

Steve Quake with
DOE JGI Director Eddy Rubin

Stephen Quake: Seizing the Moment

SABIN RUSSELL

Stanford bioengineering professor Stephen Quake, renowned as an inventive scientist and creator of new businesses, has much to do with the current success of single-cell genomics.

Quake launched the U.S. Department of Energy Joint Genome Institute's (DOE JGI) 9th Annual Genomics of Energy & Environment Meeting on March 18, 2014 with a keynote address on single-cell genomics before a full house in Walnut Creek. "It is a field that is having its moment, right now," he said.

Quake has been in force in single-molecule biophysics since he was a postdoc in the 1990s at Stanford, where his adviser was Steven Chu. Quake pioneered automated single-molecule gene sequencing, large-scale integrated microfluidic circuits, and co-founded half a dozen companies, including leading microfluidic chipmaker Fluidigm. He is now mentor to a new generation of innovators.

"Steve specializes in taking laboratory insights and converting them into very practical things," said DOE JGI Director Eddy Rubin. "His work is very rapidly applied."

In January, single-cell sequencing was named the 2013 Method of the Year by the journal *Nature Methods*. "Once considered a technical challenge for a few specialized labs, single-cell transcriptome and genome sequencing is becoming robust and broadly accessible," the editors declared. "Last year marked a turning point in the widespread adoption of these methods."

Quake told the DOE JGI audience that a confluence of three technologies is responsible for the surge in interest. Advances in polymerase chain reaction (PCR) and multiple displacement amplification (MDA) enabled amplification of whole genomes. High-throughput DNA sequencers dramatically lowered costs. Automated single-cell manipulation technologies, *continued on page 2*

in this issue

Genomic Gluttony	4
Great Prairie	5
DOE JGI Highlights	6
More Meeting Notes	8

Surviving the Really, Really Big One

Like most who study the fascinatingly morbid topic of mass extinctions, Annalee Newitz has her favorite. Hers would be the Permian-Triassic extinction event, which wiped out 95 percent of life of Earth some 250 million years ago.

"There was so much death going on," as she describes it, "the only things propagating were slimes... Food webs collapsed. There was lots of death in the oceans, lots on land. In the end, it was...Slime World."

Offbeat, funny, iconoclastic, and thought-provoking, Newitz is the editor of io9, a highly acclaimed San Francisco-based blog that serves up a digital buffet of popular science, science fiction, and wildly imaginative visions of the future. An unabashed fan of science, she was delighted to share her thoughts on March 20, 2014 as the closing keynote speaker for the 9th Annual Genomics of Energy & Environment Meeting.

"JGI has been a huge inspiration for me and my work," said Newitz, who has a PhD in English and American Studies from the University of California, Berkeley, and a tattoo of the double helix on her left forearm. "This is not just because it's an awesome place to work and you have all these cool machines," she added. "You are doing a terrific job bridging *continued on page 3*

Stephen Quake

continued from page 1

such as fluorescence-activated cell sorting and microfluidic chips, have made meticulous molecular isolations practical. “Put all three of these things together, and you can do a lot,” he said.

In a wide-ranging review of recent work in the field of single-cell genomics, Quake noted that microfluidic chips were dramatically improving their capacity. Originally able to carry out 400 isolated reactions, the new 96 x 96 chips can perform 9,216. Increasing the number of individual reactions that can be observed helps to overcome the “technical noise” that can confound results obtained with fewer reactions.

Interest in the application of genomics for energy and environmental research drew more than 460 registrants to the DOE JGI annual meeting. Single-cell genomics is playing a central role in the analysis of microbial communities in soils or in extreme environments such as geyser pools. “Different species and different cell types are mixed all

together,” Quake explained following his talk. Single-cell methods help you to dissect that complexity.”

