtail squid and the luminescent bacteria, one she's studied for 30 years. "It's my opinion model systems provide insights into mechanisms underlying symbiosis," she said, speaking of the power of simple systems for understanding more complex ones.

Watch her talk at **http://bit.ly/JGI2016McFallNgai**.

Other Featured Speakers *(in order of appearance):*

Chris Bowler, *École Normale Supérieure (France)*

Manpreet Dhami, *Stanford University*

Charles Chiu, *University of California, San Francisco*

Ben Cole, *DOE Joint Genome Institute*

Sarah Lebeis, *University of Tennessee*

Gene Myers, *Max Planck Institute of Molecular Cell Biology and Genetics (Germany)*

Rhona Stuart, *Lawrence Livermore National Laboratory*

José Dinneny, *Carnegie Institution for Science*

Jeanine Olsen, *Groningen Institute of Evolutionary Life Sciences*

Jeffrey Ross-Ibarra, *University of California, Davis*

Elizabeth Kellogg, *Donald Danforth Plant Science Center*

Christine Queitsch, *University of Washington*

Tom Juenger, *University of Texas, Austin*

Christopher Mason, *Weill Cornell Medical College*

Tim Donohue, *Great Lakes Bioenergy Research Center*

Kirsten Hofmockel, *Pacific Northwest National Laboratory*

Reshma Shetty, *Ginkgo Bioworks*

Robert Riley, *DOE Joint Genome Institute*

Michelle O'Malley, *University of California, Santa Barbara*

Mary Voytek, *NASA*

Orli Bahcall, *Nature Publishing*

Hamilton Smith, *JCVI*

Kelly Wrighton, *The Ohio State University*

Appendix F
2016 Publications

Ai C et al. Complete genome sequences of evolved arsenate-resistant *Metallosphaera sedula* strains. *Genome Announc*. 2015 Oct 1;3(5). pii: e01142-15. doi: 10.1128/genomeA.01142-15.

Aizenberg-Gershtein Y et al. High quality permanent draft genome sequence of *Phaseolibacter flectens* ATCC 12775(T), a plant pathogen of French bean pods. *Stand Genomic Sci.* 2016 Jan 13;11:4. doi: 10.1186/s40793-015-0127-5. eCollection 2016.

Alfaro M et al. Comparative and transcriptional analysis of the predicted secretome in the lignocellulose-degrading basidiomycete fungus *Pleurotus ostreatus*. *Environ Microbiol.* 2016 Apr 26. doi: 10.1111/1462-2920.13360.

Alivisatos AP et al. Microbiome. A unified initiative to harness Earth's microbiomes. *Science.* 2015 Oct 30;350(6260):507-8. doi: 10.1126/science.aac8480.

Anderson IJ et al. Complete genome sequence of the Antarctic *Halorubrum lacusprofundi* type strain ACAM 34. *Stand Genomic Sci.* 2016 Sep 10;11(1):70. doi: 10.1186/s40793-016-0194-2. eCollection 2016.

Arango Isaza RE et al. Combating a global threat to a clonal crop: banana black sigatoka pathogen *Pseudocercospora fijiensis* (synonym *Mycosphaerella fijiensis*) genomes reveal clues for disease control. *PLoS Genet.* 2016 Aug 11;12(8):e1005876. doi: 10.1371/journal.pgen.1005876.

Ardley J et al. High-quality permanent draft genome sequence of *Ensifer medicae* strain WSM244, a microsymbiont isolated from *Medicago polymorpha* growing in alkaline soil. *Stand Genomic Sci.* 2015 Dec 10;10:126. doi: 10.1186/s40793-015-0119-5. eCollection 2015.

Bashir M et al. Functional metagenomics of spacecraft assembly clean rooms: presence of virulence factors associated with human pathogens. *Front Microbiol.* 2016 Sep 9;7:1321. doi: 10.3389/fmicb.2016.01321. eCollection 2016.

Beall BF et al. Ice cover extent drives phytoplankton and bacterial community structure in a large north-temperate lake: implications for a warming climate. *Environ Microbiol.* 2016 Jun;18(6):1704-19. doi: 10.1111/1462-2920.12819.

Beam JP et al. Assembly and succession of iron oxide microbial mat communities in acidic geothermal springs. *Front Microbiol.* 2016 Feb 15;7:25. doi: 10.3389/fmicb.2016.00025. eCollection 2016.

Belmann P et al. Bioboxes: standardised containers for interchangeable bioinformatics software. *Gigascience.* 2015 Oct 15;4:47. doi: 10.1186/s13742-015-0087-0. eCollection 2015.

Bendall ML et al. Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations. *ISME J.* 2016 Jul;10(7):1589-601. doi: 10.1038/ismej.2015.241.

Berben T et al. Complete genome sequence of *Thioalkalivibrio paradoxus* type strain ARh 1(T), an obligately chemolithoautotrophic haloalkaliphilic sulfur-oxidizing bacterium isolated from a Kenyan soda lake. *Stand Genomic Sci.* 2015 Nov 19;10:105. doi: 10.1186/s40793-015-0097-7. eCollection 2015.

Berben T et al. Partial genome sequence of the haloalkaliphilic soda lake bacterium *Thioalkalivibrio thiocyanoxidans* ARh 2(T). *Stand Genomic Sci.* 2015 Oct 26;10:85. doi: 10.1186/s40793-015-0078-x. eCollection 2015.

Berben T et al. Partial genome sequence of *Thioalkalivibrio thiocyanodenitrificans* ARhD 1(T), a chemolithoautotrophic haloalkaliphilic sulfur-oxidizing bacterium capable of complete denitrification. *Stand Genomic Sci.* 2015 Oct 26;10:84. doi: 10.1186/s40793-015-0080-3. eCollection 2015.

Bewick AJ et al. On the origin and evolutionary consequences of gene body DNA methylation. *Proc Natl Acad Sci USA.* 2016 Aug 9;113(32):9111-6. doi: 10.1073/pnas.1604666113.

Billings AF et al. Genome sequence and description of the anaerobic lignin-degrading bacterium *Tolumonas lignolytica* sp. nov. *Stand Genomic Sci.* 2015 Nov 19;10:106. doi: 10.1186/s40793-015-0100-3. eCollection 2015.

Blow MJ et al. The epigenomic landscape of prokaryotes. *PLoS Genet.* 2016 Feb 12;12(2):e1005854. doi: 10.1371/journal.pgen.1005854. eCollection 2016 Feb.

Boden R et al. Permanent draft genome of *Thermithiobaclillus tepidarius* DSM 3134T, a moderately thermophilic, obligately chemolithoautotrophic member of the Acidithiobacillia. *Stand Genomic Sci.* 2016 Sep 26;11:74.

