THE PRIMER

JGI
JOINT GENOME INSTITUTE
DEPARTMENT OF ENERGY

spring 2019
volume 16
issue 2

Study senior author Diane Dickel in her lab. (Thor Swift/Berkeley Lab)

in this issue

Technique for Studying Gene Expression Takes Root in Plants

Excerpted from the Berkeley Lab release by Aliyah Kovner

Lawrence Berkeley National Laboratory (Berkeley Lab) researchers, including scientists at the U.S. Department of Energy (DOE) Joint Genome Institute (JGI), a DOE Office of Science User Facility, have successfully used an open-source RNA analysis platform on plant cells for the first time. The technology, called Drop-seq, allows scientists to see what genes are being expressed and how this relates to the specific functions of different cell types. Prior to this study, published in *Cell Reports*, Drop-seq had only been used in animal cells.

"For many of the genes in plants, we have little to no understanding of what they actually do," explained senior author Diane Dickel, who studies mammalian genomics in the Lab's Biosciences Area. "But by knowing exactly what cell type or developmental stage a specific gene is expressed in, we can start getting a toehold into its function. In our study, we showed that Drop-seq can help us do this."

Dickel has been using Drop-seq on animal cells for several years. An immediate fan of the platform's ease of use and efficacy, she soon began speaking to her colleagues working on plants about trying to use it on plant cells.

However, some were skeptical that such a project would work as easily. First off, to run plant cells through a single-cell RNA-seq analysis, they must be protoplasted — meaning they must be stripped of their cell walls using unique cocktails of enzymes. Secondly, some plant biologists have expressed concern that cells are altered too significantly by protoplasting to provide insight into normal functioning. And finally, some plant cells are simply too big to be put through existing single-cell RNA-seq platforms.

continued on page 7

Dimensions of Omics Diversity

Notes from the 2019 JGI Meeting

Natural Selection in Rice

Kicking off the 14th Annual JGI Genomics of Energy & Environment Meeting in San Francisco, Calif., Michael D. Purugganan, Dean of Science at New York University, focused on two projects to better understand how selection acts on the rice genome.

One project involves developing a rice fitness consequence map to help crop breeders looking for regions with mutations and phenotypic variations. It's based on the INSIGHT (Inference for Natural Selection from Interspersed Genomically CoHerent elemenTs) method developed for human genome studies. The greenINSIGHT method Puruggunan's lab developed over four years calculates the fitness impact of mutations that may be neutral, nearly neutral or advantageous. The method allowed the team to identify the rare gene variants and map them back to the genome. The Rice FitCons Map is expected to be available summer 2019.

continued on page 2

continued from page 1

Purugganan also cited the use of rice gene expression selection to measure the contribution of the transcription levels to fitness. This is a project he waited more than a decade to execute. Partnering with the International Rice Research Institute in the Philippines, Purugganan's team used samples from plants that were either under well-watered or intermittent drought conditions. Among their early findings from the resulting dataset of more than 15,000 genes: they've found a "drought-escape" gene that causes the plant to flower early.

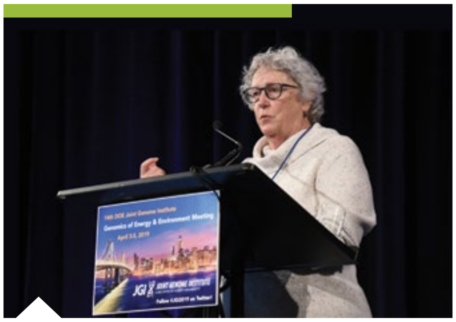
Michael Purugganan

Fully Understanding Soil Microbiome Interactions

In her keynote address, Mary Firestone of the University of California, Berkeley, emphasized the need to understand the full picture of soil microbiome interactions. The complexity was underscored by the plethora of research collaborations, and lab members past and present who have helped describe the interactions within and across domains thus far to better understand the composition and function of soil microbiomes.

Firestone's long-term goal is to learn about the interactions among plants, bacteria, fungi, viruses and mesofauna in the soil. She talked about work using dominant grasses in northern and southern California to study the effects of drought, using stable isotope probing to find and map pathways of root carbon consumers into the soil, and using the data for network analyses and food web modelling. The team has gone from isotopically labelling grasses in greenhouse chambers to field sites at the Hopland Research and Extension Center in northern California.