Single-cell genomics was enthusiastically embraced by numerous presenters at the annual meeting’s poster sessions. Andrzej Joachimiak, a structural biologist at Argonne National Laboratory, studies archaea inhabiting marine sediments. “Ninety percent of the biosphere exists at temperatures below 10°C,” he said. “These organisms represent a significant gene pool, and we should be sequencing millions of them.” Gaining understanding of these microbes and the biochemical processes involving them may have important implications for climate science, as a major portion of the world’s carbon is thought to be locked in sediments beneath the ocean floor.

Senthil Murugapiran, a postdoc at University of Nevada, Las Vegas, is using single-cell measurements to study *Aigarchaeota*, a vast phylum of archaeal microbes found in hot springs. Given that Eukaryota may have evolved from these primitive

extremophiles, single-cell sequencing will help to shed light on these “deep evolutionary relationships,” Murugapiran said.

Graduate student researcher Timothy Alba, who works with Murugapiran at UNLV’s School of Life Sciences, said single-cell genomics is a powerful tool for exploring newly discovered regions of biology. He studies *Aigarchaeota* from the hot pools of Great Boiling Springs, in northern Nevada, where the water temperature hovers around 87°C. “We don’t understand the role of an organism that is the most dominant life form at high temperatures,” he said. “A phylum is a huge branch of the tree of life, and these microbes are likely to have a significant impact on the planet’s carbon and nitrogen cycles.”

Watch Quake’s keynote on the DOE JGI’s YouTube channel at http://bit.ly/JGIUM9_Quake. 📺

Sabin Russell is a Berkeley Lab science writer.

The Big One

continued from page 1

the gap between basic science and the stuff that the public can understand has applications to their lives.”

Unfortunately, the news from the scientific frontiers has been alarming. Carbon dioxide levels in the atmosphere are soaring, and the planet is threatened with irreversible climate change. There have been five known mass extinctions in the past 450 million years, and as author Elizabeth Kolbert has written, many scientists see evidence that a sixth is already under way.

Newitz has tackled the topic as well, and her good news — for the DOE JGI delegates and humanity at large — is that we *Homo sapiens* have the chops to ride out this next one, if not on Earth, than somewhere beyond. She lays out a strong case for the role biological science will play in getting us out of this mess in her latest book, *Scatter, Adapt, and Remember: How Humans Will Survive a Mass Extinction*.

From out of the Permian-Triassic mass extinction, Newitz discovered a poster critter of resilience: *Lystrosaurus*, a little pig-like reptile that somehow survived those dark times, and became the dominant land creature on Pangaea for some 10 million years. “These guys were everywhere,” Newitz exclaimed. “They survived by adapting. These humble pig-dogs were willing to eat crap, which was about all that was left. Humans might survive, too. We are really good at adapting and eating crap.” Like those little survivors of 250 million years ago, Newitz suggested, humans are a successful invasive species.

To the researchers at the DOE JGI meeting, Newitz suggested that genome science might be an important part of that survival toolkit. “Synthetic biology is awesome,” she said. “An area we need to think about, if we want to survive as a species, is to make our cities function more

like an organism; how to make them more robust against a disaster.”

She listed alternative fuels, biofuels, solar fuels, and artificial photosynthesis as adaptations that could make cities more livable, more carbon neutral. On a more speculative note, she suggested that biological tools could be harnessed so we can grow future buildings from seeds: “We’re building our houses out of plants already. How do we improve that process?”

Given that we may have no control over an asteroid strike or a mega-volcano eruption of the kind that triggered the Permian-Triassic unpleasantness, Newitz suggested a Plan B: escape. She notes that current rocket technology just won’t get us to Mars in time to geo-engineer a new home planet; but the seemingly wacky idea of a 60-mile long “space-elevator” might be our ticket out. It’s a high-tech ladder out of our gravity

well. Once free of the bonds of Earth, the instincts of a successful invasive species should take us far.

Happily, Newitz believes we are going to make it through the next mass extinction. “We have all the characteristics of a survivor species,” she said. “We could end up underground dwellers eating bugs and moss, or we could think about how to build better cities. But we should take inspiration from how ecosystems function... We’re all going to end up evolving into something else.”