Bowers RM et al. Impact of library preparation protocols and template quantity on the metagenomic reconstruction of a mock microbial community. *BMC Genomics.* 2015 Oct 24;16(1):856. doi: 10.1186/s12864-015-2063-6.

Bredeson JV et al. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nat Biotechnol.* 2016 May;34(5):562-70. doi: 10.1038/nbt.3535. Epub 2016 Apr 18.

Bremges A et al. MeCorS: Metagenome-enabled error correction of single cell sequencing reads. *Bioinformatics.* 2016 Jul 15;32(14):2199-201. doi: 10.1093/bioinformatics/btw144.

Browne P et al. Genomic composition and dynamics among Methanomicrobiales predict adaptation to contrasting environments. *ISME J.* 2016 Aug 23. doi: 10.1038/ismej.2016.104.

Brumm PJ et al. Complete genome sequence of *Geobacillus thermoglucosidasius* C56-YS93, a novel biomass degrader isolated from obsidian hot spring in Yellowstone National Park. *Stand Genomic Sci.* 2015 Oct 5;10:73. doi: 10.1186/s40793-015-0031-z. eCollection 2015.

Brumm PJ et al. Complete genome sequences of *Geobacillus* sp. WCH70, a thermophilic strain isolated from wood compost. *Stand Genomic Sci.* 2016 Apr 27;11:33. doi: 10.1186/s40793-016-0153-y. eCollection 2016.

Buels R et al. JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biol.* 2016 Apr 12;17(1):66. doi: 10.1186/s13059-016-0924-1.

Cadillo-Quiroz H et al. Complete genome sequence of *Methanosphaerula palustris* E1-9CT, a hydrogenotrophic methanogen isolated from a minerotrophic fen peatland. *Genome Announc.* 2015 Nov 5;3(6). pii: e01280-15. doi: 10.1128/genomeA.01280-15.

Castanera R et al. Transposable elements versus the fungal genome: impact on whole-genome architecture and transcriptional profiles. *PLoS Genet.* 2016 Jun 13;12(6):e1006108. doi: 10.1371/journal.pgen.1006108. eCollection 2016 Jun.

Chen IM et al. Supporting community annotation and user collaboration in the integrated microbial genomes (IMG) system. *BMC Genomics.* 2016 Apr 26;17(1):307. doi: 10.1186/s12864-016-2629-y.

Choi DH et al. High-quality draft genome sequence of *Gracilimonas tropica* CL-CB462(T) (DSM 19535(T)), isolated from a *Synechococcus* culture. *Stand Genomic Sci.* 2015 Nov 11;10:98. doi: 10.1186/s40793-015-0088-8. eCollection 2015.

Coleman-Derr D et al. Plant compartment and biogeography affect microbiome composition in cultivated and native *Agave* species. *New Phytol.* 2016 Jan;209(2):798-811. doi: 10.1111/nph.13697.

Collier R et al. Use of *Agrobacterium rhizogenes* strain 18r12v and Paromomycin selection for transformation of *Brachypodium distachyon* and *Brachypodium sylvaticum*. *Front Plant Sci.* 2016 May 24;7:716. doi: 10.3389/fpls.2016.00716. eCollection 2016.

Corrochano LM et al. Expansion of signal transduction pathways in fungi by extensive genome duplication. *Curr Biol.* 2016 Jun 20;26(12):1577-84. doi: 10.1016/j.cub.2016.04.038.

Couturier M et al. Enhanced degradation of softwood versus hardwood by the white-rot fungus *Pycnoporus coccineus*. *Biotech Biofuels.* 2015 Dec 18;8:216. doi: 10.1186/s13068-015-0407-8. eCollection 2015.

Daly RA et al. Microbial metabolisms in a 2.5-km-deep ecosystem created by hydraulic fracturing in shales. *Nat Microbiol.* 2016 Sep 5;1:16146. doi: 10.1038/nmicrobiol.2016.146.

David AS et al. Draft genome sequence of *Microdochium bolleyi*, a dark septate fungal endophyte of beach grass. *Genome Announc.* 2016 Apr 28;4(2). pii: e00270-16. doi: 10.1128/genomeA.00270-16.

De Meyer SE et al. Symbiotic Burkholderia species show diverse arrangements of *nif/fix* and *nod* genes and lack typical high-affinity cytochrome cbb3 oxidase genes. *Mol Plant Microbe Interact.* 2016 Aug;29(8):609-619.

De Meyer SE et al. High-quality permanent draft genome sequence of the *Mimosa asperata*-nodulating *Cupriavidus* sp. strain AMP6. *Stand Genomic Sci.* 2015 Oct 16;10:80. doi: 10.1186/s40793-015-0074-1. eCollection 2015.

De Meyer SE et al. High-quality permanent draft genome sequence of the *Lebeckia ambigua*-nodulating *Burkholderia* sp. strain WSM4176. *Stand Genomic Sci.* 2015 Oct 16;10:79. doi: 10.1186/s40793-015-0072-3. eCollection 2015.

Denef VJ et al. Chloroflexi CL500-11 populations that predominate deep-lake hypolimnion bacterioplankton rely on nitrogen-rich dissolved organic matter metabolism and C1 compound oxidation. *Appl Environ Microbiol.* 2016 March;82(5):1423-32. doi: 10.1128/AEM.03014–15.

Denef VJ et al. Seasonal succession leads to habitat-dependent differentiation in ribosomal RNA:DNA ratios among freshwater lake bacteria. *Front Microbiol.* 2016 Apr 29;7:606. doi: 10.3389/fmicb.2016.00606. eCollection 2016.

Deng K et al. Development of a high throughput platform for screening glycoside hydrolases based on Oxime-NIMS. *Front Bioeng Biotechnol.* 2015 Oct 13;3:153. doi: 10.3389/fbioe.2015.00153. eCollection 2015.

Duncan DS et al. Detection of short-term cropping system-induced changes to soil bacterial communities differs among four molecular characterization methods. *Soil Biol Biochem.* 2016 May. 96, 160–168. doi:10.1016/j.soilbio.2016.02.002

Eiler A et al. Tuning fresh: radiation through rewiring of central metabolism in streamlined bacteria. *ISME J.* 2016 Aug;10(8):1902-14. doi: 10.1038/ismej.2015.260.

Eloe-Fadrosh EA et al. Global metagenomic survey reveals a new bacterial candidate phylum in geothermal springs. *Nat Commun.* 2016 Jan 27;7:10476. doi: 10.1038/ncomms10476.

Eloe-Fadrosh EA et al. Metagenomics uncovers gaps in amplicon-based detection of microbial diversity. *Nat Microbiol.* 1 Feb 2016;1:15032. doi:10.1038/nmicrobiol.2015.32

Eshraghi L et al. High-quality permanent draft genome sequence of *Bradyrhizobium* sp. strain WSM1743 — an effective microsymbiont of an *Indigofera* sp. growing in Australia. *Stand Genomic Sci.* 2015 Oct 26;10:87. doi: 10.1186/s40793-015-0073-2. eCollection 2015.