"We can't understand any interactions without recognizing interactions among soil inhabitants," she said. Most interactions studied have been in isolation, such as plants and bacteria, or bacteria and viruses, but these studies don't give researchers the full picture. "We now have the tools that enable exploration of the complex interactions that occur in soil," she added, listing an arsenal that includes omics, isotopic tools, mass spectroscopy, statistics, and modelling.



Mary Firestone

Preventing a Second Reptilian Age

In his closing keynote, Arturo Casadevall of the Johns Hopkins School of Public Health led with the human existential questions, "Why are we here?" and "Why do we eat so much?" to explain why dinosaurs no longer roam the earth.

He briefly recapped how people have come to understand the origins of illnesses before they were attributed to microbial infections, and how these understandings have been linked to terms used to describe illness, starting with the use of "virulent" to mean "full of poison." Twenty years ago, Casadevall began thinking of disease as an outcome of an interaction between a pathogen and its host. With that definition, he began trying to find a mathematical way to track the pathogenic potential of microbes, including fungi.

The fungal kingdom, Casadevall reminded the audience, has more than six million species and includes several pathogens currently doing significant damage to plants, insects and whole ecosystems. Humans and mammals, he noted, are comparatively resistant to fungal pathogens; credit goes to adaptive immunity and high body temperatures.

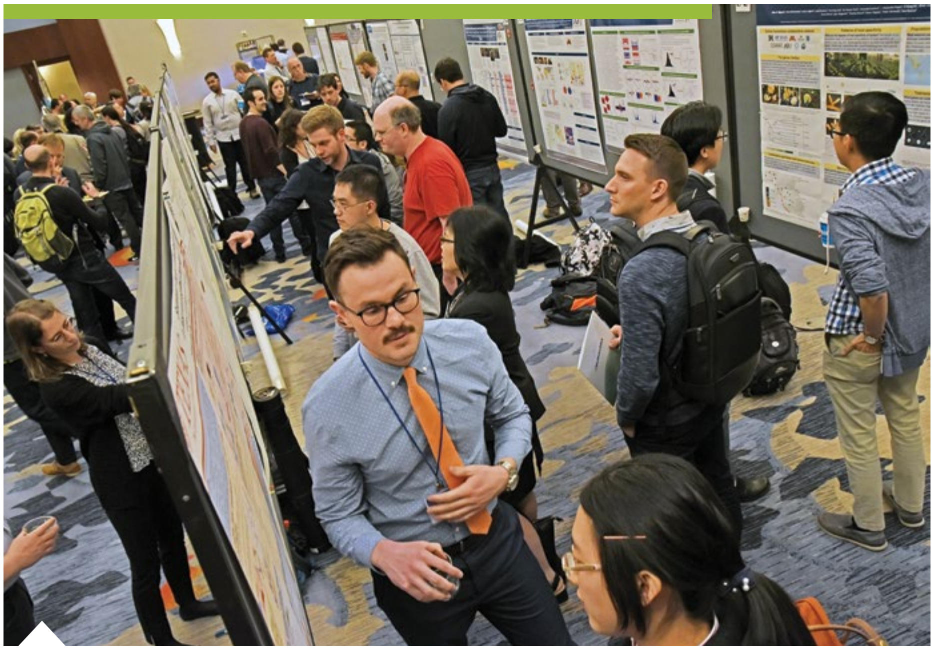
All of this led Casadevall back to the question of the great lizards: "Don't ask what killed the dinosaurs, ask 'What kept the reptiles from having a Second Reptilian Age?'" The answer is the same as the question of why humans are here and others he raised at the beginning of his talk: fungi.

Arturo Casadevall

Diversity in All Dimensions

Geri Richmond, noted chemist at the University of Oregon and a strong advocate for diversity in the scientific workforce, asked attendees to consider serving as a mentor to the next-generation workforce. "Good innovation and high impact research thrives on diversity in all dimensions," she reminded them. Richmond should know; alongside her research career, she is founding director of the grassroots organization COACh, which has helped advance the careers of thousands of women scientists and engineers around the world. Among her accolades are the National Medal of Science and the 2018 Priestley Medal, the American Chemical Society's highest honor.