Watch Newitz’s keynote on the DOE JGI’s YouTube channel at http://bit.ly/JGIUM9_Newitz. 📺

— SR



More than 460 attendees heard speakers cover topics such as forest microbiomes, grasses for bioenergy, synthetic biology and microbes. Selected Meeting videos are available at http://bit.ly/JGI_UM9videos.



Annalee Newitz (David Gilbert)

Sprawling shrub's genomic gluttony

Off the coast of Australia lies a single South Pacific island, the only place in the world on which to find *Amborella trichopoda*, a shrub that is the lone member of its family and genus. In an epic demonstration of horizontal gene transfer detailed in the December 20, 2013 issue of the journal *Science*, *Amborella's* mitochondria, the plant's energy-producing organelles, have acquired six genome equivalents of foreign DNA. One genome came from a moss, three from green algae and two from other flowering plants. It is the first time that an organelle has captured entire "foreign" genomes, those from other organisms, and the first description of a land plant acquiring genes from green algae.

"It swallowed whole genomes from other plants and algae as well as retained them in remarkably whole forms for eons," said Indiana University's Jeffrey Palmer, senior author of the study. Along with colleagues from the DOE JGI, Penn State University, and the Institute of Research for Development in New Caledonia, he determined a remarkable expansion of the genome of the plant's critical energy-generating structures.

Amborella is one of the oldest flowering plants, having branched off from others about 200 million years ago. To collect samples for study, researchers had to make the long trek to the only place the plant grows: Grande Terre, the main island of New Caledonia, a small nation 750 miles off the eastern coast of Australia. Sequencing the entire mitochondrial genome, initiated under the DOE JGI's Community Sequencing Program, was an effort that took eight years and meant producing two independent draft assemblies that were then combined.

"In short, getting a complete assembly of the *Amborella* mitochondrial genome took years and the effort of several people," said Palmer. "But

it was well worth it, because the genome is so utterly unusual, amazing and full of surprises and mechanistic insights." The complete sequence was also necessary to rule out the possibility that any of the foreign DNA was the result of contamination or was integrated in the nucleus rather than the mitochondrion, he added.

The DNA that the *Amborella* mitochondria absorbed and retained through horizontal gene transfer amounts to at least one million base pairs, with its mitochondrial genome now swollen to an enormous size of 3.9 million base pairs. In comparison, a typical plant mitochondrial genome has just around half a million base pairs.

The DOE JGI Plant Program focuses on fundamental biology of photosynthesis, the conversion of solar to chemical energy, and the role of terrestrial plants and oceanic phytoplankton in global carbon cycling. Generating models such as the one cited in this study provides

an improved framework for further detailing the origin and mechanisms of mitochondrial fusion and gene transfer across eukaryotes — plants, fungi, algae. The transfer of genes through wounding, as described in the paper, not only applies to those plants that live in harmony with *Amborella* but also parasites. By extension, the group of plant parasites that that drain resources from *Amborella* are representatives of the largest group that sap flowering plants worldwide.

Another emphasis of the Plant Program is to understand healthy growth of candidate "biofuel crops" and the threats that plant pests present. The DOE JGI and its collaborators are building a compendium of gene annotation data sets that come to light through its collaborations and are made available through Phytozome at <http://bit.ly/Phytozome10> and deposited into the DOE's Systems Biology Knowledgebase (KBase) at <http://kbase.us/>. ☘



Amborella in fruit. (Jeffrey D. Palmer)

Sifting through data from 1st soil sequencing study

"It's one of the most diverse microbial habitats on Earth, yet we know surprisingly little about the identities and functions of the microbes inhabiting soil," said Jim Tiedje, Distinguished Professor at the Center for Microbial Ecology at Michigan State University.