Evans J et al. Diversity and population structure of northern switchgrass as revealed through exome capture sequencing. *Plant J.* 2015 Oct 1. doi: 10.1111/tpj.13041.

Flynn JD et al. Draft genome sequences of gammaproteobacterial methanotrophs isolated from marine ecosystems. *Genome Announc.* 2016 Jan 21;4(1). pii: e01629-15. doi: 10.1128/genomeA.01629-15.

Fonseca-García C et al. The cacti microbiome: interplay between habitat-filtering and host-specificity. *Front Microbiol.* 2016 Feb 12;7:150. doi: 10.3389/fmicb.2016.00150. eCollection 2016.

Frei CS et al. Analysis of amino acid substitutions in AraC variants that respond to triacetic acid lactone. *Protein Sci.* 2016 Apr;25(4):804–14. doi: 10.1002/pro.2873.

Fujimoto M et al. Spatiotemporal distribution of bacterioplankton functional groups along a freshwater estuary to pelagic gradient in Lake Michigan. *J Great Lakes Res.* 2016 Sep 10. http://dx.doi.org/10.1016/j.jglr.2016.07.029

Gach PC et al. A Droplet Microfluidic Platform for Automating Genetic Engineering. *ACS Synth Biol.* 2016 May 20;5(5):426–33. doi: 10.1021/acssynbio.6b00011. Epub 2016 Feb 22.

Gaudana SB et al. Bioinformatic analysis of the distribution of inorganic carbon transporters and prospective targets for bioengineering to increase Ci uptake by cyanobacteria. *Photosynth Res.* 2015 Oct;126(1):99–109. doi: 10.1007/s11120-014-0059-8.

Gazis R et al. The genome of *Xylona heveae* provides a window into fungal endophytism. *Fungal Biol.* 2016 Jan;120(1):26–42. doi: 10.1016/j.funbio.2015.10.002.

Gehlot HS et al. High-quality permanent draft genome sequence of Ensifer sp. PC2, isolated from a nitrogen-fixing root nodule of the legume tree (Khejri) native to the Thar Desert of India. *Stand Genomic Sci.* 2016 Jun 23;11:43. doi: 10.1186/s40793-016-0157-7. eCollection 2016.

Ghodhbane-Gtari F et al. Permanent draft genome sequence of *Nocardia* sp. BMG111209, an actinobacterium isolated from nodules of *Casuarina glauca*. *Genome Announc.* 2016 Aug 4;4(4). pii: e00770-16. doi: 10.1128/genomeA.00770-16.

Ghodhbane-Gtari F et al. Permanent improved high-quality draft genome sequence of Nocardia casuarinae strain BMG51109, an endophyte of *Actinorhizal Root Nodules of Casuarina glauca. Genome Announc.* 2016 Aug 4;4(4). pii: e00799-16. doi: 10.1128/genomeA.00799-16.

Goodwin SB et al. The mitochondrial genome of the ethanol-metabolizing, wine cellar mold *Zasmidium cellare* is the smallest for a filamentous ascomycete. *Fungal Biol.* 2016 Aug;120(8):961–74. doi: 10.1016/j.funbio.2016.05.003.

Goordial J et al. Cold adaptive traits revealed by comparative genomic analysis of the eurypsychrophile *Rhodococcus* sp. JG3 isolated from high elevation McMurdo Dry Valley permafrost, Antarctica. *FEMS Microbiol Ecol.* 2016 Mar 17;4(2). pii: e00069-16. doi: 10.1128/genomeA.00069-16.

Goordial J et al. Improved high-quality draft genome sequence of the eurypsychrophile *Rhodotorula* sp. JG1b, isolated from permafrost in the hyperarid upper-elevation McMurdo Dry Valleys, Antarctica. *Genome Announc.* 2016 Mar 17;4(2). pii: e00069-16. doi: 10.1128/genomeA.00069-16.

Grabowski PP et al. Genome-wide associations with flowering time in switchgrass using exome-capture sequencing data. *New Phytol*. 2016 Jul 22. doi: 10.1111/nph.14101.

Guibert LM et al. Diverse bacterial groups contribute to the alkane degradation potential of chronically polluted subantarctic coastal sediments. *Microb Ecol.* 2016 Jan;71(1):100–12. doi: 10.1007/s00248-015-0698-0. Epub 2015 Nov 7.

Gunsalus RP et al. Complete genome sequence of *Methanospirillum hungatei* type strain JF1. *Stand Genomic Sci.* 2016 Jan 6;11:2. doi: 10.1186/s40793-015-0124-8. eCollection 2016.

Gutierrez T et al. Genome sequence of *Marinobacter* sp. strain MCTG268 isolated from the cosmopolitan marine diatom *Skeletonema costatum*. *Genome Announc.* 2016 Sep 8;4(5). pii: e00937-16. doi: 10.1128/genomeA.00937-16.

Gutierrez T et al. Genome sequence of *Arenibacter algicola* strain TG409, a hydrocarbon-degrading bacterium associated with marine eukaryotic phytoplankton. *Genome Announc.* 2016 Aug 4;4(4). pii: e00765-16. doi: 10.1128/genomeA.00765-16.

Hatzenpichler R et al. Visualizing *in situ* translational activity for identifying and sorting slow-growing archaeal-bacterial consortia. *Proc Natl Acad Sci USA.* 2016 Jul 12;113(28):E4069-78. doi: 10.1073/pnas.1603757113.

Hausmann B et al. Consortia of low-abundance bacteria drive sulfate reduction-dependent degradation of fermentation products in peat soil microcosms. *ISME J.* 2016 Oct;10(10):2365-75. doi: 10.1038/ismej.2016.42. Epub 2016 Mar 25.

Hedlund BP et al. High-quality draft genome sequence of *Kallotenue papyrolyticum* JKG1T reveals broad heterotrophic capacity focused on carbohydrate and amino acid metabolism. *Genome Announc.* 2015 Dec 3;3(6). pii: e01410-15. doi: 10.1128/genomeA.01410-15.

Hemme CL et al. Comparative metagenomics reveals impact of contaminants on groundwater microbiomes. *Front Microbiol.* 2015 Oct 31;6:1205. doi: 10.3389/fmicb.2015.01205. eCollection 2015.

Hillson NJ et al. Improving synthetic biology communication: recommended practices for visual depiction and digital submission of genetic designs. *ACS Synth Biol.* 2016 Jun 17;5(6):449-51. doi: 10.1021/acssynbio.6b00146.