"Social diversity causes discomfort and we have to work harder," she acknowledged while highlighting the economic benefits of diversity, "but that brings out the creativity." She asked attendees to consider mentoring, serving in the short-term either as a coach who can help refine specific skills, or as a sponsor who can help advance someone's career, or else as a long-term mentor who supports and encourages someone's growth and development.



Annual JGI Meeting Poster Session

"All You Need is LOV"

Some fungi can sense to environmental factors such as light and gravity, and respond by changing their growth patterns accordingly. Luis Larrondo of the Pontificia Universidad Catolica de Chile combines synthetic biology and what he calls "optogenetics," how cellular processes can be activities or repressed by light and genetics, to learn more the circadian rhythms in fungi.

Larrondo demonstrated how fungi streaked on a plate can distinguish light intensities. The image is projected onto a plate on which the fungi are growing. The plate acts as a photographic film, and after the fungi have genetically processed the light, they emit bioluminescence in a form that replicate the image. To do that, Larrondo told the audience, the fungi need LOV — the light, oxygen & voltage protein domain in *Neurospora crassa*.

Aesthetics aside, he added, the images produced through fungal bioluminescence raise several scientific questions, including transcriptional dynamics, cell-cell communications, and fungal adaptation to light.

Continuing with the photography analogy, Larrondo's team built a pinhole camera to allow the "blind" *Neurospora* to sense light intensities. They did this by developing a promoter that responds to light, and then combining it with *Neurospora's* circadian clock that controls daily rhythms in gene expression. The result: an image that slowly appears and then fades over the course of a day, and then reappears following the fungal cycle.

Worldbuilding with Automation

Emma Frow from the University of Arizona opted to focus on the sociological implications of automation in biofoundries and multiomics facilities, including the JGI. She pointed out that in the challenge of scaling up production and reducing human error, researchers may not recognize that they're shifting their thinking from how to conduct the experiment to how they can optimize and efficiently use the equipment. Building a facility and the underlying infrastructure, she reminded the audience, is about giving the team's vision and aspirations form, and making design choices that reflect this goal.

Frow asked attendees to consider the relationships between machines and humans, adding that while the human component isn't often taken into account when developing workflows, somehow the information and skills need to be transferred. "Infrastructure for bioengineering requires the complex integration of mechanical, digital, biological and human systems," she concluded. "Automation… is a widely accepted choice, but it's not inevitable."

continued on page 5

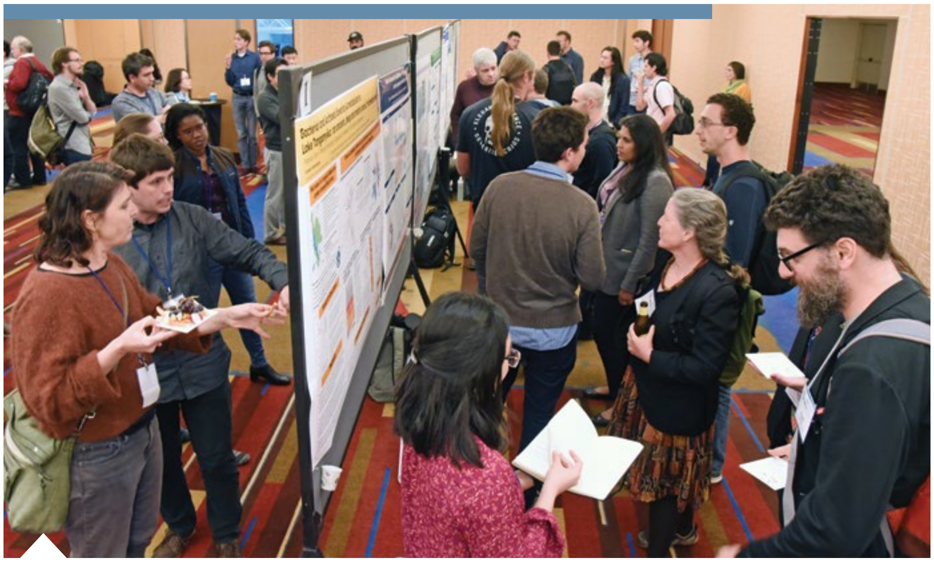
Highlights from the New Lineages of Life Symposium

A Paradigm Shift on Eukarya

Thijs Ettema of Wageningen University opened the JGI's NeLLi 2019 Symposium: From New Lineages of Life to New Functions, with the proposal that there are two domains of life — and eukarya is not one of them. Countering Carl Woese's 40-year old argument that the Tree of Life has three domains of life, he suggested that eukaryotes evolved from a specific clade of archaea rather than being a separate branch with shared common ancestry.