Tiedje, along with MSU colleagues and collaborators from the DOE JGI and Lawrence Berkeley National Laboratory (Berkeley Lab), published the largest soil DNA sequencing effort to date in the March 10, 2014, issue of *Proceedings of the National Academy of Sciences (PNAS)*. "The Great Prairie represents the largest expanse of the world's most fertile soils, which makes it important as a reference site and for understanding the biological basis and ecosystem services of its microbial community," he said. "It sequesters the most carbon of any soil system in the U.S. and produces large amounts of biomass annually, which is key for biofuels, food security, and carbon sequestration. It's an ecosystem that parallels the large ocean gyres in its importance in the world's primary productivity and biogeochemical cycles."

In the pilot study launched by the DOE JGI, MSU researchers compared the microbial populations of different soils sampled from Midwestern corn fields, under continuous cultivation for 100 years, with those sourced from pristine expanses of the Great Prairie.

The DOE JGI provided the raw sequencing power for the project, generating nearly 400 billion letters of code, which amounts to more than 130 human genome equivalents. However, the lack of reference genomes that can serve as "Rosetta stones" to help sift through these data makes it difficult to identify genes involved in traits such as agricultural productivity, carbon cycling, or disease and drought resistance. The sheer scale of the data generation and analysis involved in the project led the study's lead author, MSU's C. Titus Brown, to liken traditional "shotgun" DNA sequencing of environment samples to "shredding the contents of an entire library and reassembling an individual volume out of that massive pile of shreds."

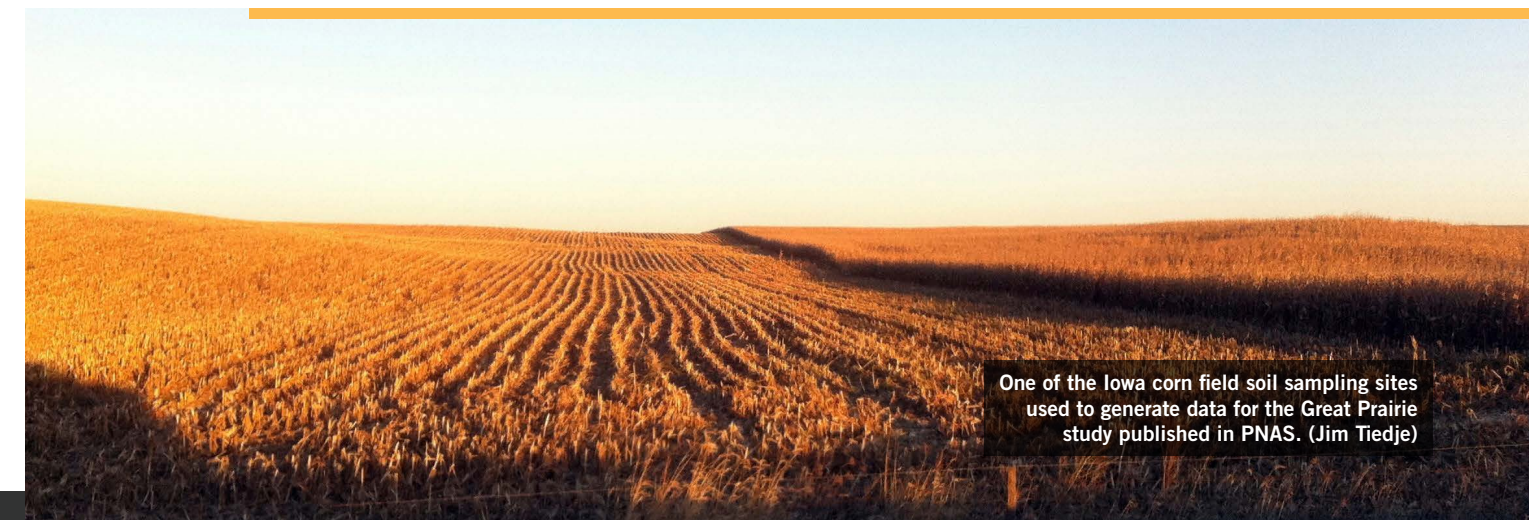
Brown and his colleague Adina Chuang Howe deployed a compression method called digital normalization that allows a substantial amount of data to be discarded without the actual data content being degraded. "These results still continue to stun me," said Brown of the results of his work on the soil dataset. "What this gives us is a 2 to 200-fold decrease in computational requirements for the actual biological analysis." He added that the process offers a data management "democratization" empowering scientists who don't have access to cloud- and high-performance computing, to analyze them.

