JGI
DOE JOINT GENOME INSTITUTE
US DEPARTMENT OF ENERGY
OFFICE OF SCIENCE

the PRIMER

Fall 2009 Volume 6 Issue 2

FUNGI IN THE FOREFRONT

New Program, Encyclopedia

Fungi play key roles in DOE-relevant missions of bioenergy production, bioremediation and carbon cycling. In bioenergy projects alone, for example, fungal genome data have been used not only to ensure the health of crops that serve as biomass feedstocks but also provide enzymes that can break down the biomass as well as help convert it to transportation fuel.

The DOE JGI has developed a Fungal Genomics Program headed by Igor Grigoriev. The program's first project, launched October 1, is the Genome Encyclopedia of Fungi (GEF). The program aims to explore fungi's ecological diversity and breadth across the Tree of Life for DOE-relevant science and applications.

DOE JGI pioneered sequencing and analysis of several fungi important for lignocellulose degradation, enzyme production and plant health, and contributes about a quarter of the fungal genome sequencing worldwide. The formation of a Fungal Genomics Program provides an opportunity to present a unified voice that can articulate the fungal genomic community's needs.

"We now stand on the precipice of a golden age of mycology enabled in large part by sequenced based research resources generated by a formal Fungal Genomics Program," said

cont. on page 8

also in this issue

Setting the Standards

BY MASSIE SANTOS BALLON

First there was Bermuda; then there is Walnut Creek.

Nearly 14 years ago, representatives of all major human sequencing centers first convened on the island of Bermuda to set the standards for sequencing the human genome. This resulted in the endorsement of what has since been referred to as the "Bermuda standard," which included a nucleotide error rate of one or less in 10,000 bases and a long-term goal of no gaps in the sequence. Most importantly, the Bermuda meeting also defined new rapid data release rules, requiring the release of sequence assemblies at least one kilobase in length as soon as possible and mandating that no human sequences be patented.

Fast-forward to the last day of the Genomic Standards Consortium's 8th meeting, held September 9-11, 2009 at DOE Joint Genome Institute in Walnut Creek Calif., GSC head Dawn Field of the UK Centre for Ecology & Hydrology reeled off some of the meeting's accomplishments.

One involved plans to incorporate standards developed by the GSC into the collaborating international nucleotide sequences

cont. on page 4

GSC8 meeting participants

Predicting a microbe's traits by its genome

Nearly three-quarters of the Earth's surface is water, and the marine microorganisms that call it home are adapted to a range of life strategies. At one end of the spectrum are microbes that thrive in nutrient-rich waters, often associated with warmer regions, known as copiotrophs. At the other end are microbes that have adapted to nutrient-poor waters, or oligotrophs.

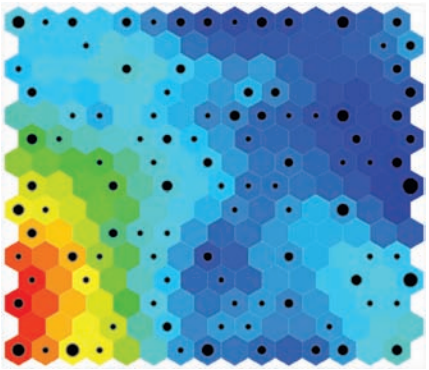
Using representative bacteria, an international team of scientists led by the University of New South Wales in Australia and the DOE JGI has developed a model that can predict whether a microbe is a copi-

image Copyright 2009 National Academy of Sciences, U.S.A.

cont. on page 2

Fall 2009 Volume 6, Issue 2

Microbial Lifestyle *cont. from page 1*



Two-dimensional map generated by University of New South Wales' Federico Lauro and Matthew Z. DeMaere shows clusters of organisms used in the study, with oligotrophs in blue, copiotrophs in red and other colors representing intermediate strategies. The distance between colors is proportional to how different they are from each other, while the size of each black dot is proportional to the number of genomes occupying a cell in the map.

otroph or an oligotroph based solely on its genome. The model allows researchers to get a better idea of the marine microbial diversity without cultivating and studying the microorganisms in a lab environment.

"Microorganisms in the oceans are subject to a broad range of selection pressures such as nutrient limitation, temperature and predation," said study senior author Rick Cavicchioli of the University of New South Wales. "The fact that trophic signatures can be gleaned from ocean microbes growing under diverse temperature conditions illustrates that this is an important selection criteria. Now that we know 'who' is present, and what their 'lifestyle' is, we can ask, what exactly are they doing — assessing function."