With this theory in mind, Ettema's lab has been working on identifying what he called the "founding archaeal lineages" of eukaryotic cells by looking for the direct descendants of the host cell, and the mitochondrial ancestor. In 2015, Ettema and fellow NeLLi Symposium keynote Christa Schleper published a paper on the eukaryotic nature of Asgard archaea (named for one of the nine worlds in Norse mythology), which represent the closest relatives of eukaryotes. The Ettema lab has focused on Asgard archaea since then, and through collaborations, they've greatly increased the number of Asgard archaeal lineages. In the current pantheon of Asgard archaeal diversity Frigga, Idunn, Frey and Hela are among those joining Odin, Thor, Loki and Heimdall.

Thijs Ettema



New Lineages of Life Symposium Poster Session

Ettema closed by recapping the argument that eukaryotes are chimeric in nature, and that one of the Asgard lineages may be a sister relation of the ancestral eukaryote cell.

Not Hot, But Lukewarm

Archaea are thought to have originated in hot springs, and generally considered extremophiles. In her keynote, Christa Schleper of the University of Vienna focused on archaea that have adapted to thrive in more moderate environments. Thaumarchaeota are found in a wide range of environments including oceans, soil, human skin, some wastewater treatment plants, and even ultraclean rooms in NASA. "They may already be on Enceladus," she noted, referring to an ice-clad moon of Saturn.

Searching for clues on how and why Thaumarchaeota thrives in a wide range of environments, the Schleper lab focused on a species called *Nitroscocaldus cavascurensis* that they isolated, cultured and sequenced.

Using this archaeon, Schleper and her colleagues are tracking down the last common ancestor of all Thaumarchaeota (LACAOA), to learn more about how early life evolved. And they are finding that this is a series of successive gains rather than a single event or leap. When analyzing the archaeal genome, they found that *N. cavascurensis* turned out to be an ammonia-oxidizing archaea (AOA) that branches early. Additionally, the genome contains mobile elements acquired through lateral gene transfer, suggesting the archaeon acquired genes that allowed it to thrive outside extreme environments.

Christa Schleper

John McCutcheon

A Patchwork of Pathways

Echoing Ettema's talk, John McCutcheon of the University of Montana focused on the role of endosymbiosis in the evolution of eukaryotes. He outlined three important endosymbiotic events in the history of life: the rise of all photosynthetic eukaryotes; the evolution of eukaryotes; and, the rise of all sap feeding insects. (Pro tip from McCutcheon: "You cannot feed an animal on plant sap alone — it's the equivalent of trying to raise a child on 7-Up.")

McCutcheon shared two anecdotes as part of his lab's efforts to understand how endosymbiont genomes can be reduced from thousands of genes down to hundreds. One of McCutcheon's examples of an endosymbiotic system involves a bacterium called *Moranella endobia* that resides within a bacterium called *Tremblaya princeps* which in turn resides within a mealybug insect cell. Unsurprisingly, he noted, the metabolic pathways in the mealybug are "patchwork patterns of genes in different places," which makes it difficult to predict the sources and locations of the functional genes.

Miscommunications about Evolution

The Twisted Tree of Life Awards began when Jonathan Eisen of the University of California, Davis became frustrated by the inaccuracies he saw in the news about evolution and evolutionary findings. One of the problems, and thus an Award category, is that people assume that Darwin is synonymous with evolution. As in, he said, "If Darwin got something wrong in the 1850s, then all evolution is wrong… We should get back to debating the tree of life."

The eponymous category of the awards arose from frequent news reports that point out life is categorized into five kingdoms of life. (For the record, Carl Woese redrew the tree of life in 1977 and distinguished three domains of life.) A third category considers misuse of terminology. "Primitive carries too many connotations," he said. "Organisms that exist today are not primitive, but may contain ancestral features." Similarly, he questioned if describing an organism as "higher" meant it was more evolved than a "lower organism."