"I think this can lead to a fundamental shift in thinking," Brown said. "We are actually converting standard, heavyweight approaches in biological sequence analysis to an ultra-efficient streaming approach." Consequently, researchers can devote more resources to extracting science from the noise, as their basic analysis expenditures have dropped.

Despite the staggering amount of sequence data generated, researchers say it's not enough to deeply characterize the microbial plays in the localized soil samples. Still, Janet Jansson, a senior staff scientist from Berkeley Lab's Earth Sciences Division who championed the project along with Susannah Tringe, head of the DOE JGI's Metagenome Program, isn't deterred by the prospect of more data to come.

"Metagenomic sequence analysis will help us establish patterns of how genes and organisms evolve in soil," said Jansson, "and how these can be used to understand and potentially manage adaptive traits such as greenhouse gas fluxes, carbon stability, and plant disease development. What we do know now is that soil microbes are responsible for cycling nutrients that are of critical importance for all higher forms of life."

For more information about C. Titus Brown's work, see the video, "Efficient streaming approaches to biological data analysis" at <http://youtu.be/rzScrmakgFA>. ☘



One of the Iowa corn field soil sampling sites used to generate data for the Great Prairie study published in *PNAS*. (Jim Tiedje)

Pond-dwelling powerhouse with biofuel potential
RINA SHAIKH-LESKO



Duckweed (Wenquin Wang)

Simple and primitive, a duckweed plant consists of a single small kidney-shaped leaf about the size of a pencil-top eraser that floats on the surface of the water with a few thin roots underwater. It grows in almost all geographic areas, at nearly any altitude, and ranks among the fastest growing plants, able to double its population in a couple of days under ideal conditions. All these and other properties make it an ideal candidate as a biofuel feedstock – a raw source for biofuel production. In particular, one trait that gives it an advantage over other candidate bioenergy crops is the plant's lack of lignin and cellulose. It has already been used to clean contaminated water and as a source to produce pharmaceuticals.

In a paper published February 19, 2014 in *Nature Communications*, researchers from Rutgers University, the DOE JGI, and several other facilities detailed the complete genome of Greater Duckweed (*Spirodela polyrhiza*) and analyzed it in comparison to several other plants, including rice and tomatoes.

S. polyrhiza turns out to have one of the smallest known plant genomes, at about 158 million base pairs and fewer than 20,000 protein-encoding genes. That's 27 percent fewer than *Arabidopsis thaliana* – which, until recently, was believed to be the smallest plant genome.

"The most surprising find was insight into the molecular basis for

genes involved in maturation – a forever-young lifestyle," said senior author Joachim Messing, director of the Waksman Institute of Microbiology at Rutgers University. Also intriguing to the research team were which genes were preserved over time and which were not. For example, despite the reduced number of total genes, *S. polyrhiza* has more copies of genes for enzymes involved in nitrogen absorption and metabolism than in other plants. This is probably linked to the plant's ability to utilize excess nitrogen in contaminated waters.

Messing estimates that duckweed will be a viable biofuel source within the next five years and points to Ceres Energy Group in New Jersey, which is already producing electricity from duckweed. "The sequencing of this genome opens new frontiers in the molecular biology of aquatic plants," he said. "This publication represents the single largest advance in this field and a new milestone in plant molecular biology and evolution, as previous studies were either classical botany or biochemistry of photosynthesis."

A video of Messing on the promise of duckweed can be seen at <http://youtu.be/PLVPfoKw2rs>.