For their model, Cavicchioli and his lab selected representative oligotrophic and copiotrophic bacteria found in the waters off Resurrection Bay, Alaska and Sydney, Australia, respectively, and compared their genomic traits. *Sphingopyxis alaskensis* thrives in low-nutrient waters and is key to carbon sequestration in the oceans, factors that led to its genome being sequenced by the DOE JGI as part of the DOE's plans to understand the roles such organisms play in the global carbon cycle. *Photobacterium angustum* on the other hand, is a copiotroph from the warmer, nutrient-rich waters off Australia.

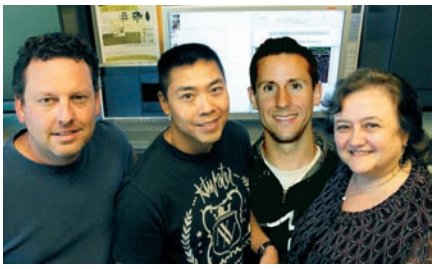
The researchers then used that information to successfully predict whether several dozen bacterial samples were copiotrophs like *P. angustum* or oligotrophs like *S. alaskensis*. Their findings supported a long-held theory that though more microbial genomic projects involve easier-to-cultivate copiotrophs, bacteria like *S. alaskensis* are actually more representative of the majority of microorganisms in the biosphere.

DOE JGI Genome Biology Program head and study co-author Nikos Kyrpides said the results accent the need to develop better techniques such as single-cell sequencing to isolate and sequence oligotrophic genomes in order to better reflect the microbial biodiversity in the biosphere. "To sequence *cont. on page 8*

Mending the Gaps

Over the last year, the DOE JGI has modified its sequencing pipeline to take advantage of the benefits next generation DNA sequencing technologies have to offer over traditional Sanger sequencing. Currently, standard 454 Titanium and paired end 454 Titanium data are generated for all microbial projects and then assembled using the Newbler genome assembler. This allows more efficient production of high-quality draft assemblies at a much greater throughput. However, it also presents new challenges—with increased throughput, comes a larger number of gaps in the Newbler genome assemblies. Gaps in these assemblies are usually caused by repeats (Newbler collapses repeat copies into individual contigs, thus creating gaps), strong secondary structures, and artifacts of the PCR process (specific to 454 paired end libraries).

To expedite gap closure and assembly improvement on the growing inventory of these assemblies, a team at the DOE JGI led by Alla Lapidus has developed software, "Gap Resolution," to address this issue. The code was written by Stephan Trong, Brian Foster, and Kurt LaButti with valuable contributions by Tom Brettin and Cliff Han. The software performs the following steps:



"Gap Resolution" development team, Brian Foster, Stephan Trong, Kurt LaButti and Alla Lapidus.

1. Identifies and distributes the reads for each captured gap into sub-projects based on read pairing information.
2. Assembles the reads associated with each sub-project using a secondary assembler, such as Newbler or PGA, then validates assembly using anchor sequences.
3. Determines whether any gaps are closed after reassembly, and either designs fakes (consensus of closed gap) for those that closed or lab experiments for those that require additional data.

Those interested in acquiring this software, please contact degilbert@lbl.gov.

the PRIMER / 3
Fall 2009 Volume 6, Issue 2

IDing Fungal Mutations to Boost Biofuel Production

During World War II, the fungus *Trichoderma reesei* was the bane of the American Army quartermasters stationed in the South Pacific. Now the same fungus is a key producer of industrial enzymes that are used, among other applications, to break down biomass for biofuel production.

The reversal in *T. reesei's* fortunes is due to the scientists who developed strains with increased production of enzymes that can break down biomass using a variety of mutagens. The problem with the mutagenic treatments, however, said DOE JGI scientist Wendy Schackwitz, is that "it's not known which of the many induced mutations are responsible for this increased production."

To help solve the problem, Schackwitz, her colleagues at the DOE Joint Genome Institute (JGI), the French applied research center IFP and the Vienna University of Technology (TU Vienna) collaborated on the first genome-wide map of *T. reesei* mutations. The current study complements last year's publication of the *T. reesei* genome, which was sequenced at the DOE JGI using the reference strain named for the Army quartermasters, QM6a.