Hirschman L et al. Crowdsourcing and curation: perspectives from biology and natural language processing. *Database* (Oxford). 2016 Aug 7;2016. pii: baw115. doi: 10.1093/database/baw115. Print 2016.

Holmfeldt K et al. Large-scale maps of variable infection efficiencies in aquatic Bacteroidetes phage-host model systems. *Environ Microbiol.* 2016 May 28. doi: 10.1111/1462-2920.13392. [Epub ahead of print]

Hudson CM et al. Lignin-modifying processes in the rhizosphere of arid land grasses. *Environ Microbiol.* 2015 Dec;17(12):4965-78. doi: 10.1111/1462-2920.13020.

Huntemann M et al. The standard operating procedure of the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v.4). *Stand Genomic Sci.* 2015 Oct 26;10:86. doi: 10.1186/s40793-015-0077-y. eCollection 2015.

Huntemann M et al. The standard operating procedure of the DOE-JGI Metagenome Annotation Pipeline (MAP v.4). *Stand Genomic Sci.* 2016 Feb 24;11:17. doi: 10.1186/s40793-016-0138-x. eCollection 2016.

Hutchinson MI et al. Genetics of mating in members of the Chaetomiaceae as revealed by experimental and genomic characterization of reproduction in *Myceliophthora heterothallica*. *Fungal Genet Biol.* 2016 Jan;86:9–19. doi: 10.1016/j.fgb.2015.11.007.

Ikeda-Ohtsubo W et al. *'Candidatus Adiutrix intracellularis'*, an endosymbiont of termite gut flagellates, is the first representative of a deep-branching clade of Deltaproteobacteria and a putative homoacetogen. *Environ Microbiol.* 2016 Sep;18(8):2548–64. doi: 10.1111/1462-2920.13234. Epub 2016 Apr 21.

Inskeep WP et al. Geomicrobiology of sublacustrine thermal vents in Yellowstone Lake: geochemical controls on microbial community structure and function. *Front Microbiol.* 2015 Oct 26;6:1044. doi: 10.3389/fmicb.2015.01044. eCollection 2015.

Javidpour P et al. Investigation of proposed ladderane biosynthetic genes from anammox bacteria by heterologous expression in *E. coli*. *PLoS One.* 2016 Mar 14;11(3):e0151087. doi: 10.1371/journal.pone.0151087. eCollection 2016.

Jiang Y et al. An expanded evaluation of protein function prediction methods shows an improvement in accuracy. *Genome Biol.* 2016 Sep 7;17(1):184. doi: 10.1186/s13059-016-1037-6.

Kan J et al. Geochemistry and mixing drive the spatial distribution of free-living archaea and bacteria in Yellowstone Lake. *Front Microbiol.* 2016 Feb 29;7:210. doi: 10.3389/fmicb.2016.00210. eCollection 2016.

Kang DD et al. Reconstructing single genomes from complex microbial communities. *it-Information Technology.* 2016 Jun 28;58(3),133–139. doi: 10.1515/itit-2016-0011.

Kanukollu S et al. Genome sequence of *Shimia* str. SK013, a representative of the Roseobacter group isolated from marine sediment. *Stand Genomic Sci.* 2016 Mar 12;11:25. doi: 10.1186/s40793-016-0143-0. eCollection 2016.

Kappler U et al. Complete genome sequence of the haloalkaliphilic, obligately chemolithoautotrophic thiosulfate and sulfide-oxidizing γ-proteobacterium *Thioalkalimicrobium cyclicum* type strain ALM 1 (DSM 14477(T)). *Stand Genomic Sci.* 2016 Jun 3;11:38. doi: 10.1186/s40793-016-0162-x. eCollection 2016.

Kato S et al. Comparative genomic insights into ecophysiology of neutrophilic, microaerophilic iron oxidizing bacteria. *Front Microbiol.* 2015 Nov 13;6:1265. doi: 10.3389/fmicb.2015.01265. eCollection 2015

Kazakov AE et al. σ54-dependent regulome in *Desulfovibrio vulgaris* Hildenborough. *BMC Genomics.* 2015 Nov 10;16:919. doi: 10.1186/s12864-015-2176-y.

Kelly WJ et al. Genomic analysis of three Bifidobacterium species isolated from the calf gastrointestinal tract. *Sci Rep.* 2016 Jul 29;6:30768. doi: 10.1038/srep30768.

Kemp DW et al. Spatial homogeneity of bacterial communities associated with the surface mucus layer of the reef-building coral *Acropora palmate. PLOS One.* 2015 Dec 14;10(12):e0143790. doi: 10.1371/journal.pone.0143790.

Kim JW et al. Genome analysis of planctomycetes inhabiting blades of the red alga *Porphyra umbilicalis. PLoS One.* 2016 Mar 25;11(3):e0151883. doi: 10.1371/journal.pone.0151883. eCollection 2016.

Kyrpides NC et al. Microbiome data science: understanding our microbial planet. *Trends Microbiol.* 2016 Jun;24(6):425-7. doi: 10.1016/j.tim.2016.02.011.

Li G et al. Genome-wide sequencing of 41 rice (*Oryza sativa* L.) mutated lines reveals diverse mutations induced by fast-neutron irradiation. *Mol Plant.* 2016 Jul 6;9(7):1078–81. doi: 10.1016/j.molp.2016.03.009.

Linshiz G et al. End-to-end automated microfluidic platform for synthetic biology: from design to functional analysis. *J Biol Eng.* 2016 Feb 2;10:3. doi: 10.1186/s13036-016-0024-5. eCollection 2016.

Liu L et al. A xyloglucan backbone 6-O-acetyltransferase from *Brachypodium distachyon* modulates xyloglucan xylosylation. *Mol Plant.* 2016 Apr 4;9(4):615–7. doi: 10.1016/j.molp.2015.11.004.

Louca S et al. Integrating biogeochemistry with multiomic sequence information in a model oxygen minimum zone. *Proc Natl Acad Sci USA.* 2016 Sep 21. pii: 201602897.

Louie TS et al. High-quality draft genome sequence of *Sedimenticola selenatireducens* strain AK4OH1T, a gammaproteobacterium isolated from estuarine sediment. *Stand Genomic Sci.* 2016 Sep 8;11:66.

Lovell JT et al. Drought responsive gene expression regulatory divergence between upland and lowland ecotypes of a perennial C4 grass. *Genome Res.* 2016 Apr;26(4):510-8. doi: 10.1101/gr.198135.115.

Mamidi S et al. Sequence-based introgression mapping identifies candidate white mold tolerance genes in common bean. *Plant Genome.* 2016 May 6;9(2). doi:10.3835/plantgenome2015.09.0092

Markowitz VM et al. Ten years of maintaining and expanding a microbial genome and metagenome analysis system. *Trends Microbiol.* 2015 Nov;23(11): 730–41. doi: 10.1016/j.tim.2015.07.012.