Experimental evolution of biofuel bacteria



Quabbin reservoir (Philip Halling, Wikimedia Commons)

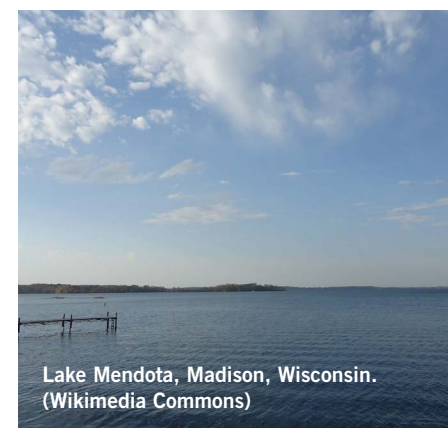
Breaking down cellulose and other woody plant materials is one of the biggest challenges of producing efficient and sustainable biofuels.

Biofuel developers depend on fungi and bacteria that are adept at breaking down these materials. *C. phytofermentans*, a soil-dwelling bacterium that decomposes leaf litter, is one of these bacteria of interest studied by researchers from the University of Massachusetts, Amherst. Their findings appeared January 22, 2014 in *PLOS ONE*. The bacterium's genome was sequenced under the aegis of the DOE JGI Community Science Program, which partially supported this study.

After letting bacteria isolated from the Quabbin Reservoir in Massachusetts grow for many generations – a technique called "experimental evolution" – on either cellulose, xylan or cellobiose, the team sequenced the populations of bacteria that had adapted to different living environments. Most mutations were found in regions of DNA that coded for or regulated genes involved with carbohydrate use or transport. The adapted bacteria have several diverse mutations, hinting that there are many ways to efficiently feed off of woody materials – and many molecular pathways that can be tweaked to break down woody materials more efficiently. - RSL

Lake-living bacteria surprisingly different from ocean relatives

Researchers are only beginning to understand the role SAR11 bacteria, which comprise half of all marine surface microbes, play in the global carbon cycle. One branch, or clade, of the SAR11 family grows in fresh water. Called LD12, the clade is difficult to study because researchers can't cultivate it in the lab. As reported in a paper published November 28, 2013 in the journal *Genome Biology*, a team of Swedish, German, and American researchers, including DOE JGI scientists, sequenced the genomes of 10 isolates of LD12 bacteria gathered



Lake Mendota, Madison, Wisconsin. (Wikimedia Commons)

at three lakes (Mendota and Sparkling Lakes in Wisconsin and Damariscotta Lake in Maine). They also took metagenome samples from all three lakes, as well as from three lakes in Sweden to complement the single-genome analysis. The single-cell genome sequences were generated at the DOE JGI.

The team was surprised to find that unlike marine SAR11 bacteria, the freshwater bacteria were much less genetically diverse. They suggested this is because LD12 bacteria probably had thinned out populations in their newly adapted environments (population bottlenecks) that reduced their genetic diversity. However, LD12 bacteria did retain hypervariable regions that researchers believe help SAR11 bacteria evade pathogens and predators. - RSL

Deciphering the virulence of a fern fungus
MASSIE SANTOS BALLON

As part of the work toward illuminating the fungal tree of life, a study published in the April 2014 issue of *New Phytologist* focused on a tiny plant pathogen called *Mixia osmundea*, a member of the Pucciniomycotina family that has had few species sequenced to date. Though it was first described 100 years ago, the fungus is rarely seen and has only been isolated on the fronds of two fern species in Japan, Taiwan and

America. The genome was sequenced and annotated and the data have been made available on the DOE JGI's fungal portal MycoCosm.

At the DOE JGI, fungal researchers are involved in several worldwide collaborations to learn more about fungi, and how they can be harnessed for roles such as improving the health of candidate biomass feedstocks for biofuels development. During the team's analysis of the fungus, they identified several carbohydrate-active enzymes that indicate *M. osmundea*'s capabilities in breaking down plant mass, as expected of the plant pathogen. They also found multiple copies of an enzyme that suggests the fungus is "especially efficient" at breaking down a particular compound in plant cell walls. However, they did not find genes that would indicate the fungus can break down xylan, or convert cellobiose into glucose. The report on the smallest fungal plant pathogen (13.6 million bases) to date provides researchers with insights into its mode of pathogenicity and reproductive biology. ☘