"With information from this study, you can begin to understand which mutations are involved in boosting cellulase production and which are just baggage," Schackwitz said. She, her fellow co-first author Stéphane Le Crom from the French institute École Normale Supérieure and their colleagues mapped the mutations found on two hyperproducing strains of *T. reesei,* NG14 and its direct descendant RUT C30, using Illumina next generation sequencing technology. Unlike Le Crom and his French colleagues, who did single end reads on the RUT C30 and NG14 isolates, Schackwitz and her DOE JGI colleagues did paired ends reads on the other RUT C30 isolate. By working with short DNA sequences on either end of an unsequenced DNA fragment of known size, the DOE JGI researchers were able to narrow down the lists of possible areas where reads might align. The information also allowed them to determine if very short DNA sequences called indels had been inserted or deleted.

Left, DOE JGI scientist Wendy Schackwitz; right, DOE JGI and PNNL scientist Scott Baker

The mutation map allows researchers to study the mutations and identify each one's effect on the *T. reesei* strains, said Schackwitz. Scott Baker, a DOE JGI scientist at Pacific Northwest National Laboratory and senior author of the study along with Christian Kubicek of TU Vienna and Antoine Margeot of IFP added, "We now have a blueprint on which we can do future studies to see which genes are related to the enzymes. If you can produce more enzyme more efficiently, that makes your process — in this case the production of biofuel — more economical."

image courtesy of Novozymes

Fermenter

Study co-author Randy Berka, a director at the Davis, Calif.-based office of the Danish bioinnovation company Novozymes, one of the largest producers of industrial enzymes, confirmed *T. reesei's* importance for biotechnical applications. "Most, if not all of the *T. reesei* strains that are used to produce cellulases today for industrial applications were derived from the ancestral QM6a isolate and its progeny," he said. "Companies have devised ways to generate improved strains from the QM6a pedigree that produce cellulase enzyme products more economically."

His colleague, Novozymes staff scientist Michael Rey, added that identifying how each mutation affects the fungal strains could lead to healthier, hyperproducing *T. reesei* strains. "We can learn a great deal about the components within *T. reesei* cells that might be further tweaked to make strains with higher productivities which translates into better economy. This could be critical in developing *T. reesei* strains that produce enzymes cheaply enough for demanding applications such as cellulosic ethanol," he said.

The paper was published online the week of August 31 in the journal *Proceedings of the National Academy of Sciences* Online Early Edition. Hear Scott Baker talk about *T. reesei* at http://www.jgi.doe.gov/News/finalTreeeipodcast090109.mp3.

Fall 2009 Volume 6, Issue 2

GSC8 meeting *cont. from page 1*

databases (INSDC) containing genomes, metagenomes and ribosomal sequences for use by the community. Another involved an agreement between three GSC board members who oversee two data management systems to test whether or not their systems can exchange information with each other in the hopes of building a new collaborative computing platform. A third collaboration involved plans to work with large sequencing projects such as the Human Microbiome Project and the Terragenome Project to see the standards developed by the GSC integrated and implemented by more members of the genomic community.

This list of planned collaborations seemingly marks the GSC's further transition from an informal gathering of genomics researchers concerned about establishing standards for collecting and capturing data on genomic projects to a broader community ready to introduce and integrate their ideas to the genomic community at large.

"This may sound like a UN speech but it's common sense and understanding now within the GSC community that's changing everything on how we move forward," said GSC board member and DOE JGI Genome Biology head Nikos Kyrpides.

The three-day gathering covered not just updates to GSC-driven projects such as developing and disseminating minimum specifications to be electronically entered for genome and metagenome projects and the first issue of the official GSC open access online publication *Standards in Genomic Sciences (SIGS)*, which introduced two novel types of articles based on GSC initiatives, but also input from outside groups who will be asked to adhere to the standards as they move forward on their projects.

First-time meeting attendee George Weinstock, Associate Director of the Genome Center at Washington University, said that though the GSC's work on developing standards was very important and its progress well thought-out, the information wasn't yet getting out. "The message now should be, 'We're here and we've matured' and start reaching out," he said.

GSC8 workshop co-organizer Dawn Field

Weinstock was invited to GSC8 to talk about the Human Microbiome Project, which he noted was one of the projects that should be engaged further by the GSC. Made possible by the advent of next-generation sequencing technologies, the project involves collecting 12,000 samples from the human body and sequencing the microbial communities found in each one, as well as sequencing at least 900 individual bacterial reference genomes and additional virus and eukaryotic microbial genomes. Weinstock said some samples would likely generate 10 Gigabases of information, and in its entirety the project will produce beyond the equivalent of a human genome at 30X coverage. In order to meet the project's computing needs, Washington University has invested in a data center that can handle several petabytes (1 petabyte = 1 billion megabytes).