McCarthy S et al. Expanding the limits of thermoacidophily in the archaeon *Sulfolobus solfataricus* by adaptive evolution. *Appl Environ Microbiol.* 2015 Nov 20;82(3):857-67. doi: 10.1128/AEM.03225-15.

Mefferd CC et al. High-quality draft genomes from *Thermus caliditerrae* YIM 77777 and *T. tengchongensis* YIM 77401, isolates from Tengchong, China. *Genome Announc.* 2016 Apr 28;4(2). pii: e00312-16. doi: 10.1128/genomeA.00312-16.

Melton ED et al. Complete genome sequence of *Desulfurivibrio alkaliphilus* strain AHT2(T), a haloalkaliphilic sulfidogen from Egyptian hypersaline alkaline lakes. *Stand Genomic Sci.* 2016 Sep 8;11(1):67. doi: 10.1186/s40793-016-0184-4. eCollection 2016.

Miettinen O et al. Draft genome sequence of the white-rot fungus *Obba rivulosa* 3A-2. *Genome Announc.* 2016 Sep 15;4(5). pii: e00976-16. doi: 10.1128/genomeA.00976-16.

Miyauchi S et al. Visual comparative omics of fungi for plant biomass deconstruction. *Front. Microbiol.* 2016 Aug 24;7:1335. doi: 10.3389/fmicb.2016.01335

Mosier AC et al. Fungi contribute critical but spatially varying roles in nitrogen and carbon cycling in acid mine drainage. *Front Microbiol.* 2016 Mar 3;7:238. doi: 10.3389/fmicb.2016.00238. eCollection 2016.

Mukai T et al. Facile recoding of Selenocysteine in nature. *Angew Chem Int Ed Engl.* 2016 Apr 18;55(17):5337–41. doi: 10.1002/anie.201511657.

Muraguchi H et al. Strand-specific RNA-seq analyses of fruiting body development in *Coprinopsis cinerea. PLoS One.* 2015 Oct 28;10(10):e0141586. doi: 10.1371/journal.pone.0141586. eCollection 2015.

Nagy LG et al. Comparative genomics of early-diverging mushroom-forming fungi provides insights into the origins of lignocellulose decay capabilities. *Mol Biol Evol.* 2016 Apr;33(4):959–70. doi: 10.1093/molbev/msv337.

Nguyen HD et al. Ochratoxin A production by *Penicillium thymicola. Fungal Biol.* 2016 Aug;120(8):1041–9. doi: 10.1016/j.funbio.2016.04.002.

Nicholson WL et al. An improved high-quality draft genome sequence of *Carnobacterium inhibens* subsp. *inhibens* strain K1(T). *Stand Genomic Sci.* 2016 Sep 8;11(1):65. doi: 10.1186/s40793-016-0193-3. eCollection 2016.

Niederhuth CE et al. Widespread natural variation of DNA methylation within angiosperms. *Genome Biol.* 2016 Sep 27;17(1):194.

Nouioui I et al. Draft genome sequence of *Frankia* Strain G2, a nitrogen-fixing actinobacterium isolated from *Casuarina equisetifolia* and able to nodulate actinorhizal plants of the order rhamnales. *Genome Announc.* 2016 May 26;4(3). pii: e00437-16. doi: 10.1128/genomeA.00437-16.

Olsen JL et al. The genome of the seagrass *Zostera marina* reveals angiosperm adaptation to the sea. *Nature.* 2016 Feb 18;530(7590):331–-5. doi: 10.1038/nature16548.

Oulas A et al. Metagenomic investigation of the geologically unique Hellenic Volcanic Arc reveals a distinctive ecosystem with unexpected physiology. *Environ Microbiol.* 2015 Oct 21. doi: 10.1111/1462-2920.13095.

Oyserman BO et al. Metatranscriptomic insights on gene expression and regulatory controls in *Candidatus Accumulibacter phosphatis. ISME J.* 2016 Apr;10(4):810–22. doi: 10.1038/ismej.2015.155.

Paez-Espino D et al. Uncovering Earth's virome. *Nature.* 2016 Aug 25;536(7617):425–-30.

Panschin I et al. Comparing polysaccharide decomposition between the type strains *Gramella echinicola* KMM 6050(T) (DSM 19838(T)) and *Gramella portivictoriae* UST040801-001(T) (DSM 23547(T)), and emended description of *Gramella echinicola* Nedashkovskaya et al. 2005 emend. Shahina et al. 2014 and *Gramella portivictoriae* Lau et al. 2005. *Stand Genomic Sci.* 2016 Jun 3;11:37. doi: 10.1186/s40793-016-0163-9. eCollection 2016.

Peña A et al. High quality draft genome sequences of *Pseudomonas fulva* DSM 17717(T), *Pseudomonas parafulva* DSM 17004(T) and *Pseudomonas cremoricolorata* DSM 17059(T) type strains. *Stand Genomic Sci.* 2016 Sep 1;11(1):55. doi: 10.1186/s40793-016-0178-2. eCollection 2016.

Pernice MC et al. Large variability of bathypelagic microbial eukaryotic communities across the world's oceans. *ISME J.* 2016 Apr;10(4):945–58. doi: 10.1038/ismej.2015.170.

Peter M et al. Ectomycorrhizal ecology is imprinted in the genome of the dominant symbiotic fungus *Cenococcum geophilum*. *Nat Commun*. 2016 Sep 7;7:12662. doi: 10.1038/ncomms12662.

Petrik DL et al. BdCESA7, BdCESA8, and BdPMT utility promoter constructs for targeted expression to secondary cell-wall-forming cells of grasses. *Front Plant Sci.* 2016 Feb 4;7:55. doi: 10.3389/fpls.2016.00055. eCollection 2016.

Piché-Choquette S et al. H2-saturation of high affinity H2-oxidizing bacteria alters the ecological niche of soil microorganisms unevenly among taxonomic groups. *PeerJ.* 2016 Mar 10;4:e1782. doi: 10.7717/peerj.1782. eCollection 2016.

Plaza DF et al. Identification of a novel nematotoxic protein by challenging the model mushroom *Coprinopsis cinerea* with a fungivorous nematode. *G3 (Bethesda).* 2015 Nov 19;6(1):87–98. doi: 10.1534/g3.115.023069.

Putnam NH et al. Chromosome-scale shotgun assembly using an *in vitro* method for long-range linkage. *Genome Res*. 2016 Mar;26(3):342–50. doi: 10.1101/gr.193474.115.

Quinn JY et al. SBOL visual: a graphical language for genetic designs. *PLoS Biol.* 2015 Dec 3;13(12):e1002310. doi: 10.1371/journal.pbio.1002310. eCollection 2015.