The fungal plant pathogen has been isolated from fronds of this fern. (Sturgis McKeever, Georgia Southern University, Bugwood.org CC BY-NC 3.0)

Genomics – in Poetry

In honor of National Poetry Month, we started #JGInPoetry. We need your help to build our poetry bank. One contributor, Daniel Levine, recently sent in the following poem:

*When it comes to debate
there's no room
Our love of the car will spell
"doom"
Unless crops we can turn
Into fuel we can burn
By using A-C-T-G to spell vroom*



Details at <http://bit.ly/JGInPoetry>

GWAS reveals promise of poplar biomass as a precursor for efficient biofuel production

LEVI GADYE



Jerry Tuskan (ORNL)

Lignocellulosic biomass, familiar to many as dry plant material such as wood, has long been used to provide humans with renewable energy in the form of fire. More recently, however, the use of this biomass for producing advanced biofuels has emerged as a carbon-neutral alternative to fossil fuel reliance. Despite the global abundance of lignocellulose in the cell walls of plants and trees, its resistance to chemical breakdown, or recalcitrance, presents a significant barrier for efficient biofuels production. Recent work, led in part by Gerald Tuskan at

Oak Ridge National Laboratory, has revealed details about cell wall biosynthesis that hold promise for identifying and breeding strains of *Populus trichocarpa*, the western black poplar, that will be more amenable for production of biofuel. Tuskan discussed the project during the 9th Annual Genomics of Energy & Environment Meeting.

In conjunction with colleagues from the University of British Columbia and West Virginia University, Tuskan set out to characterize genotypic and phenotypic variation amongst over a thousand poplar trees representing native populations found within a region spanning 15 degrees of latitude in the Pacific Northwest. Combining DNA sequencing of single-nucleotide polymorphisms, as well as RNA-sequencing data, the group generated a 'parts list' for genes involved in cell wall synthesis for each strain, and correlated variations in this biosynthetic pipeline.

This genome-wide association study revealed that variation in final lignin content was determined not by the terminal lignin biosynthetic pathway, but rather by upstream production of early lignin precursors – a logical means for preventing unnecessary cell wall synthesis well before significant

energy had already been invested in the synthesis of later-stage lignin precursors. "Our preconceived notions of how to modify cell wall chemistry seem wrong," Tuskan noted. A dominant-negative gene was found to be responsible for inhibition of early synthetic steps in low-lignin strains, and furthermore, this decrease in lignin content was functionally reflected in decreased recalcitrance of the poplar cell wall, and increased yield in subsequent biofuel production.

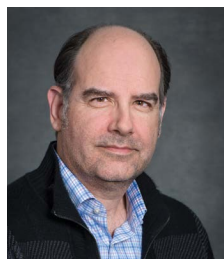
Though future work may tease out a more subtle association between lignin phenotypes and terminal steps in lignin synthesis, Tuskan is hopeful that these low-lignin poplar strains will soon lead to more efficient and economically viable biofuel production, as well as providing useful insights into regulation of cell wall biosynthesis. Low-lignin poplar variants might be expected to suffer in structural integrity, but thus far, these variants appear healthier than their high-lignin counterparts, suggesting that low lignin content may in fact be a beneficial trait in poplars. Given the rapid generation time and growth of poplars – over 20 feet in just three years - efficient, industrial biofuel production from poplars may soon be on the horizon.

Watch the video of Jerry Tuskan's talk at http://bit.ly/JGIUM9_Tuskan. 🌟

Levi Gadye is a neuroscience graduate student at the University of California, Berkeley who contributes to the Berkeley Science Review.

Kyripides elected to AAM

Congratulations to Prokaryote Super Program head Nikos Kyripides – one of 88 microbiologists elected a 2014 Fellow to the American Academy of Microbiology. He is now one of more than 2,400 Fellows representing all subspecialties of microbiology selected based on their records of scientific achievement and original contributions that have advanced microbiology.



Contact The Primer

David Gilbert, Editor
DEGilbert@lbl.gov