Not all sequencing centers are similarly equipped to ramp up their data centers to keep up with genome projects' increasing computing needs, noted Folker Meyer, a GSC Board Member and Associate Division Director of the Institute of Genomics and Systems Biology at Argonne National Laboratory, To help ease the bottleneck, in July, after a GSC Special Interest Group meeting in Stockholm, Meyer and other researchers within the GSC created a group called M5, which is working toward gaining the genomic community access to large-scale computing centers such as Oak Ridge National Laboratory, Argonne National Laboratory and the National Energy Research Scientific Computing Center.

"There are computers out there that can solve all our problems in a minute, but we can't get access to them because we can't communicate our needs," Meyer said, noting that large-scale computing power is traditionally associated with the physics community, which deals with petabytes of information, typically from just a handful of sources. "We need to work together to 'speak with one voice' to be able to ask for solutions to real problems."

While genome projects may not yet be on the order of petabyte or two, the Terragenome project, "Tera Terra", under which umbrella the DOE JGI is leading the sequencing of metagenome samples from

Other GSC8 Highlights:

- **Authoritative database INSDC will include a GSC: MIGS, MIMS, MIENS keyword in their records and the completion of the MIGS 2.1 checklist, now including MIENS**
- **Idea of a Global Census of Microbes to be led by DOE JGI's Nikos Kyrpides launched**
- **Agreement to work on gene calling standards for publication in SIGS. The effort is led by DOE JGI's Nikos Kyrpides but involves all the genome annotation groups present at the meeting.**

Midwestern prairies to learn more about the microbial communities in the soil, is expected to generate a Terabase (1 million Mb) of sequence.

Next generation sequencing has increased the genomic data stream to make such large genome projects possible, but it also means researchers need more computing power. Meyer said the genomic community needs to come up with a common language to articulate their needs to large scale computing facilities. As a first "baby step," Meyer and fellow GSC Board members Kyrpides and Berkeley Lab's Biological Data Management and Technology Center head Victor Markowitz are setting up a test to find out whether their respective data management systems, MG-RAST and IMG, can exchange information. A second, related project calls for all those working on metagenome projects to develop a common workflow language and a data file exchange standard in addition to the successful exchange standard that has already been agreed for descriptive information about the genomes and metagenomes themselves.

"I believe that one of the greatest achievements of the GSC could be the M5 initiative, which, if it succeeds, will create the badly needed culture of collaboration, open exchange of methods, approaches and ideas and spirit of working together for common goals," said Kyrpides.

Beyond seeing standards for electronically capturing and collecting data on large genome projects, the meeting also featured speakers concerned with standardizing the finishing and annotation pipelines.

DOE JGI's Patrick Chain presented a proposal to add several new levels of finishing standards to fit all sequencing projects. By the end of the meeting, the GSC Board had endorsed this activity and invited this group of experts to work with this GSC as part of its wider consensus-building efforts.

Presenting a concise version of his keynote speech during the "Sequencing, Finishing, Analysis in the Future" conference held earlier this year, University of Maryland's Owen White also gave a talk proposing the idea of consensus annotation.

Before the meeting ended, Field raised the possibility of leaving a few seats open to interested newcomers at the next meeting, to be held in spring of 2010. Though the GSC has always welcomed new members and has progressively grown, for the

cont. on page 8

GSC members reel off acronyms to describe projects and groups with ease but those unfamiliar with the organization may find the following list helpful.