Eisen conceded that communicating science accurately can be a challenge, but consistently making exceptions is not helpful. "It's important to be aware of the risks of doing it wrong," he said, recommending that researchers pay attention to how the science is presented in press releases as they are likely to be the foundation for media coverage and thus, public interest.

Jonathan Eisen

continued from page 3

2019 JGI Meeting

Genomics at the Exascale

Dan Jacobson spoke about being part of the team at Oak Ridge National Laboratory that broke the exascale barrier in 2018. They used the Summit supercomputer to carry out nearly 2 quadrillion (a billion billion) calculations per second while conducting genomic data analyses with datasets containing millions of genomes to uncover hidden networks of genes.

Having a supercomputer that can process population-scale genomic datasets meant Jacobson and his colleagues naturally transitioned to human health applications, specifically, to uncover the genes involved in opioid addiction. The work done as a joint initiative between the Department of Energy and the US Veterans Administration aims to improve healthcare outcomes for Veterans.

The ORNL team that included Jacobson and JGI's Chief Informatics Officer Kjiersten Fagnan won a 2018 ACM Gordon Bell at the International Conference for High Performance Computing, Networking, Storage and Analysis (SC18) Conference. Addressing the JGI meeting attendees, Jacobson counted the win as the first ever for genomics and systems biology, outlining the work that applied high-performance computing to population-scale genomic datasets.

Select videos from the Meeting are available on JGI's YouTube channel at http://bit.ly/JGI2019videos.

Dan Jacobson and Kjiersten Fagnan

Left to Right: Rex Malmstrom, head of Micro-Scale Applications; Emiley Eloe-Fadrosh, head of the Metagenome Program; and, research scientist Simon Roux.

Perspectives of Innovation

In a special issue of *mSystems* published May 14, 2019, JGI researchers were among the authors who offered perspectives on what the next five years of innovation could look like.

In one article, Micro-Scale Applications head Rex Malmstrom and Metagenome Program head Emiley Eloe-Fadrosh outlined more targeted approaches to reconstruct individual microbes in an environmental sample beyond the tried-and-true method of traditional shotgun metagenomics. One approach uses fluorescence-activated cell sorting (FACS) to sort cells into several small pools before sequencing to recover rare microbes. The technique was applied to a recent Nature Communications paper, uncovering giant virus sequences in soil for the first time. (Learn more at jgi.doe.gov/hidden-giant-virus-forest-soils.)

A second approach known as DNA stable isotope probing (DNA SIP) relies on incubating DNA with isotopically-labeled compounds to link active microbes with specific metabolic processes. A third approach steps away from reconstructing genomes directly from the environment and focuses on sequence reads to link mobile elements such as plasmids to microbial host cells.

In a separate article, research scientist Simon Roux, a member of Eloe-Fadrosh's Environmental Genomics group, made a pitch for readers to get involved in the developing field of virus ecogenomics. While microbes regulate the planet's biogeochemical cycles, viruses infecting microbes shape global ecological and environmental processes, especially by modifying host cell metabolism and altering cell fitness. As the vast majority of viruses remain uncultivated, Roux outlined what's been achieved thus far based on computational predictions, and future directions integrating computational and experimental approaches with the aim of eventually studying virus-host dynamics in nature.

A comprehensive map of viruses including host information, Roux said, would be the foundation upon which researchers can develop high-throughput and high-resolution studies of interactions between host cells and viruses in the environment.