Raissig MT et al. Grasses use an alternatively wired bHLH transcription factor network to establish stomatal identity. *Proc Natl Acad Sci USA.* 2016 Jul 19;113(29):8326–31. doi: 10.1073/pnas.1606728113.

Rice MC et al. Complete genome of *Nitrosospira briensis* C-128, an ammonia-oxidizing bacterium from agricultural soil. *Stand Genomic Sci.* 2016 Jul 28;11:46. doi: 10.1186/s40793-016-0168-4. eCollection 2016.

Riley R et al. Comparative genomics of biotechnologically important yeasts. *Proc Natl Acad Sci USA.* 2016 Aug 30;113(35):9882-7. doi: 10.1073/pnas.1603941113.

Rivers AR et al. Experimental identification of small non-coding RNAs in the model marine bacterium *Ruegeria pomeroyi DSS-3. Front Microbiol.* 2016 Mar 29;7:380. doi: 10.3389/fmicb.2016.00380. eCollection 2016.

Roehner N et al. Sharing structure and function in biological design with SBOL 2.0. *ACS Synth Biol.* 2016 Jun 17;5(6):498–506. doi: 10.1021/acssynbio.5b00215.

Rojas-Rojas FU et al. Draft genome sequence of heavy metal-resistant *Cupriavidus alkaliphilus* ASC-732T, isolated from agave rhizosphere in the northeast of Mexico. *Genome Announc.* 2016 Sep 22;4(5). pii: e01013-16. doi: 10.1128/genomeA.01013-16.

Roux S et al. Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature.* 2016 Sep 21. doi: 10.1038/nature19366.

Salazar G et al. Particle-association lifestyle is a phylogenetically conserved trait in bathypelagic prokaryotes. *Mol Ecol.* 2015 Oct 13. doi: 10.1111/mec.13419.

Salman V et al. Insights into the single cell draft genome of "Candidatus Achromatium palustre". *Stand Genomic Sci.* 2016 Mar 23;11:28. doi: 10.1186/s40793-016-0146-x. eCollection 2016.

Schneebeli K et al. *Brachypodium distachyon* genotypes vary in resistance to *Rhizoctonia solani* anastomosis group 8. *Funct Plant Biol.* 2016 Feb 1;43 (2),189–198. doi: 10.1071/FP15244

Seshadri R et al. Discovery of novel plant interaction determinants from the genomes of 163 root nodule bacteria. *Sci Rep.* 2015 Nov 20;5:16825. doi: 10.1038/srep16825.

Shaw A.J. et al. Chapter five – the Sphagnum genome project: a new model for ecological and evolutionary genomics. *Adv Bot Res.* 2016 Feb 22. 78,167–187. http://dx.doi.org/10.1016/bs.abr.2016.01.003

Shi L et al. Performance evaluation and tuning of BioPig for genomic analysis. Proceeding DISCS '15 *Proceedings of the 2015 International Workshop on Data-Intensive Scalable Computing Systems.* 2015 Nov 11; 9. doi:10.1145/2831244.2831252

Singer E et al. Next generation sequencing data of a defined microbial mock community. *Sci Data.* 2016 Sep 27;3:160081. doi: 10.1038/sdata.2016.81.

Singer E et al. Similar microbial communities found on two distant seafloor basalts. *Front Microbiol.* 2015 Dec 16;6:1409. doi: 10.3389/fmicb.2015.01409. eCollection 2015.

Singer E et al. High-resolution phylogenetic microbial community profiling. *ISME J.* 2016 Aug;10(8):2020–32. doi: 10.1038/ismej.2015.249. Epub 2016 Feb 9

Small GE et al. Large differences in potential denitrification and sediment microbial communities across the Laurentian great lakes. *Biogeochemistry.* 2016 Apr 23. 128: 353. doi:10.1007/s10533-016-0212-x

Solomon KV et al. Early-branching gut fungi possess a large, comprehensive array of biomass-degrading enzymes. *Science.* 2016 Mar 11;351(6278):1192–5. doi: 10.1126/science.aad1431.

Solomon KV et al. Robust and effective methodologies for cryopreservation and DNA extraction from anaerobic gut fungi. *Anaerobe.* 2016 Apr;38:39–46. doi: 10.1016/j.anaerobe.2015.11.008.

Song Q et al. Construction of high resolution genetic linkage maps to improve the soybean genome sequence assembly Glyma1.01. *BMC Genomics.* 2016 Jan 6;17(1):33. doi: 10.1186/s12864-015-2344-0.

Spatafora JW et al. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia.* 2016 Sep;108(5):1028–1046.

Stevenson SR et al. Genetic analysis of Physcomitrella patens identifies ABSCISIC ACID NON-RESPONSIVE (ANR), a regulator of ABA responses unique to basal land plants and required for desiccation tolerance. *Plant Cell.* 2016 Jun;28(6):1310–27. doi: 10.1105/tpc.16.00091.

Strassert JF et al. Genome analysis of *'Candidatus Ancillula trichonymphae',* first representative of a deep-branching clade of Bifidobacteriales, strengthens evidence for convergent evolution in flagellate endosymbionts. *Environ Microbiol Rep.* 2016 Aug 12. doi: 10.1111/1758-2229.12451.

Strepis N et al. Description of *Trichococcus ilyis* sp. nov. by combined physiological and in silico genome hybridization analyses. *Int J Syst Evol Microbiol.* 2016 Jul 11. doi: 10.1099/ijsem.0.001294.

Susanti D et al. Permanent draft genome sequence of *Desulfurococcus mobilis* type strain DSM 2161, a thermoacidophilic sulfur-reducing crenarchaeon isolated from acidic hot springs of Hveravellir, Iceland. *Stand Genomic Sci.* 2016 Jan 13;11:3. doi: 10.1186/s40793-015-0128-4. eCollection 2016.

Tashkandy N et al. High-quality draft genome sequence of *Flavobacterium suncheonense* GH29-5(T) (DSM 17707(T)) isolated from greenhouse soil in South Korea, and emended description of *Flavobacterium suncheonense* GH29-5(T). *Stand Genomic Sci.* 2016 Jun 16;11:42. doi: 10.1186/s40793-016-0159-5. eCollection 2016.

Tiwari R et al. Genome sequence of Bradyrhizobium sp. WSM1253; a microsymbiont of *Ornithopus compressus* from the Greek Island of Sifnos. *Stand Genomic Sci.* 2015 Nov 30;10:113. doi: 10.1186/s40793-015-0115-9. eCollection 2015.

Tsementzi D et al. SAR11 bacteria linked to ocean anoxia and nitrogen loss. *Nature.* 2016 Aug 11;536(7615):179–83.