Glossary of Terms

CAMERA: Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis.
CBoL: Consortium for the Barcode of Life.
DDBJ: DNA Database of Japan.
EBI: European Bioinformatics Institute, a non–profit academic organization under EMBL.
EMBL: European Molecular Biology Laboratory.
EML: Ecological Metadata Language.
ENA: European Nucleotide Archive.
EnvO: Environment Ontology Consortium.
GCDML: Genomic Contextual Data Markup Language implemented using XML.
GEM Catalogue: GEnomes and Metagenome Catalogue, a centralized database of rich sets of metadata about the complete genome and metagenome collection.
GSC: Genomic Standards Consortium.
ICoMM: International Census of Marine Microbes.
INSDC: International Nucleotide Sequence Databases Collaboration composed of DDBJ/EMBL/GenBank.
IMG: Integrated Microbial Genomes system developed and maintained at the DOE JGI.
ISA: Investigation/Study/Assay.
ISA-Tab: format to support the submission and exchange of the metadata and to render, in an effective way other XML-based metadata in user-friendly readable, common format.
M3: Metagenomics, Metadata and Metaanalysis.
M5: Metagenomics, Metadata, Metaanalysis, Models and MetaInfrastructure.
MG-RAST: Metagenome Rapid Annotation using Subsystem Technology, a fully-automated service for annotating metagenome samples.
MIENS: Minimum Information about an ENvironmental Sequence.
MIGS: Minimum Information about a Genome Sequence.
MIMS: Minimum Information about a Metagenome Sequence.
NCBI: U.S. National Center for Biotechnology Information.
NERC: Natural Environment Research Council, the UK's main agency for funding and managing research, training and knowledge exchange in the environmental sciences.
OMICS: peer-reviewed journal that covers work in "all the fields ending in '-omics'."
RCN4GSC: Research Coordination Network for the Genomic Standards Consortium that just received a five-year, $500,000 grant from the U.S. National Science Foundation to fund GSC workshops and collaborations.
SIGS: Standards in Genomic Sciences (www.standardsingenomics.org), open-access publication of the GSC established to share genome and metagenome reports in compliance with MIGS/MIMS standards.
SRA: Short Read Archive.
XML: Extensible Markup Language rules for encoding documents electronically.

Fall 2009 Volume 6, Issue 2

Extra Chromosomes for Extra Habitat Diversity

A fungus that affects several dozen crops has been sequenced and analyzed by an international team, including scientists from the DOE JGI and the University of Arizona and other institutions, and published in the August 28 issue of the journal *PLoS Genetics.*

Nectria haematococca Mating Population (MP)VI, is also known by the asexual name *Fusarium solani,* and the fungus is the most extensively studied member of the *F. solani* fungal species complex. The published sequence of *N. haematococca* is expected to help not just agricultural researchers who can work on increasing fungal tolerance to avoid declining crop yields, but bioenergy researchers working on more robust plant biofuel feedstocks.

"Understanding plant pathogens, mechanisms of pathogenicity and their control is essential for obtaining a sustainable growth of biofuel feedstock," said DOE JGI's Igor Grigoriev, one of the study's senior authors.

Like *N. haematococca,* other members of the *F. solani* species complex have been found in a wide range of habitats, ranging from crops such as legumes to humans with weakened immune systems to the radioactive sections of the damaged Chernobyl nuclear reactor. The *F. solani* species complex's ability to thrive in such diverse environments also means the fungi are difficult to remove once they've set it. Ongoing efforts to eradicate the species complex from the cave of Lascaux, France without affecting the prehistoric wall paintings, for example, have so far proved unsuccessful.

"The current research found that the *N. haematococca* fungus has an unusually large genome (54 megabases)," said University of Arizona plant scientist and study senior author Hans VanEtten. "This large genome most likely helps explain

image courtesy of LNPV UMAF

the *F. solani* species complex's ability to thrive in such diverse habitats, infecting more than 100 different plants."

One reason *N. haematococca* MPVI has such a large genome is that it has "extra" or supernumerary chromosomes which increases the number of genes it has by about a thousand. The extra chromosomes are also known as "conditionally dispensable" (CD) chromosomes because they contains genes not needed for growth in the laboratory but needed for different biological habitats.

"*Nectria* is a plant pathogen whose pathogenicity is linked to genes located on a conditionally dispensable (CD) chromosome," Grigoriev said. "Losing this chromosome leads to its inability to invade a plant host. Through genome sequencing we not only decoded a genetic footprint of this chromosome but found two more chromosomes with similar properties that were then confirmed as dispensable in other strains."

VanEtten noted that based on the genetic information obtained from sequencing the fungal genome, all three CD chromosomes might have come from another organism through horizontal transfer. The process may not be a common occurrence, but it might explain why both the *N. haematococca* MPVI fungus and the *F. solani* species complex in general can thrive in so many habitats and prove resistant to compounds normally toxic to fungi such as antibiotics and heavy metals.