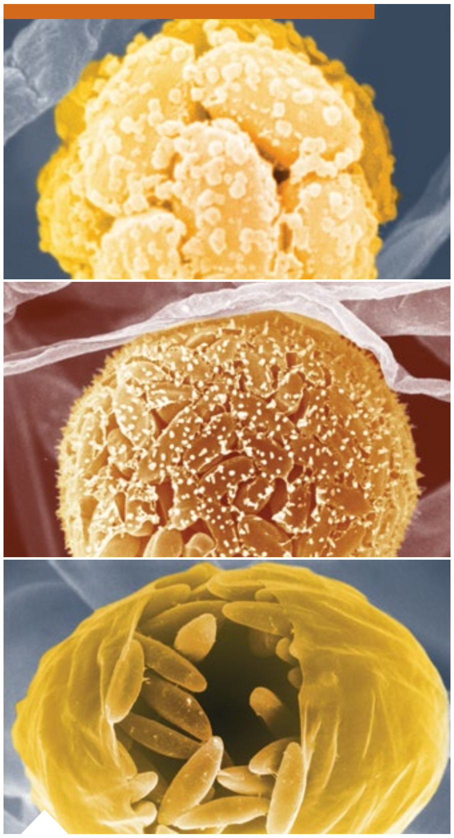
Roux and Eloe-Fadrosh are also among the JGI researchers who led a community effort to develop guidelines and best practices for defining virus data quality. In a *Nature Biotechnology* paper, JGI partnered with a number of virus experts; as well as representatives from the Genomic Standards Consortium (GSC), an open-membership working body that engages the research community in the standards development process; and the International Committee on Taxonomy of Viruses, the premier authority on the official taxonomy of viruses which is currently re-evaluating virus classification based on sequence-based information. (Learn more at jgi.doe.gov/defining-quality-virus-datasets.)

Unwrapping Complex Mechanisms of Genome-Wide Regulation

The JGI's 1000 Fungal Genomes initiative is uncovering the true diversity of the fungal kingdom. The abundance of fungal genomes underscores the need for a more comprehensive exploration of genomic DNA modifications (epigenome). The best-known example of a modified DNA base is a methylation of the base cytosine at the 5th position on its carbon ring, or 5mC. In many organisms, 5mC is an important piece of how they regulate their genomes, but is not well understood in fungi.

In *Nature Ecology & Evolution*, a team led by Robert Schmitz at the University of Georgia and including JGI researchers conducted the largest analysis to date of 5mC, an important piece of the epigenome. Across the fungal tree of life, they tracked the evolution of 5mC pathways across 528 species.

The researchers found that unlike most flowering plants and animals, fungi lack gene body methylation. The analysis revealed that 5mC in some fungal species is abundant



Top: *Spinellus fusiger.* Center: *Parasitella parasitica.* Bottom: *Helicostylum cordense.* (ZyGoLife Research Consortium on Flickr, CC BY-SA 2.0)

in repetitive DNA and transposons in methylated cytosine clusters, rather than in genes, which suggests that 5mC may play a role in genome defense. Additionally, the team found a correlation between genome-wide 5mC abundance and the abundance of specific 5mC methyltransferase (MTase) variants.

A Model System for Perennial Grasses

In *Nature Communications*, a team led by Tom Juenger of the University of Texas (UT) at Austin and including JGI researchers reported the culmination of nearly a decade of work to develop genomic resources for drought tolerance in perennial grasses. The team is applying the resources developed for *Panicum hallii* towards stress tolerance improvement in its more complex relative, the candidate bioenergy crop, switchgrass (*P. virgatum*).

Through the Community Science Program, JGI sequenced and assembled near-complete genomes of *P. hallii* var. *hallii* (99.2% complete) and *P. hallii* var. *filipes* (94.8% complete) and resequenced a host of natural collections from across the species range. The HAL2 and FIL2 genomes are available on the JGI's plant portal Phytozome.

With these high-quality reference genomes for *P. hallii*, researchers can identify and characterize the regulatory elements that influence adaptation and tolerance to stressors such as drought. This information can be applied toward improving crop yields in other grasses.

Panicum hallii var. *filipes* (left) and *Panicum hallii* var. *hallii* (right) grown under controlled greenhouse conditions in Austin, Texas. (Amalia Díaz)

Technique for Studying Gene Expression Takes Root in Plants

continued from page 1

Undeterred by these challenges, Dickel and her JGI colleagues teamed up with researchers from UC Davis who had perfected a protoplasting technique for root tissue from *Arabidopsis thaliana* (mouse-ear cress), a species of small flowering weed that serves as a plant model organism.

After preparing samples of more than 12,000 *Arabidopsis* root cells, the group was thrilled when the Drop-seq process went smoother than expected.