Utturkar SM et al. Application of long sequence reads to improve genomes for *Clostridium thermocellum* AD2, *Clostridium thermocellum* LQRI, and *Pelosinus fermentans* R7. *Genome Announc.* 2016 Sep 29;4(5). pii: e01043-16. doi: 10.1128/genomeA.01043-16.

van Baren MJ et al. Evidence-based green algal genomics reveals marine diversity and ancestral characteristics of land plants. *BMC Genomics.* 2016 Mar 31;17(1):267. doi: 10.1186/s12864-016-2585-6.

Vavourakis CD et al. Metagenomic insights into the uncultured diversity and physiology of microbes in four hypersaline Soda Lake brines. *Front Microbiol.* 2016 Feb 25;7:211. doi: 10.3389/fmicb.2016.00211. eCollection 2016.

Vigneron A et al. Complementary microorganisms in highly corrosive biofilms from an offshore oil production facility. *Appl Environ Microbiol.* 2016 Feb 19. pii: AEM.03842-15.

Wagner MR et al. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat Commun.* 2016 Jul 12;7:12151. doi: 10.1038/ncomms12151.

Walker AK et al. Full genome of *Phialocephala scopiformis* DAOMC 229536, a fungal endophyte of spruce producing the potent anti-insectan compound rugulosin. *Genome Announc.* 2016 Feb 25;4(1). pii: e01768-15. doi: 10.1128/genomeA.01768-15.

Waltman PH et al. Identifying aspects of the post-transcriptional program governing the proteome of the green alga *Micromonas pusilla*. *PLoS One.* 2016 Jul 19;11(7):e0155839. doi: 10.1371/journal.pone.0155839. eCollection 2016.

Weinmaier T et al. A viability-linked metagenomic analysis of cleanroom environments: eukarya, prokaryotes, and viruses. *Microbiome.* 2015 Dec 8;3:62. doi: 10.1186/s40168-015-0129-y.

Winkel M et al. Single-cell sequencing of thiomargarita reveals genomic flexibility for adaptation to dynamic redox conditions. *Front Microbiol.* 2016 Jun 21;7:964. doi: 10.3389/fmicb.2016.00964. eCollection 2016.

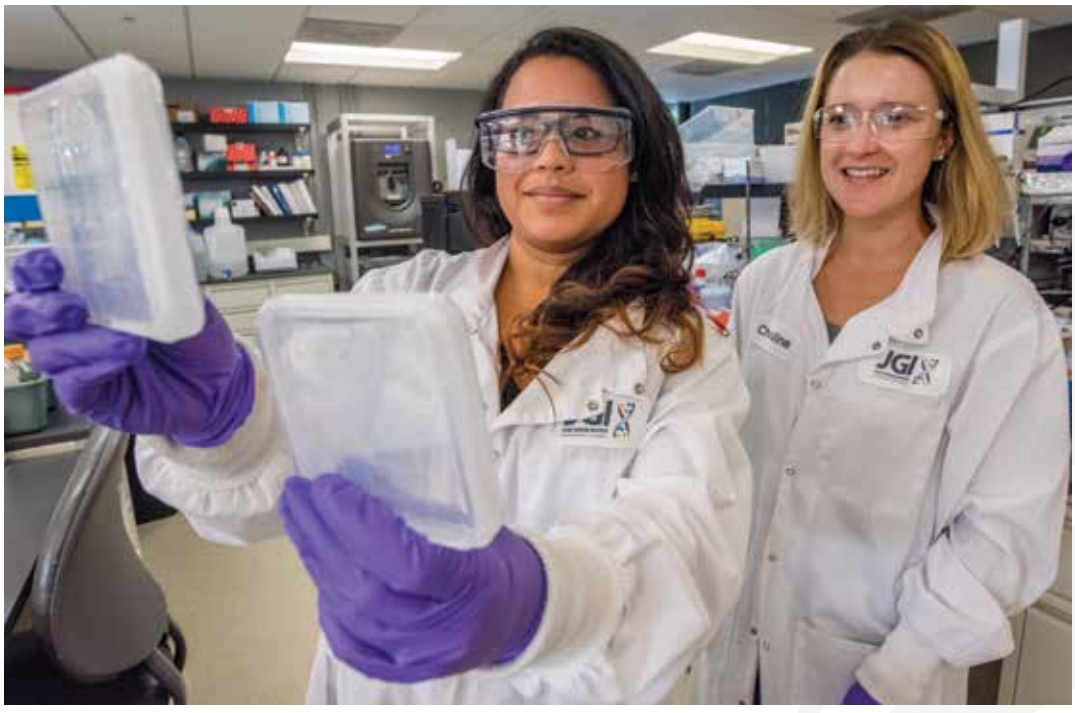
Yang S et al. Functionally conserved enhancers with divergent sequences in distant vertebrates. *BMC Genomics.* 2015 Oct 30;16:882. doi: 10.1186/s12864-015-2070-7.

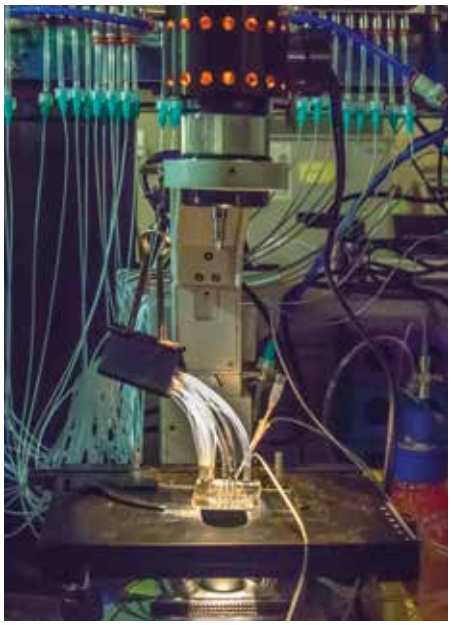
Yang Y et al. Genome-wide analysis of lectin receptor-like kinases in *Populus*. *BMC Genomics.* 2016 Sep 1;17:699. doi: 10.1186/s12864-016-3026-2.