Since other studies have suggested that *N. haematococca* isn't the only fungus to have acquired genes via horizontal gene transfer, vanEtten and his colleagues plan further genetic investigations.

New DOE JGI Informatics Department Head

Svilen Tzonev (right) joined the DOE JGI in early August with a decade of experience managing bioinformatics and informatics technology departments in industry and an in-depth knowledge of next-generating sequencing analysis systems.

Tzonev most recently served as senior director of Software and Bioinformatics with Hayward, Calif.-based Illumina, where he was involved in developing and commercializing the Genome Analyzer system. He also served in a similar capacity at Solexa and Amersham (now GE Healthcare).

As head of the Informatics department, Tzonev is in charge of defining, assessing, implementing and enhancing the DOE JGI's informatics infrastructure plan to ensure, he said, that the user facility maintains its position on the frontiers of science while focusing on projects related to the U.S. Department of Energy missions of bioenergy, carbon cycling and the environment.

Tzonev has a Ph.D. in physics form the University of Illinois at Urbana-Champaign and did his postdoctoral studies at the University of California at San Francisco Medical School.

the PRIMER / 7
Fall 2009 Volume 6, Issue 2

HudsonAlpha's Rick Myers

For 15 years, Richard Myers (left) was a genetics professor at the Stanford University School of Medicine and director of the Stanford Human Genome Center. Last year, in a cross-country move that was also a homecoming, he left California for Alabama to take on the role of president and director of HudsonAlpha Institute for Biotechnology in Huntsville.

While at Stanford, Myers worked with the DOE JGI to sequence part of the human genome. Now Myers' team at HudsonAlpha works with the DOE JGI on closing the gaps in non-human eukaryotic projects (plants, fungi, etc.) and has sequenced more than 40 organisms related to bioenergy, agriculture and the environment.

In a profile done by GenomeWeb, Myers described his own style of overseeing the HudsonAlpha team and what he looks for in potential new members. The Myers lab at Stanford had more than 50 people working on a number of projects. Myers sees HudsonAlpha as a bigger lab and working on keeping the atmosphere pleasant for everyone.

Myers also serves on the DOE's Biology and Biotechnology Program Advisory Committee, the HapMap Advisory Committee and the Review Group for Large-Scale DNA Sequencing Centers of the National Human Genome Research Institute.
http://www.genomeweb.com/sequencing/leading-large-lab

photo by Gary Meek, Georgia Tech

Finding the Right Cleaner for the Job

Areas contaminated with high levels of metals and compounds such as uranium, sulfur, iron and nitrates can be harmful to humans, but to *Shewanella* bacteria, such places are smorgasbords. The microbes take in these compounds for energy and then expel them in an altered, less toxic form. This ability highlights *Shewanella's* importance for bioremediation as the bacteria can both limit the size of the contaminated area as well as clean it up.

Collaborating with researchers at the Pacific Northwest National Laboratory (PNNL) who are part of the *Shewanella* Federation, a consortium involving researchers from academia, national labs and industry under the DOE's Genomics: GTL, the DOE JGI has sequenced nearly two dozen Shewanella bacteria.

In a study published online the week of August 31 in the journal *Proceedings of the National Academy of Sciences Early Edition,* researchers from the Georgia Institute of Technology, Michigan State University and PNNL used 10 *Shewanella* strains to complete the first system-level assessment of these bacteria.

The study is expected to help direct researchers into selecting the right strain of *Shewanella* for future bioremediation projects based on the conditions in the area. Under the microscope, one strain of *Shewanella* may look identical to another, but when Kostas Konstantinidis (bottom left) and his colleagues compared the microbial genomes, the researchers found that while some of the strains shared 98 percent of the same genes, other strains only shared 70 percent. Additionally, nearly half of the protein-coding genes were strain-specific.

The differences reflected adaptations to specific environmental and ecological conditions, which led to Konstantinidis' suggestion that the species classification shift from traditional methods to a more genomics-based approach.
http://www.gatech.edu/newsroom/release.html?id=3310

photo by Nicholas Benner, Mizzou Magazine, University of Missouri

Odor Alert: Bacteria at Work

South of the Grand Canyon lies Arizona's Lost Orphan Mine, which produced uranium until the 1960s. The location makes it a natural tourist stop, but the nearby creek has been contaminated by radioactive residues. As a result, hikers passing through the area are warned not to drink the water, and to avoid the entire mining operation.