"When I first spoke to Diane about trying Drop-seq in plants I recognized the huge potential, but I thought it would be difficult to separate plant cells rapidly enough to get useful data," said John Vogel, JGI's head of plant functional genomics. "I was shocked to see how well it worked and how much they were able to learn from their initial experiment. This technique is going to be a game changer for plant biologists because it allows us to explore gene expression without grinding up whole plant organs, and the results aren't muddled by signals from the few most common cell types."

The authors anticipate that the platform, and other similar RNA-seq technologies, will eventually become routine in plant investigations. "We also showed that you can use these technologies to understand how plants respond to different environmental conditions at a cellular level — something many plant biologists at Berkeley Lab are interested in because being able to grow crops under poor environmental conditions, such as drought, is essential for our continued production of food and biofuel resources," Dickel said.

Full story at newscenter.lbl.gov/2019/05/16/new-technique-plant-rna.

JGI JOINT GENOME INSTITUTE
A DOE OFFICE OF SCIENCE USER FACILITY

The 15th Annual DOE Joint Genome Institute

Genomics of Energy & Environment Meeting

Oakland Marriott City Center

Watch this site!
usermeeting.jgi.doe.gov

SAVE THE DATES!

MARCH 23 – 27, 2020

The Integrative Genomics Building (IGB), the new home of JGI and KBase at Berkeley Lab

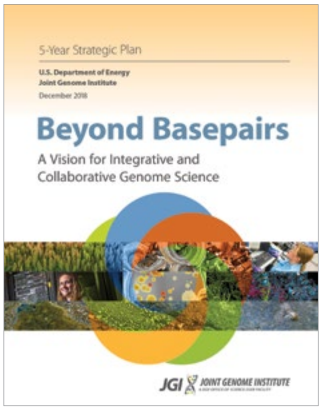
BERKELEY LAB · BioSciences · U.S. DEPARTMENT OF ENERGY Office of Science

19-JG-5821

Congratulations to the JGI 2019 Annual Meeting Poster Winners!

Recognizing scientific posters that presented particularly compelling or novel research in the field of microbial genomics, *Microbial Genomics*, an open access and open data journal published by the Microbiology Society, awarded a prize at the New Lineages of Life (NeLLi) Symposium to Haley Sapers of CalTech (right), seen with NeLLi organizer Tanja Woyke (left). Another prize was awarded to Matthew Deaner (second from right in the group shot) of the University of Texas-Austin during the JGI Genomics of Energy & Environment Meeting. Both received a $200 cash price and a certificate.

At the JGI Genomics of Energy & Environment Meeting, the Best Student Poster Award went to Lazarus Takawira (right) of the University of Pretoria, while the Best Overall Poster Award went to Erin Nuccio of Lawrence Livermore National Laboratory.



5-Year Strategic Plan
U.S. Department of Energy Joint Genome Institute
December 2018

Beyond Basepairs
A Vision for Integrative and Collaborative Genome Science

JGI JOINT GENOME INSTITUTE

JGI Strategic Plan

JGI has released its newest 5-Year Strategic Plan: Beyond Basepairs — A Vision for Integrative and Collaborative Genome Science. New areas of emphasis include JGI's emerging leadership role in a Microbiome Data Science effort (pages 37-39) and the launching of a Secondary Metabolites research group (pages 47-49). Download a copy at 2019_JGI-Strategic-Plan.pdf.

Register Now!

Find out which upcoming conferences will have a JGI presence:

ASM Microbe 2019
June 20–24, 2019
in San Francisco, CA:
www.asm.org/Events/ASM-Microbe/Home
Attend the IMG Workshop and make sure to stop by Booth 5132!

2019 Synthetic Biology: Engineering, Evolution & Design (SEED)
June 23–27, 2019
in New York City, NY
synbioconference.org/2019

Society for Industrial Microbiology and Biotechnology (SIMB) Annual Meeting
July 21–24, 2019
in Washington DC
www.simbhq.org/annual/

Plant Biology 2019
August 3–7, 2019
in San Jose, CA:
plantbiology.aspb.org
Make sure to stop by the JGI/KBase Booth 607!

Microbial Genomics & Metagenomics (MGM) Workshop
November 4–8, 2019
in Berkeley, CA:
Register now:
mgm.jgi.doe.gov

Contact The Primer

David Gilbert, Managing Editor
DEGilbert@lbl.gov
Massie Santos Ballon, Editor

19-JGI-60030