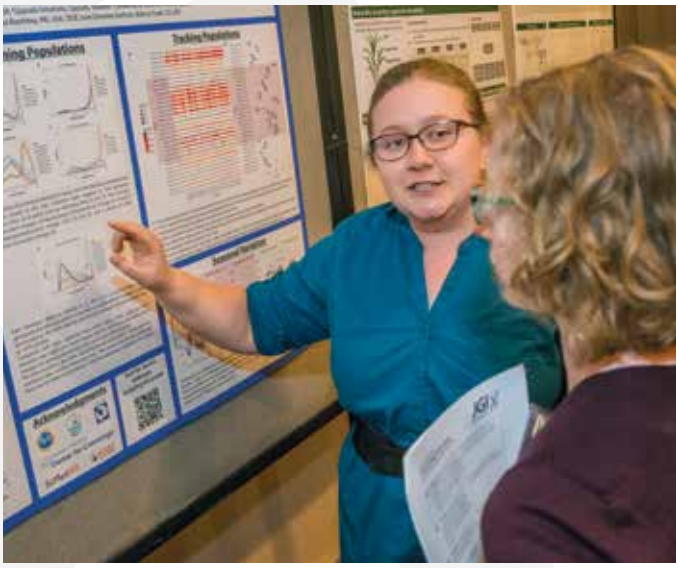
Zargar K et al. *In vitro* characterization of phenylacetate decarboxylase, a novel enzyme catalyzing toluene biosynthesis in an anaerobic microbial community. *Sci Rep.* 2016 Aug 10;6:31362. doi: 10.1038/srep31362.

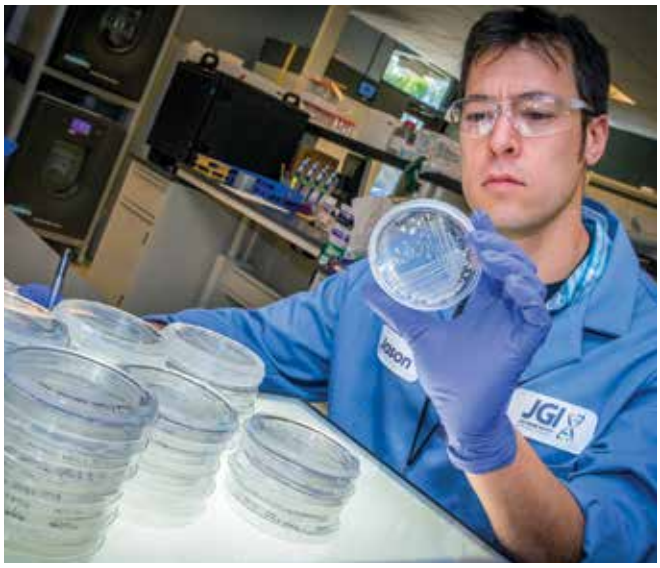
Zeiner CA et al. Comparative analysis of secretome profiles of Manganese(II)-Oxidizing Ascomycete fungi. *PLoS One.* 2016 Jul 19;11(7):e0157844. doi: 10.1371/journal.pone.0157844. eCollection 2016.

Zhou EM et al. High-quality draft genome sequence of the *Thermus amyloliquefaciens* type strain YIM 77409(T) with an incomplete denitrification pathway. *Stand Genomic Sci.* 2016 Feb 27;11:20. doi: 10.1186/s40793-016-0140-3. eCollection 2016.









Comments?

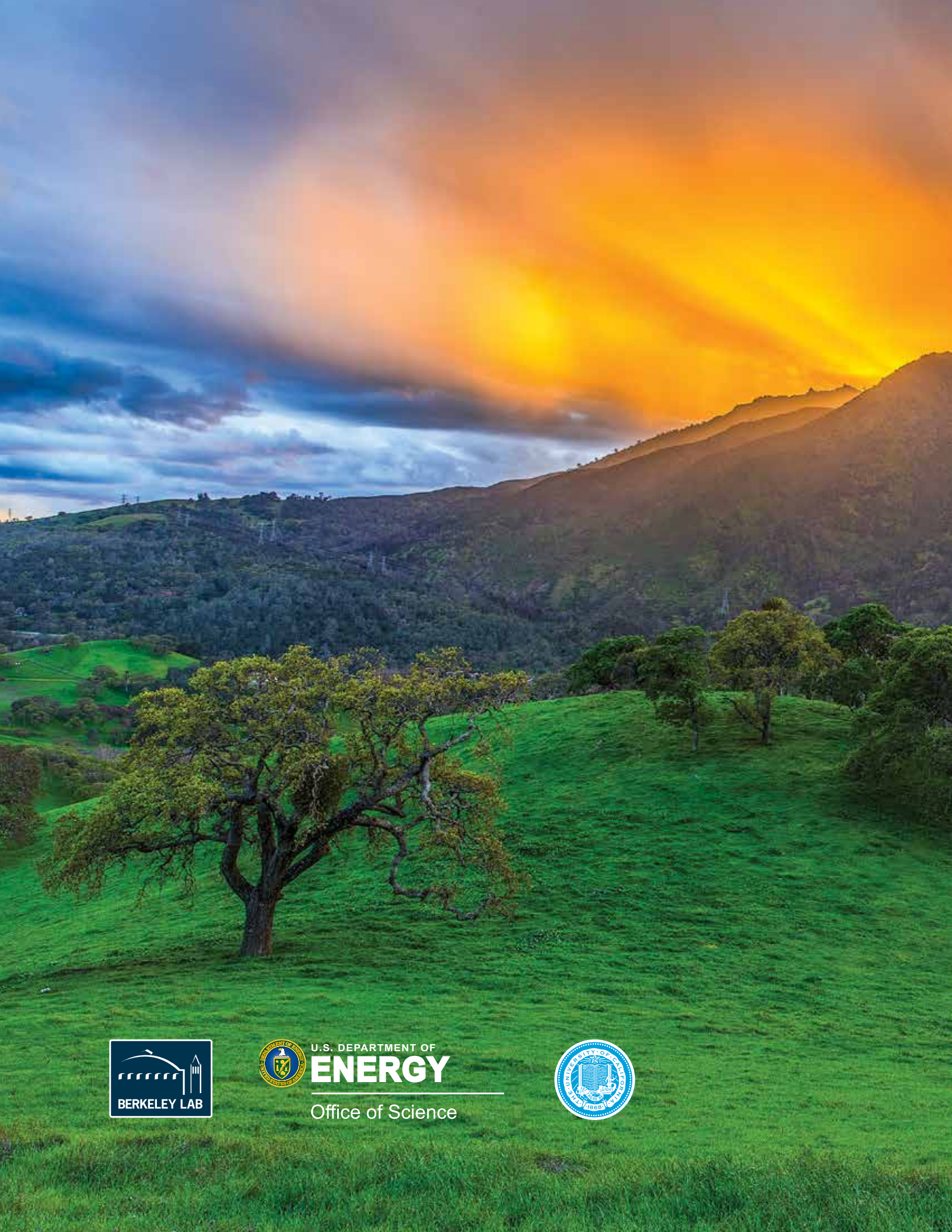
Massie S. Ballon, DOE JGI Public Affairs, Editor
mlballon@lbl.gov

Follow Us

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.

All photography, unless otherwise noted, by Roy Kaltschmidt, Berkeley Lab.

17-JGI-3589


BERKELEY LAB
U.S. DEPARTMENT OF
ENERGY
Office of Science