If University of Missouri biochemist and DOE JGI collaborator Judy Wall (above) has her way, however, unsuspecting people who pass by areas like the Lost Orphan Mine may find themselves steering well clear of these sites thanks to stinky bacteria that could warn the unwary while providing a low-cost, low-maintenance method of cleaning up contaminants.

Wall works with *Desulfovibrio*, anaerobic bacteria that can corrode iron and steel pipes and "sour" petroleum. The DOE JGI has sequenced more than a dozen strains of *Desulfovibrio* bacteria, including one proposed by Wall called *Desulfovibrio desulfuricans*, a strain that feasts on uranium and chromium.

Desulfovibrio bacteria can also degrade sulfates and toxic metals. And they give off a distinctive, rotten-egg odor when they're eating sulfates for energy. Wall said working with the microbes in airless conditions can be daunting to all but "the most stubborn and committed."
http://cafnr.missouri.edu/news/stories2009/uranium.php?utm_source=web&utm_medium=banner&utm_campaign=homepage

Smells like Sweaty Spirits

photo courtesy of North Carolina State University

Considered a winemaker's worst enemy, a yeast species that contaminates the entire fermentation process produces a characteristic odor most associate with horse sweat. Trevor Phister (left) of North Carolina State University says the yeast *Dekkera bruxellensis,* known to wine enthusiasts as

cont. on page 8

Fall 2009 Volume 6, Issue 2

Fungi *cont. from page 1*

the program coordination team, Grigoriev, DOE JGI's Alla Lapidus, and Scott Baker, a DOE JGI researcher at Pacific Northwest National Laboratory in a statement.

Baker noted in his paper last year the need for a Fungal Genomics Program: "Fungal genomes deserve more attention given the scientific and economic benefits derived from studying them."

Potential collaborators interested in working on projects under the Fungal Genomics Program will be solicited through a fungal-specific Community Sequencing Program application.

Microbial Lifestyle *cont. from page 2*

microbial genomes that are representative of the environments in which they were collected and for a more systematic and comprehensive sampling of the Tree of Life, researchers need to increasingly develop and rely on other techniques such as single-cell sequencing," he said.

His words were echoed by DOE JGI collaborator Ramunas Stepanauskas of the Bigelow Laboratory for Ocean Sciences, who said the model developed by Cavicchioli's team supports his own group's use of cultivation-independent, single-cell sequencing of marine microbes.

Cavicchioli said the researchers plan to apply the model to studying the microbial communities' relationships with climate change, such as the end of phytoplankton bloom, which is believed to be relatively nutrient rich even in the absence of a chlorophyll spike. "By analyzing and comparing the strategies of the dominant organisms we should have an idea of the carbon flux going through the environment which in turn will be affected by global warming," he said.

The work by Cavicchioli and his colleagues is detailed on the September 14 cover of the journal *Proceedings of the National Academy of Sciences.*

GSC8 meeting *cont. from page 5*

past four years, the approximately 50-person gathering has been by-invitation-only, in part, Field explained afterward, to maintain a core group that could keep projects moving while soliciting new perspectives. By opening the conference in the future, the GSC hopes to reach out to even more of the genomic community and develop further collaborative efforts while also building a sustainable model for continuation of GSC activities. Unsurprisingly, the idea met with approval.

To learn more about the Genomic Standards Consortium, go to their wiki page at http://gensc.org/. To see talks from the GSC8, go to http://www.scivee.tv/node/12786.

News *cont. from page 7*

Brettanomyces custersii or simply brett, is sneaky. As he explained in *Wine Spectator, Dekkera* has a tolerance for high alcohol levels, which allows it to wait for brewer's yeast (*Saccharomyces*) to get rid of other microbes in the winemaking process before stepping in to usurp the other yeast's position to infect and spoil the entire batch of wine.

To understand how *Dekkera* can spread so quickly, and develop methods of controlling it, Phister is working with the DOE JGI under CSP 2010 to sequence the yeast's genome. The DOE is interested in *Dekkera* because biofuel production involves converting organic matter into alcohols such as ethanol and the yeast can significantly reduce the yield.

Sequencing *Dekkera* could also provide biofuel producers with information to develop yeast strains that could more efficiently handle lignocellulose fermentations, making the biofuel production process more cost-effective. http://www.winespectator.com/webfeature/show/id/40447

Contact The Primer
David Gilbert, Editor / DEGilbert@lbl.gov / (925) 296-5643